Path Calculations and Option Pricing

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

By

Min Wang

Department of Mathematics

University of Leicester

December 2019

Abstract

The thesis is worked in the areas of the intersection of probability, combinatorics and analytical combinatoric. The research is motivated from the need of producing new methodologies and financial models in global market resulted from the lesson of 2007-2009 global financial market and a quantum tool called Feynman path integral method which has been applied to model path-dependent option pricing model by Hao and Utev. Path calculation method deal with models by analysing each possible individual asset price paths which broaden the methodology of modelling financial market and can solve some unusual or complex models which is difficult to model by using non path-dependent calculation method.

My research has focused on developing combinatorial structure and path calculation methods and then apply them to model individual share price path and calculate option prices. The share price can be modelled as a path with a given share price changes and the expiry date. We have applied Flajolet symbolic method, generating functions and path calculation method to model a set of typical finite restricted share price paths with restriction not allowing k consecutive down steps and derived a calculation of option prices in the model. Besides, applying the Flajolet symbolic method, we constructed a relationship between individual share price and generating function, analysed the transformed share price paths via different operations on generating functions. In addition, we applied path calculation method to solve winning probability in the classical gambler ruin problem which contributes the same result as the solution solved by establishing the recurrence equation method. Furthermore, we have solved a different gambler's ruin problem using the path calculation method which cannot be solved by the recurrence equation method.

Counting paths with combinatoric can be studied from two ways, one way is to label and the other is computation. Labelling is a part of representation of objects. We have developed a graphical theoretical construction of individual share price path via general binary trees and matroid. In addition, We have developed a method to solve some kinds of pattern avoiding path counting combinatorial problem by modifying certain probability methods. Two working papers including modelling of paths via matroids and counting via Markov-type technique is now being produced.

Acknowledgements

Undertaking the PhD mathematics has been an evaluable life-changing experience for me. There are many people who have helped me as a mathematician student over the last four years. It is pleasure to acknowledge the guidance and support I have received from many people.

Firstly, I praise the Lord Jesus has chosen me to know him in my first year of PhD, with his almighty for guiding and providing me the wisdom, knowledge, strength, blessing and opportunity to complete the study. It is life-changing for me by knowing him and his words. His words is a lamp for my feet, a light on my path.

I would like to express my utmost and sincere gratitude to my supervisor, Professor Sergey Utev, for his instructions, time, patience, motivation, inspiration, encouragement and immense knowledge I received throughout my study and research work. I would like to thank him for encouraging me to start the PhD, guided me to do the research, pass viva and amend thesis. I feel very fortunate to have been a PhD student supported and guided by him. He not only supported me for the study but also encouraged me many times in general, without his guidance and constant feedback this PhD would not have been achievable.

I would like to thank the thesis committee: Dr Bogdan Grechuk(University of Leicester), Dr Fraser Daly(Heriot-Watt University), Professor Sergei Petrovskii(University of Leicester) for their insightful comments and time. I want to thank Dr Bogdan Grechuk with his kind help and useful advice that helped in addressing the shortcomings in the thesis.

I would like to thank all of the staff members in the Department of Mathematics at University of Leicester for their help and providing convenient environment throughout my research study. I want to thank Charlotte Langley for her kind help.

I am very much indebted to my parents for their unconditional love, endless prayer, support and words of encouragement. It would not be possible without them to complete the thesis and succeed.

Last but not least, I am very grateful to all my friends in Michael Atiyah building, Stephen, Wendy, Pattharaporn(Aon), Simran, Eliyas, Nur, Ka Kin, Num, Jehan, Areej et al. for their support and friendship. I want also give thanks to Xiaoyu Yu, Siqi Wang, Nan Yan, Phoenix, Chu Xiao, Hana, Haige Yuan, Yixin Xie, Raojun Li, Shuai Zhao, Rui Luo et al. for their friendship in my study and usual life.

Contents

List of Figures

Chapter 1

Introduction

The lessons from 2007-09 global financial crisis leads to needs of reasonable responses in financial sector and the real economy [1] [2]. New methodologies, financial models need to be designed in the global market [1] [2]. More volatility of asset prices in financial market during the crisis bring in excessive volatility of option prices which impose the challenges to option pricing model [3].

Option pricing approach and model is an active research area in financial market, started from the well-known Black Scholes model, which calculate theoretical European option price with certain assumptions [4]. One of the limits of the model is that the model does not cover the pricing of American option which could be exercised before the expiration date [7]. On the other hand, Cox-Ross-Rubinstein Binomial Model provides a discrete teaching model to continuous model, which was first proposed by Cox, Ross and Rubinstein and then become a widely used model in the literature [5]. The binomial model adopts an iterative procedure and can calculate option prices based on the decisions at each period before the expiration date[5].

Quantum mechanics approaches can provide a way to study the behaviour of unpredictable stock market [2][8]. Use quantum tools to do financial modelling initiated in Ma and Utev [9] and developed in Karadeniz and Utev [10].

One of the famous quantum tools, Feynman path integral quantum mechanic approach can be applied to stock option pricing [8]. The Dirac-Feynman quantum mechanics technique for insurance risk modelling was initiated and developed in Tamturk and Utev [11]. The quantum approach is adapted to formalise the path-dependent option pricing [2].

Two simple quantum non-life risk models was studied in [11, section 2], in the quantum risk models, the claim amounts are assumed to have two point distributions and the observed data are treated as having small claims u and significant claims d. The two quantum risk models study the cumulative claim amounts collected at regular time intervals where each interval was set up reasonable small to have a maximum two claims [11]. One quantum risk model studies cumulative claim amounts which assume repeated claims was not observed, in the case of interval which allow maximum two claims, it is a model not allowing two small claims $2d$

and two significant claims $2u$ [11].

Motivated from the interesting quantum risk models in [11, section 2], in the thesis, the main idea in the dissertation is to consider restricted paths, the typical restricted path we studied is the share price paths with restriction not allowing 2 consecutive down steps.

Motivated from the Feynman path integral method [12] [13], in the thesis, our approach to finance is based on path calculation. The path calculation idea is to identify all possible truncated individual paths in the model, associate each possible path with a probability and then sum up the probabilities of the valid paths according to different truncate time and different arriving time to finally get the required probability in the model.

This dissertation presents variety of discrete time models for the financial stock market.

The share price is modelled as a path

$$
\begin{aligned}
\sigma_1 &= S_0 \to S_1 = S_0 y_1 \to s_2 = S_1 y_2 \\
&\to \ \ldots \ \ S_{T-1} = S_{T-2} y_{T-1} \to S_T = S_{T-1} y_T,
\end{aligned}
$$

where y_i are the share price changes, T is the expiry day. In the classical Cox-Rubinstein Binomial model [5] [6], for Black-Scholes approach to option pricing [4], $y_1 \in \{u, d\}$ where u is the jump up and d is the jump down.

The path calculation approach implies that

$$
E(f(S_T)) = \sum_{\sigma} P(\sigma) \cdot f(S_T(\sigma)), \quad \text{where} \quad \sum_{\sigma} P(\sigma) = 1.
$$

The main idea in the dissertation is to consider restricted paths. More exactly, a typical example is that the share price is modelled by a restricted two step binomial model with restriction not allowing two consecutive down steps, the model is also a restricted two step binomial model with restriction that the number of maximum consecutive down steps is 1.

Various models are constructed in this way. In chapter 2, several examples of option pricing using path calculation method in binomial model was introduced. In chapter 3, section 3.4.3, motivated by Flajolet symbolic method ([15, ?]) and apply the idea taken from [16] [14], we use the Flajoet symbolic method to give an more constructive explanation and solution using symbolic method to solve general bitstrings not allowing 3 consecutive 0's less than 3. In chapter 4.5, an example of path calculation using combinatorial method not touching the given bold line segment is derived and techniques is stated in the example. In chapter 6 motivated by Flajolet symbolic method and generating function approach[14] [15], and motivated by quantum path calculation method in [2], we calculated the option price in the finite restricted binomial model. Having different ways of counting is

necessary to do path counting. In chapter 8, we developed a method of counting restricted share price paths via markov-type stochastic modelling.

Moreover, counting paths with combinatoric can be studied from two ways, one way is to label and the other is computation. Labelling is a part of representation of objects and generating function can provide many different ways of representing financial paths.

Like reflection problem in random walk, clearly, path calculation depends on path representation. Different way of path representation have been studied. Specifically, In chapter 5, motivated from the exercise in [17], we developed the algorithm of constructing share price via binomial trees and analysed the modelling of share price paths via binomial trees. Furthermore, in chapter 7, drawing ideas from [30, page 179], the modelling of share price paths via polyominos was introduced. In chapter 10, based on the relationship between binomial trees and share prices we derived in chapter 5, section[5.3.1], motivated from the computation of matroid using tutte polynomial in the paper [60, section 6], modelling of share prices via matroid and tutte polynomial was analysed.

In addition, the major part of path calculation is counting no of paths with equal probability, which is combinatorial problems. Extensive combinatorial problem was studied in the chapter 3, 4. Specifically, motivated from the definitions of generating functions and operations on generating functions by Flajolet and Sedgewick in the book [14, chapter 5] [15, chapter 1], we give examples of representing share price paths using generating function in section 3.1, 3.2. It follows by the preliminary knowledges for using Flajolet symbolic method in section 3.3, which is taken from the chapters in the book [15, chapter I. 2.1.] and presented in a more compact and constructive way.

Well-know problem of counting trees and forests using symbolic method is stated in the section 3.4 in the thesis. Counting binary trees using generating function derived from combinatorial method is described by Sedgewick, Robert, and Philippe Flajolet in the chapter 3.8 in the book [14]. Counting binary trees using symbolic method is described in chapter 5.2 in the book [14]. Relationship between enumeration of forests and trees is stated as a theorem 6.2 in the book [14]. In the section 3.4 of the thesis, along the lines of description of the generating function and symbolic method techniques described in the chapter 3.8, chapter 5.2 and theorem 6.2 by Sedgewick, Robert, and Philippe Flajolet in the book [14], we gave more detailed explanation of counting trees and forests via Flajoet symbolic method. The counts results for counting trees and forests is the well-known catalan numbers.

Motivated from the statement in the paper that a geometry behind the stock market transaction can be related to combinatorial object permutations which provides a connection of combinatoric problem to financial market [20]. According to the chapter 1.3 by Stanley in the book [21], counts for cycle structure, inversion structure and decent structure statistics on permutation are discussed in the section 3.5 of the thesis. Specifically, the exposition in the section is based on the presentation in the section [21, page 29-39]. In section 3.5.1, following the definitions of disjoint cycle

notation, standard form of disjoint cycle notation, diagraph form of permutation [21, page 29-39], using the same example, we give detailed algorithm how disjoint union of directed cycles is obtained from a disjoint cycle notation of a permutation. It follows by a proposition which is taken from [21, Proposition 1.3.1] follows the argument in the page [21, page 30] but with more clear and readable proof.

In section 3.5.2 and 3.5.3, following the counts of permutation with given cycle type taken from [21, Proposition 1.3.2], exponential generating function of the counts taken from [21, Theorem 1.3.3] and an example of applying the exponential generating function taken from [21, Example 1.3.5], counting number of permutations with a fixed cycle statistic is stated as a derivation of recurrence equation of the counts which is taken from [21, Lemma 1.3.6] but we gave a readable proof and added the argument how to set up the bounding conditions of the equation. In addition, counting number of permutations with a fixed cycle statistic is solved using generating function, the solution is taken from the second proof of [21, Proposition 1.3.7] but in which we added the derivation of $\sum_{k=0}^{n} c(n,k)t^k$ using Flajolet symbolic method.

In section 3.5.4, counting number of permutations with a fixed inversion statistic is discussed, which is mainly based on [21, page 35-37]. Specifically, motivated from the discussion of associating a permutation with a given integer sequence $(a_1, a_2, \ldots, a_n)$, with $0 \leq a_i \leq n-i$, we gave a detailed and applicable algorithm of the natural correspondence from the given integer sequence to a permutation and also give the argument to justify the algorithm is reasonable by analysing the relation between the possible values of a_{n-i} and the number of positions in the permutation u_i. Following the definitions of inversions taken from [21, page 36], motivated from the discussion regarding the integer sequence $(a_1, a_2, \ldots, a_n)$, with $0 \leq a_i \leq n-i$ on the page [21, page 34-36], we derived a proposition that claims each index i, with $0 \leq i \leq n$ in a permutaiton $w \in \mathfrak{S}_n$ can be uniquely characterized by an integer a_i in the integer sequence $(a_1, a_2, \ldots, a_n)$, with $0 \leq a_i \leq n-i$. The proposition can contribute to the proof of bijection between permutations $\mathfrak{S}_n$ and integer sequences $(a_1, a_2, \ldots, a_n)$, with $0 \leq a_i \leq n-i$, which is a proposition taken from [21, Proposition 1.3.12] after introducing the definition of inversion table taken from [21, Page 36]. Next, counting number of permutations with a fixed inversion statistic is solved using generating function which is taken from [21, Corollary 1.3.13] but in the proof we added the derivation of the derivation of $\sum_{w \in \mathfrak{S}_n} q^{inv(w)} = \sum_{(a_1, a_2, \ldots, a_n) \in \mathcal{T}_n} q^{a_1 + a_2 + \ldots + a_n}$ using Flajolet symbolic method, say Cartesian product symbolic method, which makes the proof more readable. Lastly, we stated the definition of $n!$ which is taken from the discussion on the page [21, Page 37]

In section 3.5.5, according to [21, Section 1.4], counting number of permutations with a fixed descent statistic is discussed. Specifically, using the same notation and definition of $A(d,k)$ taken from [21, Page 39], motivated from the example 1.4.2 on [21, Page 38] and the first few examples of Eulerian polynomial stated [21, Page 39], we derive a formula $A(n,k)$ for counting the number of permutations $w \in \mathfrak{S}_n$

with a fixed number of descents $k-1$. Besides, definitions of descents, descent set, the number of descents of a permutation $w \in \mathfrak{S}_n$ are taken from [21, Page 38-39]. Next, given a finite set S in increasing order, definitions of the two statistics $\alpha(S)$, $\alpha(S)$, $\beta(S)$ on permutations from descent set are taken from the page [21, Page 38]. It follows by a proposition taken from [21, Proposition 1.4.1] which states the combinatorial method of finding $\alpha(S)$ with a given finite set S but in the proof of the proposition, we added more readable argument of the combinatorial method. Next, following from the discussion on the page [21, Page 39], we summarize the definition of alternating permutation and reverse alternating permutation.

In chapter 4, motivated from a question of counting the number of lattice paths from $(0,0)$ to (n,m) which is advised by supervisor, several ways of counting unrestricted lattice paths are discussed. Applying the method of solving recurrence equations using generating functions taken from [22, Lec-31], counting unrestricted paths for the question are presented in section 4.4; the technique is to solve $a(n,m)$, try to fix one variable n and solve it using one variable generating function, then, extract the coefficient of x^m in the generating function to get the coefficient $a(n,m)$. The question can be related to the counting bitstrings of fixed length and fixed number of bits 1, which was solved using symbolic method by Sedgewick, Robert, and Philippe Flajolet in the book [14, Chapter 3.8]. Counting number of bitstrings of fixed length using symbolic method is also presented in the book [14], in which two ways of construction of bitstrings of fixed length is stated in the chapter 5.2 of the book [14]. In the thesis, taking the method from [14, chapter 3.8, 5.2], in section 4.1 we gave the two method of counting lattice path of length N and the solution of counting lattice path of length N and k up steps. In addition, another question advised by supervisor which counts the restricted lattice paths from $(0,0)$ to (N,N) with two choice steps and not going above the diagonal line is presented in section 4.2, which is motivated from [23, Example 2.7] and can also refer to Flajolet in the book [15, Page 319].

In real financial market, consider a portfolio consisting of one bond and one share, if setting the bond price stay, the portfolio price will depend on the share price only, which is a random walk. However, if setting the bond price is always up during some periods, consider the portfolio price denoted by a pair of share price and bond price, then the portfolio price would be a self-avoiding price path because it cannot go back to the same price state. This gives a motivation to counting self-avoiding walk, which was stated in the section 4.3 in the thesis. The enumeration method and presentation of counting self-avoiding walks is mainly taken from the paper [24, section7, 10] and detailed explanation of derivation of the general recursive method was added before deriving the two variable generating function $\mathcal{G}(t,v)$ and extracting the coefficient $g(n,m)$.

1.1   Structure and results

This thesis investigated combinatorial structures, path calculation methods and applied them to model individual share price path and calculate option prices. The contributions of this thesis are covered in Chapter 2, 5, 6, 7, 8, 9 and several sections

in Chapter 3, 4. This thesis is organized as follows.

Chapter 1. It contains an Introduction with the Structure of the thesis and the results.

Chapter 2. We apply path calculation method to binomial option pricing model. It starts with the Binomial Model from the path calculation approach. (Results)

Chapter 3. Motivated by Flajolet symbolic method ([15]), generating functions approach is investigated. The Path Calculations and Interpretations are important for the financial modelling. Generating function and operation on generating function can provide more ways to represent financial paths (Results).

Chapter 4. One thesis regarding a combinatoric problem is summarized, which provide new combinaotric technique that might be used for path counting. We derived the solution of counting a path not touching the given bold line segment and the technique is illustrated using an example. We applied the path calculation method to solve the winning probability in the gambler ruin problem, in which we got the same answer as using the classical method (Results).

Chapter 5, we studied the modelling of share price paths via Binary trees (Results). It is based on the possibility comes from the Bijection between the set of binary trees and the set of Dyck paths. We provided the algorithm of constructing share price path from a set of full binary trees in Section 5.1.1 and gave the share price path interpretation and gave the algorithm of constructing share price path from a set of general trees. We also provided an algorithm to represent a share price via a general binary ordered tree. Moreover, we also Count all paths not allowing given down steps in Binomial model.

Chapter 6. We studied the representation of paths for the restricted model(Results).

Chapter 7. We studied the representation of paths via polyominos. More counting is done in Count paths using Parallelogram polyominoes (Results).

Chapter 8. We studied the path calculation via the markov-type stochastic modelling, which can be generalised to counting any pattern avoiding path counting. We calculate the number of restricted share price paths not allowing consecutive down steps at time $1 \leq T \leq N$ using Markov-type technique stochastic modelling, which can be generalized to counting any pattern avoiding path counting(Results).

Chapter 9. Path modelling via matroid and tutte polynomial is studied. A new relationship between lattice integer points and tutte polyno-mials of transversal

matroid is developed and modelling of share price paths via matroids is introduced using examples(Results).

Chapter 10. The detailed results of the thesis is outlined for Chapter 2, 5, 6, 7, 8, 9 and several sections in Chapter 3, 4. The possible future works is discussed.

Chapter 2

Binomial Model

In this chapter, motivated from the Feynman path integral application in the paper [2], path calculation argument is proposed motivated from [13], [12]. Then, binomial model is studied using path calculation method which can be referred to [11], [13].

2.1 Path calculation argument

2.1.1 Motivation

the section is motivated from the section 2 [13, Page 6-8] and the introduction from the paper [12].

Consider a particle movement experiment starting from a fixed given position and observe its final position at a later time.

In classical mechanics, if the experiment is repeatedly performed in the same way such as same movement velocity, each realization would result in same final position measurement.

However, in quantum world, the situation is different. As the quantum particle has a wave like property, the result of the observation of the final position of the particle will have different outcomes in each realization when perform the experiment in the same way.

Therefore, in quantum world, for the particle movement experiment, the interest is to fix a final position of the particle and predict the probability of the quantum particle starting from a fixed given position and ending at the fixed final position.

From a fixed given position to one final position, a free quantum particle can travel in any way with any time, there might have infinite time path. So, paths truncated at some final time T will be the interested paths for the quantum particle.

Consider each possible path from one point to another point has an equal probability, each path has different possible realization time, therefore, assign each path with a complex amplitude, where its modulus square denotes the probability of the path realized in the experiment.

The different time spent on each path denotes the different direction of its probability amplitude. Suppose $x_1, x_2, \ldots, x_{n-1}$ denotes the state (eigen)vectors at each time $t_1, t_2, \ldots, t_{n-1}$, the quantum particle is at x_0 at time t_0, and in the end it is observed at x at time $T = t_n$.

In discrete case, sum up the probability amplitude of all possible paths from t_{k-1} to t_k along the state vector x_k, in each divided time period, the probability amplitude is independent, so, the probability amplitude of each possible path from t_0 at x_0 to T at x can be calculated by multiplying each period probability amplitude together.

The Feynman path integral can be summarized as identifying all possible path, truncate them at final time T and attach a probability amplitude to each possible individual path, then sum up all the probability amplitude of the valid possible path which gives the probability of a quantum particle starting from position x_0 at time t_0 to final observed position x at time $t_n = T$

2.1.2 Path Calculation Method

Motivated by Feynman path integral method in [2] [13], a path calculation method is introduced in the section.

Consider a discrete model of share price starting from a fixed initial state, the model is constructed from a random experiment which claim that share price can either go up with probability p or go down with probability p in each unit time step.

The path calculation method is that considering the model be a class of all possible share price paths, the class is denoted by A.

The path calculation idea is to identify all possible truncated individual paths in the model, associate each possible path with a probability and then sum up the probabilities of the valid paths according to different truncate time and different arriving time to finally get the required probability in the model.

It can be summarized in mathematical notation as follows,

Step 1: Truncate all possible paths before the termination time $T = n$, associate each possible individual path with a probability.

Consider the probability measure $p = q = \frac{1}{2}$ in each step, all possible paths in the model has equal probability $P(\sigma) = \frac{1}{2^n}$.

Step 2: Consider the possible paths from the fixed initial state S_0 to a final state q, sum up the probabilities of all possible paths of $(S_0 \to S_n = q)$ according to their different truncated visiting time(observing time) n. That is,

$$\mathcal{A}(q; A) = \sum_n \sum_{\sigma \in A, \sigma(n) = q} P(\sigma)$$

Step 3: Sum up the probabilities of all possible paths $(S_0 \to S_n)$ in the model

according to their different possible arriving states q, that is,

$$\mathcal{A}(A) = \sum_q \sum_n \sum_{\sigma \in A, \sigma(n)=q} P(\sigma)$$

Step 4: Consider the normalized probability of the class associated with the model, that is,

$$\mathcal{S}(A) = \sum_q \sum_n C_n \sum_{\sigma \in A, \sigma(n)=q} P(\sigma)$$

2.2   Two step binomial model

The examples and notations in this section are motivated from the chapter 2 in the lecture notes of financial mathematics [28, Page 8-21].

Suppose a sample space of a random experiment is denoted by S, and the restricted sample space of the same random experiment with additional condition is denoted by L. It is known that $|L| < |S|$.

Suppose one element in the sample space is denoted by σ, probability that the element happens is a function of the element, denoted by $P(\sigma) = Prob(\sigma)$.

Suppose another function of the element is defined and denoted by $f(\sigma)$

Example1: Two step binomial model

Suppose a random experiment is tossing a coin twice, if the value of a coin tossed is head, then share price goes up by a factor u. If the value of a coin tossed is tail, then the share price goes down by a factor d.

Therefore, the share price is modelled by a two step binomial model.

Suppose the model use the probability measure defined by $p(\text{up}) = p_u$ and $p(\text{down}) = p_d$.

Denote the value of share price at time t by S_t, and S_t is a variable. $S_t \rightarrow S_{t+1}$ means that share prices move from time t to $t+1$. The model starts from time $t = 0$ and suppose share price at time 0 is fixed and denoted by S_0.

Consider the value of share price at time $t = 2$, it is a random variable and denoted by S_2. The two step sample space corresponding to the variable S_2 is denoted by Ω_2, it is the set of all possible share price paths from S_0 to S_2.

$$\Omega_2 = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$$

where,

$$
\begin{aligned}
\sigma_1 &= S_0 \to S_0 u \to S_0 u^2 \\
\sigma_2 &= S_0 \to S_0 u \to S_0 ud \\
\sigma_3 &= S_0 \to S_0 d \to S_0 du \\
\sigma_4 &= S_0 \to S_0 d \to S_0 d^2
\end{aligned}
$$

Then, the variable S_2 is a variable which maps from Ω_2 to real number $\mathbb{R}$ (the value of the share price at time $t = 2$).

Since the option value is based on the underlying stock price, suppose the option value function at time $t = 2$ in the two step binomial model is denoted by f, it is a function depending on the value of S_2.

The possible paths σ_2 and σ_3 are two ways from the starting share price state S_0 to time 2 share price state $S_0 ud$, since their corresponding share price values S_2 at the time $t = 2$ are the same value $S_0 ud = S_0 du$.

Suppose the option claim(payoff) for the share at time $t = 2$ is C_2, then,

$$
C_2 = f(S_2)
$$

The all possible option values at time $t = 2$ is the set

$$
\begin{aligned}
C_2 &= \{f(S_2(\sigma_1)),\ f(S_2(\sigma_2)),\ f(S_2(\sigma_3)),\ f(S_2(\sigma_4))\} \\
&= \{f(S_0 u^2),\ f(S_0 ud),\ f(S_0 d^2)\} \\
&= \{C_{uu},\ C_{ud},\ C_{dd}\} \\
&\quad \text{where,} \quad C_{uu} = f(S_0 u^2),\ C_{ud} = f(S_0 ud),\ C_{dd} = f(S_0 d^2)
\end{aligned}
$$

The probability measure of the two step binomial model is defined by $p(\text{up}) = p_u$ and $p(\text{down}) = p_d$.

Suppose the two coin tosses are independent experiments from each other, the probability of outcomes in the model sample space $\Omega_2 = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$ is calculated by $P(\text{path}) = p_u{}^{\sharp \text{ups}} p_d{}^{\sharp \text{downs}}$. So,

$$
\begin{aligned}
P(\sigma_1) &= P(S_0 \to S_0 u \to S_0 u^2) = P(2\ \text{ups}) = p_u^2 \\
P(\sigma_2) &= P(S_0 \to S_0 u \to S_0 ud) = P(1\text{up},\ 1\text{down}) = p_u p_d \\
P(\sigma_3) &= P(S_0 \to S_0 d \to S_0 du) = P(1\text{up},\ 1\text{down}) = p_u p_d \\
P(\sigma_4) &= P(S_0 \to S_0 d \to S_0 d^2) = P(2\ \text{downs}) = p_d^2
\end{aligned}
$$

Suppose the defined probability is $P(\text{up}) = P(\text{down}) = \frac{1}{2}$, then, each outcome has equal probability

$$
P(\sigma_1) = P(\sigma_2) = P(\sigma_3) = P(\sigma_4) = \frac{1}{|\Omega_2|} = \frac{1}{4}
$$

and the probability of share price changes from S_0 at time 0 and $S_2 = S_0 ud$ at time 2 is

$$P(S_0 \rightarrow S_0 ud) = P(\sigma_2) + P(\sigma_3) = 2 p_u p_d = \frac{1}{2}$$

Now, share price value S_2 and the share option payoff $C_2 = f(S_2)$ at time t are two random variables and can be summarized by

$$S_2 = \begin{cases} S_0 u^2 & \text{with} \quad \frac{1}{4} \\ S_0 ud & \text{with} \quad \frac{1}{2} \\ S_0 d^2 & \text{with} \quad \frac{1}{4} \end{cases} \qquad \text{and} \qquad C_2 = \begin{cases} C_{uu} & \text{with} \quad \frac{1}{4} \\ C_{ud} & \text{with} \quad \frac{1}{2} \\ C_{dd} & \text{with} \quad \frac{1}{4} \end{cases}$$

Therefore, when each share price path has equal probability, the expected share option value at time $t = 2$ in the two step binomial model is calculated using the formula

$$\begin{aligned} E(C_2) = E(f(S_2)) \quad &= \quad \sum_{\sigma \in \Omega_2} P(\sigma) \cdot f(S_2(\sigma)), \quad \text{where} \quad \sum_{\sigma \in \Omega_2} P(\sigma) = 1 \\ &= \quad \frac{1}{4} \sum_{\sigma \in \Omega_2} f(S_2(\sigma)) \end{aligned}$$

The path calculation method in the second equation is also motivated from [2, 11]

Next, the option price in the two step binomial model is stated as

Lemma 1 *In the two step binomial model, calculate the option value at time $t = 0$ using a given risk-free interest rate r, thus, the option value is*

$$\begin{aligned} OP(C_2) = OP(f(S_2)) \quad &= \quad \frac{E(f(S_2))}{(1+r)^2} = \frac{1}{4(1+r)^2} \sum_{\sigma \in \Omega_2} f(S_2(\sigma)) \\ &= \quad \frac{1}{4(1+r)^2} (f(S_0 u^2) + f(S_0 ud) + f(S_0 du) + f(S_0 d^2)) \\ &= \quad \frac{1}{4(1+r)^2} (C_{uu} + 2 C_{ud} + C_{dd}) \end{aligned}$$

Example2: Example 1 with different probability measure

Suppose a random experiment is tossing a coin twice, same assumption as the Example 1 except that the two step stock price model is the binomial model using the probability measure defined by

$$p(\text{up}) = p_u, \ p(\text{down}) = p_d \quad \text{with} \quad p_u + p_d = 1$$

Then, the all possible option values $f(S_2)$ at time $t = 2$ is the set

$$\begin{aligned} C_2 \quad &= \quad \{f(S_2(\sigma_1)), \ f(S_2(\sigma_2)), \ f(S_2(\sigma_3)), \ f(S_2(\sigma_4))\} \\ &= \quad \{f(S_0 u^2), \ f(S_0 ud), \ f(S_0 d^2)\} \\ &= \quad \{C_{uu}, \ C_{ud}, \ C_{dd}\} \\ &\qquad \text{where,} \quad C_{uu} = f(S_0 u^2), \ C_{ud} = f(S_0 ud), \ C_{dd} = f(S_0 d^2) \end{aligned}$$

Suppose the two coin tosses are independent experiments from each other, using the formula $P(\text{path}) = p_u{}^{\sharp\text{ups}} p_d{}^{\sharp\text{downs}}$, then, the probabilities of possible share prices in the sample space Ω_2 are

$$
\begin{aligned}
P(S_0 u^2) &= P(\sigma_1) = p_u^2 \\
P(S_0 ud) &= P(\sigma_2) + P(\sigma_3) = 2 p_u p_d \\
P(S_0 d^2) &= P(\sigma_3) = p_d^2
\end{aligned}
$$

Then, using the formula

$$
E(f(S_2)) = \sum_{\sigma \in \Omega_2} P(\sigma) \cdot f(S_2(\sigma)), \quad \text{where} \quad \sum_{\sigma \in \Omega_2} P(\sigma) = 1
$$

the expected option value at time $t = 2$ is calculated as

$$
\begin{aligned}
E(C_2) = E(f(S_2)) &= \sum_{\sigma \in \Omega_2} P(\sigma) \cdot f(S_2(\sigma)) \\
&= p_u^2 f(S_0 u^2) + p_u p_d f(S_0 ud) + p_u p_d f(S_0 du) + p_d^2 f(S_0 d^2) \\
&= p_u^2 C_{uu} + 2 p_u p_d \cdot C_{ud} + p_d^2 C_{dd}
\end{aligned}
$$

The path calculation method in the second equation is also motivated from [2, 11]

Next, the option value in the model is stated as

Lemma 2 *In the two step binomial model with general probabiliy measure, the option value at time $t = 0$ is calculated using the same given risk-free interest r, and it is*

$$
\begin{aligned}
OP(C_2) = OP(f(S_2)) &= \frac{E(f(S_2))}{(1+r)^2} \\
&= (1+r)^{-2} \left(C_{uu} p_u^2 + C_{ud} \cdot 2 p_u p_d + C_{dd} p_d^2 \right)
\end{aligned}
$$

Example3: Two step binomial model not allowing two downs

Suppose a random experiment is tossing a coin twice and not allowing two consecutive downs.

If the value of a coin tossed is head, then share price goes up by a factor u. If the value of a coin tossed is tail, then the share price goes down by a factor d.

Therefore, the share price is modelled by a restricted two step binomial model with restriction not allowing two consecutive down steps, the model is also a restricted two step binomial model with restriction that the number of maximum consecutive down steps is 1.

Suppose the model use the probability measure defined by $P(\text{up}) = P(\text{down}) = \frac{1}{2}$.

Consider the value of share price at time $t = 2$, it is a variable and denoted by S_2. The restricted two step sample space corresponding to the variable S_2 is denoted by A_2, which is the set of all possible share price paths from S_0 to S_2 not allowing two consecutive down step,

$$A_2 = \{\sigma_1, \sigma_2, \sigma_3\}$$

where,

$$
\begin{aligned}
\sigma_1 &= S_0 \to S_0 u \to S_0 u^2 \\
\sigma_2 &= S_0 \to S_0 u \to S_0 ud \\
\sigma_3 &= S_0 \to S_0 d \to S_0 du
\end{aligned}
$$

Then, the variable S_2 is a random variable which maps from A_2 to real number $\mathbb{R}$ (the value of the stock price at time $t = 2$).

Suppose the two coin tosses are independent experiments from each other, then, each share price path has equal probability ,

$$
\begin{aligned}
P(\sigma_1) &= P(S_0 \to S_0 u \to S_0 u^2) = P(2 \text{ ups}) = \frac{1}{4} \\
P(\sigma_2) &= P(S_0 \to S_0 u \to S_0 ud) = P(1\text{up, 1down}) = \frac{1}{4} \\
P(\sigma_3) &= P(S_0 \to S_0 d \to S_0 du) = P(1\text{up, 1down}) = \frac{1}{4}
\end{aligned}
$$

Since the option value is based on the underlying stock price, suppose the option value function at time $t = 2$ in the restricted two step binomial model is denoted by f, it is a function depending on the value of S_2. Suppose the option claim(payoff) for the share at time $t = 2$ is C_2, then,

$$C_2 = f(S_2)$$

C_2 depends on the value of the time 2 share price S_2. S_2 depends on the share price paths in the given model sample space and share paths $S_2(\sigma_2) = S_2(\sigma_3) = S_0 ud$. The state probability $P(S_2 = S_0 ud) = P(\sigma_2) + P(\sigma_3) = \frac{1}{2}$.

In the restricted model, the all possible option values at time $t = 2$ is the set

$$
\begin{aligned}
C_2 &= \{f(S_0 u^2), \; f(S_0 ud)\} \\
&= \{C_{uu}, \; C_{ud}\} \\
&\quad \text{where,} \quad C_{uu} = f(S_0 u^2), \; C_{ud} = f(S_0 ud)
\end{aligned}
$$

Next is to calculate the expected option value at time $t = 2$ in the restricted two step binomial model using the formula

$$E(f(S_2)) = \sum_{\sigma \in A_2} \tilde{P}(\sigma) \cdot f(S_2(\sigma)), \quad \text{where} \quad \sum_{\sigma \in A_2} \tilde{P}(\sigma) = 1$$

since the probability $\tilde{P}(\sigma)$ of share price path at time $t = 2$ is assumed equal for every path $\sigma \in A_2 = \{\sigma_1, \sigma_2, \sigma_3\}$, suppose it equals to a number $\tilde{P}(\sigma) = q$, then, it is $\tilde{P}(\sigma) = \frac{1}{|A_2|} = \frac{1}{3}$.

It is noted that $\tilde{P}(\sigma) = \dfrac{P(\sigma)}{\sum_{\sigma \in A_2} P(\sigma)} = \dfrac{1/4}{3/4} = \dfrac{1}{3}$, which is a normalized probability that can be used to compute the expected option value at time $t = 2$ in the restricted model.

Therefore, the new normalized probability for the time 2 share price S_2 is calculated by

$$\begin{aligned}
\tilde{P}(S_0 u^2) &= \frac{1/4}{3/4} = \frac{1}{3} \\
\tilde{P}(S_0 ud) &= \frac{2/4}{3/4} = \frac{2}{3}
\end{aligned}$$

Now, in the model of two step binomial model not allowing two consecutive down steps, the share price value S_2 and the share option payoff $C_2 = f(S_2)$ at time t can be summarized by

$$S_2 = \begin{cases} S_0 u^2 & \text{with } \frac{1}{3} \\ S_0 ud & \text{with } \frac{2}{3} \end{cases} \quad \text{and} \quad C_2 = \begin{cases} C_{uu} & \text{with } \frac{1}{3} \\ C_{ud} & \text{with } \frac{2}{3} \end{cases}$$

So, in this case,

$$E(C_2) = E(f(S_2)) = \frac{1}{3} \sum_{\sigma \in A_2} f(S_2(\sigma)) = \sum_{S_2(\sigma), \sigma \in A_2} f(S_2(\sigma)) \cdot \tilde{P}(S_2(\sigma))$$

Next, the option price in the given restricted model is stated as

Lemma 3 *In the restricted two step binomial model not allowing two downs, the option value at time $t = 0$ is calculated using a given risk-free interest r, and it is*

$$\begin{aligned}
OP(C_2) = OP(f(S_2)) &= \frac{E(f(S_2))}{(1+r)^2} = \frac{1}{3(1+r)^2} \sum_{\sigma \in A_2} f(S_2(\sigma)) \\
&= \frac{1}{3(1+r)^2} (f(S_0 u^2) + 2f(S_0 ud)) \\
&= \frac{1}{3(1+r)^2} (C_{uu} + 2C_{ud})
\end{aligned}$$

Example4:   Two steps restricted binomial model with general probability measure

Suppose a random experiment is tossing a coin twice and not allowing two consecutive downs. It has same assumptions as the Example 3 except that the restricted two step stock price model using the probability measure defined by

$$p(\text{up}) = p_u, \ p(\text{down}) = p_d \quad \text{with} \quad p_u + p_d = 1$$

suppose the option claim for the share at time 2 is C_2, then, $C_2 = f(S_2)$, the all possible option values $f(S_2)$ at time $t = 2$ is the set

$$C_2 = \{f(S_2(\sigma_1)), f(S_2(\sigma_2)), f(S_2(\sigma_3))\}$$

the set of time 2 option values is evaluated using the price change factor u and d as

$$
\begin{aligned}
C_2 \ &= \ \{f(S_0 u^2), f(S_0 ud), f(S_0 du)\} \\
&= \ \{C_{uu}, \ C_{ud}\} \\
&\quad \text{where,} \quad C_{uu} = f(S_0 u^2), \ C_{ud} = f(S_0 ud) = f(S_0 du)
\end{aligned}
$$

Suppose the two coin tosses are independent experiments from each other, and the defined probability is $p(\text{up}) = p_u, \ p(\text{down}) = p_d \quad \text{with} \quad p_u + p_d = 1$
then, the probabilities of outcomes in the restricted two step sample space A_2 are

$$
\begin{aligned}
P(\sigma_1) \ &= \ P(S_0 \to S_0 u) \cdot P(S_0 u \to S_0 u^2) = p_u^2 \\
P(\sigma_2) \ &= \ P(S_0 \to S_0 u) \cdot P(S_0 u \to S_0 ud) = p_u p_d \\
P(\sigma_3) \ &= \ P(S_0 \to S_0 d) \cdot P(S_0 d \to S_0 ud) = p_u p_d
\end{aligned}
$$

Since the sum of the probabilities in the restricted two step sample space A_2 is not equal to one, when calculating the expected option value $E(f(S_2))$ at time $t = 2$, the normalized probabilities $\{\tilde{P}(\sigma_1), \tilde{P}(\sigma_2), \tilde{P}(\sigma_3)\}$ is used instead of $\{P(\sigma_1), P(\sigma_2), P(\sigma_3)\}$, and calculated by $\quad \tilde{P}(\sigma) = \frac{P(\sigma)}{\sum_{\sigma \in A_2} P(\sigma)}$

To summarize, the share price value S_2 and the share option payoff $C_2 = f(S_2)$ at time t can be summarized by

$$
S_2 = \begin{cases} S_0 u^2 \ \text{with} \ \frac{p_u^2}{p_u^2 + 2 p_u p_d} \\ S_0 ud \ \text{with} \ \frac{2 p_u p_d}{p_u^2 + 2 p_u p_d} \end{cases} \quad \text{and} \quad C_2 = \begin{cases} C_{uu} \ \text{with} \ \frac{p_u^2}{p_u^2 + 2 p_u p_d} \\ C_{ud} \ \text{with} \ \frac{2 p_u p_d}{p_u^2 + 2 p_u p_d} \end{cases}
$$

thus,

$$
\begin{aligned}
E(C_2) = E(f(S_2)) \ &= \ \sum_{\sigma \in A_2} \tilde{P}(\sigma) \cdot f(S_2(\sigma)) \\
&= \ \sum_{\sigma \in A_2} \left(\frac{P(\sigma)}{\sum_{\sigma \in A_2} P(\sigma)} \right) \cdot f(S_2(\sigma)) \\
&= \ \frac{p_u^2}{p_u^2 + 2 p_u p_d} f(S_0 u^2) + \frac{2 p_u p_d}{p_u^2 + 2 p_u p_d} \cdot f(S_0 ud)
\end{aligned}
$$

Next, the option price in the given restricted model using general probability measure is stated as

Lemma 4 *In the restricted two step binomial model not allowing two downs, using general probability measure, the option value at time $t = 0$ is calculated using the same given risk-free interest r, and it is*

$$
\begin{aligned}
OP(f(S_2)) &= \frac{E(f(S_2))}{(1+r)^2} \\
&= \frac{1}{(p_u^2 + 2p_u p_d)(1+r)^2} \left(p_u^2 f(S_0 u^2) + 2p_u p_d \cdot f(S_0 ud) \right) \\
&= \frac{1}{(p_u^2 + 2p_u p_d)(1+r)^2} \left(p_u^2 C_{uu} + 2p_u p_d \cdot C_{ud} \right)
\end{aligned}
$$

Example5:   Three steps binomial model with general probability measure

Suppose a random experiment is tossing a coin three times and not allowing two consecutive downs. It has same assumptions as the Example 4 except that the restricted stock price model is three steps stock price model.

From the restricted two step binomial model, the sample space of the new random experiment is denoted by A_3, which is

$$ A_3 = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5\} $$

where,

$$
\begin{aligned}
\sigma_1 &= S_0 \to S_0 u \to S_0 u^2 \to S_0 u^3 \\
\sigma_2 &= S_0 \to S_0 u \to S_0 u^2 \to S_0 u^2 d \\
\sigma_3 &= S_0 \to S_0 u \to S_0 ud \to S_0 udu \\
\sigma_4 &= S_0 \to S_0 d \to S_0 du \to S_0 duu \\
\sigma_5 &= S_0 \to S_0 d \to S_0 du \to S_0 dud
\end{aligned}
$$

Then, the variable S_3 is a variable which maps from A_3 to real number $\mathbb{R}$ (the value of the stock price at time $t = 3$).

Suppose the option claim for the share at time 3 is C_3, then, $C_3 = f(S_3)$, the all possible option values $f(S_3)$ at time $t = 3$ is the set the all possible option values $f(S_3)$ at time $t = 3$ is the set

$$ C_3 = \{f(S_3(\sigma_1)), f(S_3(\sigma_2)), f(S_3(\sigma_3)), f(S_3(\sigma_4)), f(S_3(\sigma_5))\} $$

the set of time 3 option values is evaluated using the price change factor u and d as

$$
\begin{aligned}
C_3 &= \{f(S_0 u^3), f(S_0 u^2 d), f(S_0 u^2 d), f(S_0 u^2 d), f(S_0 ud^2)\} \\
&= \{C_{uuu}, \ C_{uud}, \ C_{dud}\} \\
&\quad \text{where,} \quad C_{uuu} = f(S_0 u^3), \ C_{uud} = f(S_0 u^2 d) \\
&\qquad\qquad C_{dud} = f(S_0 ud^2)
\end{aligned}
$$

Suppose the three coin tosses are independent experiments from each other, and the defined probability is $p(\text{up}) = p_u$, $p(\text{down}) = p_d$ with $p_u + p_d = 1$ then, the probabilities of outcomes in the restricted three steps sample space A_3 are

$$
\begin{aligned}
P(\sigma_1) &= P(S_0 \to S_0 u)P(S_0 u \to S_0 u^2)P(S_0 u^2 \to S_0 u^3) = p_u^3, \\
P(\sigma_2) &= P(\sigma_3) = P(\sigma_4) = P(S_0 \to S_0 u)(S_0 u \to S_0 u^2)P(S_0 u^2 \to S_0 u^2 d) = p_u^2 p_d \\
P(\sigma_5) &= P(S_0 \to S_0 d)P(S_0 d \to S_0 duu^2)P(S_0 d \to S_0 dud) = p_u p_d^2
\end{aligned}
$$

Since the sum of the probabilities in the restricted three step sample space A_3 is not equal to one, when calculating the expected option value $E(f(S_3))$ at time $t = 3$, the normalized probabilities

$$
\{\tilde{P}(\sigma_1), \tilde{P}(\sigma_2), \tilde{P}(\sigma_3), \tilde{P}(\sigma_4), \tilde{P}(\sigma_5))\}
$$

are used instead of $\{P(\sigma_1), P(\sigma_2)P(\sigma_3), P(\sigma_4), P(\sigma_5)\}$. They are calculated by $\tilde{P}(\sigma) = \frac{P(\sigma)}{\sum_{\sigma \in A_3} P(\sigma)}$, where $\sigma \in A_3$.

To summarize, the share price value S_3 and the share option payoff $C_3 = f(S_3)$ at time t can be summarized by

$$
S_3 = \begin{cases} S_0 u^3 & \text{with} & \frac{p_u^3}{p_u^3 + 3p_u^2 p_d + p_u p_d^2} \\ S_0 u^2 d & \text{with} & \frac{3p_u^2 p_d}{p_u^3 + 3p_u^2 p_d + p_u p_d^2} \\ S_0 dud & \text{with} & \frac{p_u p_d^2}{p_u^3 + 3p_u^2 p_d + p_u p_d^2} \end{cases} \quad \text{and} \quad C_3 = \begin{cases} C_{uuu} & \text{with} & \frac{p_u^3}{p_u^3 + 3p_u^2 p_d + p_u p_d^2} \\ C_{uud} & \text{with} & \frac{3p_u^2 p_d}{p_u^3 + 3p_u^2 p_d + p_u p_d^2} \\ C_{dud} & \text{with} & \frac{p_u p_d^2}{p_u^3 + 3p_u^2 p_d + p_u p_d^2} \end{cases}
$$

thus,

$$
\begin{aligned}
E(C_3) = E(f(S_3)) &= \sum_{\sigma \in A_3} \tilde{P}(\sigma) \cdot f(S_3(\sigma)) \\
&= \sum_{\sigma \in A_3} \left(\frac{P(\sigma)}{\sum_{\sigma \in A_3} P(\sigma)} \right) \cdot f(S_3(\sigma)) \\
&= \frac{1}{p_u^3 + 3p_u^2 p_d + p_u p_d^2} \left(p_u^3 f(S_0 u^3) + 3p_u^2 p_d \cdot f(S_0 u^2 d) + p_u p_d^2 \cdot f(S_0 ud^2) \right)
\end{aligned}
$$

Next, the option price in the given three steps restricted model using general probability measure is stated as

Lemma 5 *In the restricted three steps binomial model not allowing two downs, using general probability measure, the option value at time $t = 0$ is calculated using the same given risk-free interest r, and it is*

$$
\begin{aligned}
OP(C_3) &= OP(f(S_3)) = \frac{E(f(S_3))}{(1+r)^3} \\
&= \frac{(p_u^3 f(S_0 u^3) + 3p_u^2 p_d \cdot f(S_0 u^2 d) + p_u p_d^2 \cdot f(S_0 ud^2))}{(p_u^3 + 3p_u^2 p_d + p_u p_d^2)(1+r)^3} \\
&= \frac{1}{(p_u^3 + 3p_u^2 p_d + p_u p_d^2)(1+r)^3} \left(p_u^3 C_{uuu} + 3p_u^2 p_d \cdot C_{uud} + p_u p_d^2 \cdot C_{dud} \right)
\end{aligned}
$$

2.3 Calculate Paths Using Probability Generating function

2.3.1 Methods and Examples

In this section, the generating function approach in the methods and examples are motivated from the definitions of generating function and probability generating function on the page 92, 129 by Sedgewick, Robert, and Philippe Flajolet in the book [14, chapter 3].

Based on the above two time steps binomial model of share prices, calculate $\sharp$ of paths from initial share price state S_0 to S_2 using path probability generating function. There are two ways to relate paths in a binomial model with path probability generating functions.

Method1, Relate paths of (initial state value $\rightarrow$ the final state value) of the binomial model with two sequence of numbers and corresponding two generating functions. In two time steps binomial model, that is, change $(S_0 \rightarrow S_2)$ to two generating functions $(\sum_{i=0} a_i z^i, \sum_{i=0} p_i w^i)$, where the sequence $\{a_i\}_{i=0}^N$ represent the possible values of share value S_2 at time 2, $\{p_i\}_{i=0}^N$ represent the corresponding probability sequence.

By definition, generating function can be associated with a sequence of numbers. Then, if initial share price state S_0 is a fixed value, path generating function from S_0 to S_2 can be related to two sequences of numbers, one is the sequence of share values of S_2 at time step $t = 2$, another is a corresponding probability sequence.

Denote number of paths from S_0 to S_2 by M, and M is finite. In the two steps binomial model $M = 2^2 = 4$.

Time 2 share value S_2 is a random variable and can take any possible state values in the model, and let $S_2 = \{S_0 u^2, S_0 ud, S_0 d^2\}$, where $S_0 ud$ contains two paths from S_0 to S_2 in the model.

Denote number of distinct share state values S_2 at time 2 by N. Since M is finite, N is finite. In the two steps binomial model $N = 2^2 - 1 = 3$.

Since S_0 is fixed, the set of possible values of $(S_0 \rightarrow S_2)$ can be uniquely represented by the set of possible values of S_2, which is a time 2 share value and can take possible values in the set $S_2 = \{S_0 u^2, S_0 ud, S_0 d^2\}$. The set S_2 can also be uniquely represented by the set $\{u^2, ud, d^2\}$. So, one possible

value of $(S_0 \to S_2)$ corresponds to one possible product of step increment factors in the set $\{u^2, ud, d^2\}$.

If denote the set $\{u^2, ud, d^2\}$ by a sequence of numbers $\{a_i\}_{i=0}^{N-1}$, where a_i represent one possible sum of step increment factors and the sequence is ordered by index i from 0 to $N-1$, which denote the total number of up steps in each path of the model.

Then, in the two steps binomial model, fix S_0, the possible values of $(S_0 \to S_2)$ can be denoted by a sequence $\{a_i\}_{i=0}^{2}$, where $\{a_0 = 2d, a_1 = u + d, a_2 = 2u\}$

In the model, suppose the defined probability measure for the two step binomial model is $p(\text{up}) = p_u$, $p(\text{down}) = p_d$ with $p_u + p_d = 1$.

Then the share price states from S_0 to S_2 has the corresponding probability sequence $\{p_i\}_{i=0}^{2}$, where $p_i = P(S_2 = a_i)$ and the probability sequence is $\{p_0 = p_d^2, p_1 = 2p_u p_d, p_2 = p_u^2\}$.

Thus, in terms of path probability generating function, the two step binomial model can be written as

$$
\begin{cases}
G_S(z) & = & \displaystyle\sum_{i \geq 0}^{2} a_i z^i = 2d + (u + d)z + 2uz^2 \\
G_P(w) & = & \displaystyle\sum_{i \geq 0}^{2} p_i w^i = \sum_{i \geq 0}^{2} P(S_2 = a_i) w^i = p_d^2 + 2p_u p_d w + p_u^2 w^2
\end{cases}
$$

Method2, Motivated from the notations in the book [27], relate each path of $(S_0 \to S_1 \to S_2 \to \ldots \to S_N)$ of the binomial model with two sequence of numbers.

One is $\{a_i\}_{i=0}^{N}$, where $a_i = $ share price S_i at time i, $N = $ number of time steps in the path.

The other is $\{p_i\}_{i=0}^{N}$, where p_i denotes probability of share change from $t = i - 1$ to $t = i$. S_0 is fixed, so, set $p_0 = 1$.

Each time period probability is defined by the probability measure $p_{i \to i+1}(\text{up}) = p_u$, $p_{i \to i+1}(\text{down}) = p_d$ with $p_u + p_d = 1$.

The probability of time t share value S_i is

$$
P(S_0 \to S_1 \to \ldots \to S_i) = p_0 \prod_{j=1}^{i} p_{(j-1 \to j)} = p_0 \prod_{j=1}^{i} p_i.
$$

The generating function approach in the next two examples is motivated by by Sedgewick, Robert, and Philippe Flajolet from the definitions of generating function in the book [14, chapter 3.1] and notations motivated from the lecture note of financial mathematics [28, Chapter 2].

Example 2.3.1 *[28, 14]:*

In the two step binomial model, if only considering one share price changes and consider each path as two generating functions,

then, take one possible value of $(S_0 \to S_1 \to S_2)$ *as an example, say,*

take a path $(S_0 \to S_0 u \to S_0 u^2)$, *since the state value* S_0 *is fixed, the path can be represented by* $(1, u, u^2)$. *If considering share price change factor at each time period, the path can also be represented by* $(1, u, u)$. *Correspondingly, the path probability sequence can be represented by* $(p_0, \ p_0 p_1, \ p_0 p_1 p_2)$, *or can be represented by a sequence* $(p_0, \ p_1, \ p_2) = (1, \ p_1, \ p_2)$.

To summarize, the path has two sequence represented as

$$\begin{cases} \{a_i\}_{i=0}^2 &= \{S_0, S_0 u, S_0 u^2\} \\ \{p_i\}_{i=0}^2 &= \{p_0, p_1, p_2\} = \{1, p_1, p_2\} \end{cases}$$

then, the path has two generating functions written as

$$\begin{cases} \sigma(x) &= \displaystyle\sum_{i \geq 0}^2 a_i x^i = S_0 + S_0 u x + S_0 u^2 x^2 \\ P_\sigma(w) &= \displaystyle\sum_{i \geq 0}^2 p_i w^i = 1 + p_1 w + p_2 w^2 \end{cases}$$

Example 2.3.2 *[28, 14]:*

In the three step binomial model, If considering price changes of two shares portfolio, and assuming only allowing one share to change upward in each time step.

Denote the two share portfolio by (S_t^1, S_t^2), *where* $t = 0, 1, 2, 3$.

Then, one possible value of $[(S_0^1, S_0^2) \to (S_1^1, S_1^2) \to (S_2^1, S_2^2) \to (S_3^1, S_3^2)]$ *represent a two share portfolio prices path.*

In each time step, the portfolio has two choices of change, either share 1 *change upward or share* 2 *change upward, that is, for* $0 \le i,\ j \le 3$,

$(S_0 u^i, S_0 u^{j+1})$ (share 2, up)

$\uparrow$

state $(S_0 u^i, S_0 u^j) \longrightarrow (S_0 u^{i+1}, S_0 u^j)$ (share 1, up)
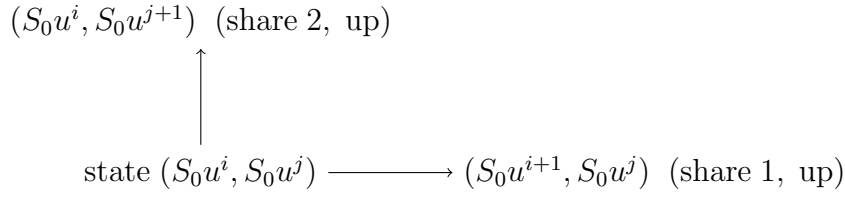
Figure 2.1: one step possible change of the portfolio

For example,

take a path $[(S_0^1, S_0^2) \to (S_0^1 u, S_0^2) \to (S_0^1 u, S_0^2 u) \to (S_0^1 u^2, S_0^2 u)]$,

since the portofolio's initial state value (S_0^1, S_0^2) *is fixed, the path can be represented by* $[(1,1) \to (u,1) \to (u,u) \to (u^2,u)]$.

If considering the powers of change factors of each portfolio state value in the path,

the path can also be represented by $[(0,0) \to (1,0) \to (1,1) \to (2,1)]$, *which is*

$(1,1) \longrightarrow (2,1)$

$\uparrow$

$(0,0) \longrightarrow (1,0)$
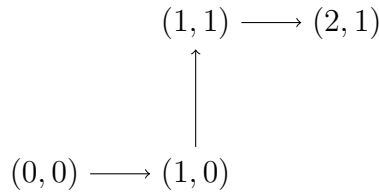
Figure 2.2: one example path

Correspondingly, the path probability sequence can be represented by $(p_{0,0},\ p_{0,0}p_{1,0},\ p_{0,0}p_{1,0}p_{1,1},\ p_{0,0}p_{1,0}p_{1,1}p_{2,1})$, *or can be represented by a sequence* $(p_{0,0},\ p_{1,0},\ p_{1,1},\ p_{2,1}) = (1,\ p_{1,0},\ p_{1,1},\ p_{2,1})$.

So, if in the three steps binomial model and considering two share portfolio price changes, the probability sequence is $\{p_{i,j}\}_{i,j=0}^{i+j=3}$, *where* $p_{i,j}$ *denotes the probability of share portfolio change from* $t = i+j-1$ *to* $t = i+j$.

The probabilities of the two choices of portfolio change are shown as

To summarize, the share portfolio path has two sequence represented as

$$\begin{cases} \{a_{i,j}\}_{i,j=0}^{3} & = \{(S_0^1, S_0^2), (S_0^1 u, S_0^2), (S_0^1 u, S_0^2 u), (S_0^1 u^2, S_0^2 u)\} \\ \{p_{i,j}\}_{i,j=0}^{i+j=3} & = \{p_{0,0},\ p_{1,0},\ p_{1,1},\ p_{2,1}\} = \{1,\ p_{1,0},\ p_{1,1},\ p_{2,1}\} \end{cases}$$

$$P_{i,j+1} \text{ (share 2, up)}$$

$$\uparrow$$

$$\text{state } (i,j) \longrightarrow P_{i+1,j} \text{ (share 1, up)}$$
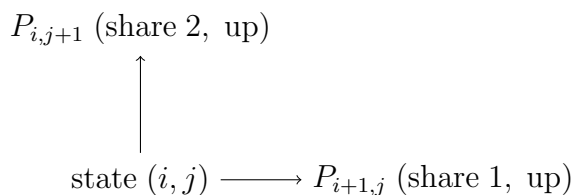
Figure 2.3: probabilities of portfolio possible change each step

then, the path has two generating functions written as

$$
\begin{cases}
\sigma(x,y) & = \displaystyle\sum_{i,j\geq 0}^{3} a_{i,j} x^i y^j \\[2mm]
& = (S_0^1, S_0^2)x^0 y^0 + (S_0^1 u, S_0^2)x^1 y^0 + (S_0^1 u, S_0^2 u)x^1 y^1 + (S_0^1 u^2, S_0^2 u)x^2 y^1 \\[2mm]
P_\sigma(w) & = \displaystyle\sum_{i,j\geq 0}^{3} p_{i,j} w^{i+j} = 1 + p_{1,0}w + p_{1,1}w^2 + p_{2,1}w^3
\end{cases}
$$

2.3.2   Represent all paths in the model using Generating function

In general term, each node state (i,j) can be related with a monomial term $x^i y^j$, consider a path
$$\sigma = (0,0) \to (1,0) \to (1,1) \to (1,2) \to (2,2) \to (3,2) \to (3,3)$$

the corresponding portfolio value generating function is
$$\sigma(x,y) = x^0 y^0 + x^1 y^0 + x^1 y^1 + x^1 y^2 + x^2 y^2 + x^3 y^2 + x^3 y^3$$

if $x=0, y=0$, then $\sigma(0,0) = 1$, which is the number of path represented by the generating function $\sigma(x,y)$.

Similarly, if a two share portfolio state value is represented as
$$\sigma = (0,0) \to (1,0) \to (2,0) \to (3,0) \to \dots$$

It means only share 1 price change and it goes up in each time step.

the generating function for the two share portfolio value is
$$\sigma(x,y) = x^0 y^0 + x^1 y^0 + x^2 y^0 + x^3 y^0 + \dots = \sum_{j=0}^{\infty} x^j = \frac{1}{1-x}$$

Lemma 6 *Suppose a binomial model representing change of two shares portfolio has N time steps.*

It can be seen that in $\sigma(x, y) = \sum_{i,j=0}^{\infty} x^i y^j$, if powers of terms in the generating function uniquely corresponds to a time step sequence $\{i + j = k\}_{k=0}^{k \leq N}$, then, the generating function can be a path.

Example 2.3.3 *For example, a polynomial $x^0 y^0 + x^2 y^0$ is not a two share portfolio path, since step $(0,0) \rightarrow (2,0)$ is not allowed by the assumption (only one share change in each time step). It is obvious that the powers of terms in the polynomial corresponds to a time step sequence $\{0, 2\}$, and it missed time step 1.*

The next example 1 is motivated from the Feynman path integral [13] and its application in finance in the page 6 of the paper [2, Theorem 2] and [11].

Example 1:

Consider one model Ω_n for one share prices of length n,

let a path $\sigma = \sigma_1 \sigma_2 \ldots \sigma_n = $ one possible value of $(S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow S_n)$, and the possible path $\sigma = \sigma_1 \sigma_2 \ldots \sigma_n \in \Omega_n$, where $\sigma_i = $ the step change $(i - 1, S_{i-1}) \rightarrow (i, S_i)$.

Let $\sigma(x)$ be the generating function for the path, and it is defined by $\sigma(x) :=$
$$\sum_{i=0}^{n} S_i(\sigma) x^{|\sigma_1 \sigma_2 \ldots \sigma_i|},$$ where $|\sigma_1 \sigma_2 \ldots \sigma_n|$ is the number of steps of the path $\sigma_1 \sigma_2 \ldots \sigma_i$.

Then, the path generating function for the model is defined by

$$A_{\Omega_n}(x) := \sum_{\sigma \in \Omega_n} \sigma(x)$$
$$= \sum_{\sigma \in \Omega_n} \sum_{i=0}^{n} S_i(\sigma) x^{|\sigma_1 \sigma_2 \ldots \sigma_i|}$$

and the path probability generating function for the model is defined by

$$P_{\Omega_n}(w) = \sum_{\sigma \in \Omega_n} \sum_{i=0}^{n} P_i(\sigma) w^{|\sigma_1 \sigma_2 \ldots \sigma_i|}$$

It is noted that

$$\begin{aligned} S_i(\sigma) &= \text{ the share value at time } i \text{ of the path } \sigma, \\ P_i(\sigma) &= \text{ the state probability at time } i \text{ in the path } \sigma \\ &= P(\sigma_1 \sigma_2 \ldots \sigma_i) \end{aligned}$$

Thus, in the model $\Omega_n,$ the expected share price value at time n can be calculated by

$$E[S_n(\sigma)] = \sum_{\sigma=\sigma_1\sigma_2...\sigma_n \in \Omega_n} S_n(\sigma)P_n(\sigma)$$

Suppose the model Ω_n is a n steps binomial model, and the risk-free interest rate is r. Next is to calculate the price of the exotic option claim in the model.

Suppose the option claim for the share price at time t is $C_t = f(S_t)$, for $1 \le t \le n$. Let T be the final time, and $1 \le t \le T$.

The exotic option claim in the model is defined by $C = \sum_{j=1}^{T} f(S_t)$

Thus, the option value of the exotic claim is stated as

Theorem 7 *The option value of the exotic claim at time $t = 0$ is calculated using the given risk-free interest rate r and it is*

$$
\begin{aligned}
OP(C) &= OP(\sum_{j=1}^{T} f(S_t)) = \sum_{j=1}^{T} OP(f(S_t)) \\
&= \sum_{j=1}^{T} [(1+r)^{-j} \cdot E(f(S_j))] \\
&= \sum_{j=1}^{T} [(1+r)^{-j} \cdot \left(\sum_{\sigma \in \Omega_n} f(S_j(\sigma))P_j(\sigma) \right)] \\
&= \sum_{\sigma=\sigma_1\sigma_2...\sigma_n \in \Omega_n} \sum_{j=1}^{T} \frac{f(S_j(\sigma))P_j(\sigma)}{(1+r)^j}
\end{aligned}
$$

The path calculation method in the theorem generalize the path calculation lemma in the paper [2] [11].

In the model $\Omega_n,$ since the starting state of the share price is fixed, for each path $\sigma \in \Omega_n,$ set $P_0(\sigma) = 1$. From the path probability generating function $P_{\Omega_n}(w)$ of the model, it can be observed that

$$
\begin{aligned}
P_{\Omega_n}(0) &= \sum_{\sigma \in \Omega_n} P_0(\sigma) \\
&= \sum_{\sigma \in \Omega_n} 1 \\
&= \sharp \text{ of the possible paths in the model } \Omega_n
\end{aligned}
$$

Another useful form of generating functions of the model Ω_n can be written as

$$
\begin{cases}
\mathcal{A}_{\Omega_n}(x) = \sum_{i=0}^{n} \sum_{\sigma \in \Omega_n} S_i(\sigma) x^i \\
P_{\Omega_n} = \sum_{i=0}^{n} \sum_{\sigma \in \Omega_n} P_i(\sigma) w^i = \sum_{i=0}^{n} \sum_{\sigma \in \Omega_n} P(\sigma_1 \sigma_2 \ldots \sigma_i) w^i
\end{cases}
$$

Then, it can be observed that the coefficients has the following meaning,

$$
\begin{cases}
\sum_{\sigma \in \Omega_n} S_i = \text{the sum of the possible state values at time } i \\
\sum_{\sigma \in \Omega_n} P_i(\sigma) = \sum_{\sigma \in \Omega_n} P(\sigma_1 \sigma_2 \ldots \sigma_i) \\
\qquad = \text{the sum of the corresponding states probabilities at time } i
\end{cases}
$$

It is noted that if is useful to write the formula as above. If we assume the sum of possible share price values is finite at time i, with $0 \le i \le n$,

Then, the corresponding generating functions for the model Ω_n is valid for counting and Ω_n becomes a combinatorial class.

Chapter 3

Generating functions approach

A generating function(GF) is a clothesline on which we hang up a sequence of numbers for display(Herbert Wilf, 1994). Generating function can provide a compact representation of a sequence and it is a method to encode a recurrence relation on a sequence. The key idea is to equivalent a recurrence relation on a sequence to a functional equation satisfied by GF.

Suppose a random experiment is tossing a coin twice, if the value of a coin tossed is head, denoted by 1, then stock price goes up by a factor u. if the value of a coin tossed is tail, denoted by -1, then the stock price goes down by a factor d.

Therefore, the stock price is modelled by a two step binomial model.

Suppose the model use the probability measure defined by

$$P(up) = p, \ P(down) = q \quad \text{with} \quad p + q = 1$$

Generating function is a way of representing a sequence of numbers, which can corresponds to one particular share price path using the two ways of correspondence mentioned in chapter 2.

3.1 Definitions of OGFs and EGFs

The definition 8 of ordinary generating function is motivated by Flajolet and Sedgewick on page 92 in the book [14, Chapter 3.1].

Definition 8 *(Ordinary generating function) [14, Chapter 3.1]*
Given a sequence $\{a_k\}_{k=0}^{\infty} = \{a_0, a_1, a_2, \ldots\}$, the function $A(z) = \sum_{k \geq 0}^{\infty} a_k \cdot z^k$, $\forall k \in \mathbb{N}$ is called the ordinary generating function(OGF) of the sequence.

Notation: Use notation $[z^k]A(z)$ to refer to the coefficient a_k. It is a formal power series, we are interested the coefficient a_k of the formal power series.

In the next example, an example of sequence $\{1, 1, \ldots, 1\}$ is taken from the page 93 in the book [14, Chapter 3.1], we provide a way to use the sequence to represent a share price path with given initial share price S_0, up and down change factor of

the share price u, and d. Then, calculate the corresponding ordinary generating function pairs.

Example 3.1.1 *A sequence is:*

$$1, 1, 1, \ldots, 1; \quad here, \quad a_k = 1, \ k = 0, 1, 2, \ldots$$

Consider the correspondence $\{a_k\}_{k=0}^{\infty} \leftrightarrow \{S_k\}_{k=0}^{\infty}$;

$$\{S_0 \to S_0 u \to S_0 u^2 \to \ldots\} \quad \leftrightarrow \quad (1, u, u^2, \ldots)$$
$$\leftrightarrow \quad (1, u, u, \ldots) \ [Consider\ share\ price\ change\ factors]$$
$$\leftrightarrow \quad (1, 1, 1, \ldots) \ [Powers\ of\ share\ price\ change\ factors]$$

$\{p_k\}_{k=0}^{\infty}$ *denotes the probability of share price change from* $t = k - 1$ *to* $t = k$. *Set* $p_0 = 1$ *which set the probability of share price at time* 0 *equal to* S_0 *is 1.*

The sequence corresponds to a path which have up at each step, and the probability of up at each step is p, *and the number of step of the path is infinite. The path can has two sequences:*

$$(S_0, S_0 u, S_0 u^2, S_0 u^3, \ldots) \quad and \quad (1, p, p^2, p^3, \ldots)$$

It has two ordinary generating function(OGF):

$$\left(\sum_{k \geq 0}^{\infty} S_0 u^k z^k, \quad \sum_{k \geq 0}^{\infty} p^k w^k \right) = (\frac{S_0}{1 - uz}, \ \frac{1}{1 - pw})$$

Sometimes it is more convenient to handle sequence by a generating function involving a normalizing factor, it is called exponential generating function.

The definition 9 of ordinary generating function is motivated by Flajolet and Sedgewick on page 97 in the book [14, Chapter 3.2].

Definition 9 *(Exponential generating function)*
Given a sequence $(a_0, a_1, \ldots, a_k, \ldots)$, *the function* $A(z) = \sum_{k \geq 0}^{\infty} a_k \cdot \frac{z^k}{k!}$, $\quad \forall k \in \mathbb{N}$ *is called the exponential generating function(EGF) of the sequence.*

Notation: Use notation $k![z^k]A(z)$ to refer to the coefficient a_k. It is a exponential power series, we are more interested the coefficient a_k of the exponential power series.

In the next example, the same example of sequence $\{1, 1, \ldots, 1\}$ is taken from the page 93 in the book [14, Chapter 3.1], we provide a way to use the sequence to represent a share price path with given initial share price S_0, up and down change factor of the share price u, and d. Then, calculate the corresponding exponential generating function pairs.

Example 3.1.2 *Sequence:*

$$1, 1, 1, \ldots, 1; \quad here, \quad a_k = 1, \; k = 0, 1, 2, \ldots$$

Consider the correspondence $\{a_k\}_{k=0}^{\infty} \leftrightarrow \{S_k\}_{k=0}^{\infty}$;

$$\{S_0 \to S_0 u \to S_0 u^2 \to \ldots\} \quad \leftrightarrow \quad (1, u, u^2, \ldots)$$
$$\leftrightarrow \quad (1, u, u, \ldots) \, [Consider \; share \; price \; change \; factors]$$
$$\leftrightarrow \quad (1, 1, 1, \ldots) \, [Powers \; of \; share \; price \; change \; factors]$$

$\{p_k\}_{k=0}^{\infty}$ *denotes the probability of share price change from* $t = k - 1$ *to* $t = k$. *Set* $p_0 = 1$ *which set the probability of share price at time* 0 *equal to* S_0 *is 1.*

The sequence corresponds to a path which have up *at each step, and the probability of* up *at each step is* p, *and the number of step of the path is infinite. The path can has two sequences:*

$$(S_0, S_0 u, S_0 u^2, S_0 u^3, \ldots) \quad and \quad (1, p, p^2, p^3, \ldots)$$

It has Exponential generating function(EGF):

$$\left(\sum_{k \geq 0}^{\infty} S_0 u^k \frac{z^k}{k!}, \quad \sum_{k \geq 0}^{\infty} p^k \frac{w^k}{k!} \right) = (S_0 e^{uz}, \; e^{pw})$$

In the next example, an example of sequence $\{1, 1, 2, 6, 24, 120, \ldots, n!\}$ is taken from the table 3.3 in the book [14, Chapter 3.2], we provide a way to use the sequence to represent a share price path with given initial share price S_0, up and down change factor of the share price u, and d. Then, calculate the corresponding exponential generating function pairs.

Example 3.1.3 *Sequence:*

$$1, 1, 2, 6, 24, 120, \ldots; \quad here, \quad a_n = n!, \; n = 0, 1, 2, \ldots$$

Consider the correspondence $\{a_k\}_{k=0}^{\infty} \leftrightarrow \{S_k\}_{k=0}^{\infty}$;

$$\{S_0 \to S_0 u \to S_0 u^{2!} \to S_0 u^{3!} \ldots\} \quad \leftrightarrow \quad (1, u, u^{2!}, u^{3!} \ldots)$$
$$\leftrightarrow \quad (1, u, u^2, u^3 \ldots) \, [Consider \; share \; price \; change \; factors]$$
$$\leftrightarrow \quad (1, 1, 2!, 3! \ldots) \, [Powers \; of \; share \; price \; factors]$$

$\{p_k\}_{k=0}^{\infty}$ denotes the probability of share price change from $t = k - 1$ to $t = k$. Set $p_0 = 1$ which set the probability of share price at time 0 equal to S_0 is 1.

It corresponds to a path which have n up steps at each time step, and the probability of up at each step is p, and the number of step of the path is infinite. The path can be represented as the two sequences:

$$(1, u, u^2, u^3, u^4, \ldots) \quad and \quad (1, p, p^2, p^3, \ldots)$$

Given a fixed initial share price S_0, the sequence can be represented as exponential generating function(EGF):

$$\left(\sum_{k \geq 0}^{\infty} u^k \frac{z^k}{k!}, \quad \sum_{k \geq 0}^{\infty} p^k \frac{w^k}{k!} \right) = (e^{uz}, \ e^{pw})$$

The benefits of exponential generating function provide a compact way to represent factorial jump share price path using exponential.

It is obvious that generating functions is a way to represent sequences. When doing some operations on a given sequence of numbers, the generating function will be changed correspondingly. Then, definitions of common operations on ordinary generating functions(OGFs) are given follows.

3.2   Definitions of common operations on OGFs

The lemma 10 of scaling operation of ordinary generating function is motivated by Flajolet and Sedgewick on page 94 in the book [14, Chapter 3.1].

Lemma 10 *(Scaling)*
If $A(z) = \sum_{k \geq 0} a_k z^k$ is the ordinary generating function of a sequence $\{a_k\}_{k=0}^{\infty} = \{a_0, a_1, a_2, \ldots\}$, when scaling the sequence by a sequence of scalar numbers $\{1, c, c^2, \ldots\}$; each term a_k is scaled with the same scalar c^k.

Then $A(cz) = \sum_{k \geq 0} a_k c^k z^k$ is the ordinary generating function of the sequence $\{c^k a_k\}_{k=0}^{\infty}$

In the next example, starting from the same example of sequence $\{1, 1, \ldots, 1\}$ which is taken from the page 93 in the book [14, Chapter 3.1], a transformed share price by scaling operation is calculated, then, we calculate the corresponding scaling share price path ordinary generating function pairs.

Example 3.2.1 *Example: Original sequence is:*

$$1, 1, 1, \ldots, 1; \quad here, \quad a_N = 1, \ N = 0, 1, 2, \ldots$$

Consider the correspondence $\{a_k\}_{k=0}^{\infty} \leftrightarrow \{S_k\}_{k=0}^{\infty}$;

From the example in section 6.1, the sequence corresponds to a path which have up at each step, and the probability of up at each step is p, and the number of step of the path is infinite. The path can has two sequences:

$$(S_0, S_0 u, S_0 u^2, S_0 u^3, \dots) \quad and \quad (1, p, p^2, p^3, \dots)$$

It has two ordinary generating function(OGF):

$$\left(\sum_{k\geq 0}^{\infty} S_k\, z^k, \sum_{k\geq 0}^{\infty} p^k w^k\right) = \left(\sum_{k\geq 0}^{\infty} S_0 u^k z^k, \sum_{k\geq 0}^{\infty} p^k w^k\right) = \left(\frac{S_0}{1 - uz},\ \frac{1}{1 - pw}\right)$$

The scaled sequence $\{2^k\}_{k=0}^{\infty}$:

$$1, 2, 4, 8, 16, 32, \dots, 2^N \cdot 1, \dots;$$

Consider the correspondence $\{b_k\}_{k=0}^{\infty} \leftrightarrow \{\hat{S}_k\}_{k=0}^{\infty}$;

then, its ordinary generating function(OGF):

$$\sum_{k\geq 0} \hat{S}_k z^k = \sum_{k\geq 0} S_0 u^k 2^k z^k = \frac{S_0}{1 - cuz} = \sum_{k\geq 0} S_0 u^k 2^k z^k$$

here, the coefficient of the generating function is $[z^k]\frac{S_0}{1-cuz} = S_0(uc)^k$

It corresponds to a path which have 2 up jump change at each time step compared to the original share price path, and considering the probability of 2up at each time step is p, and the number of step of the path is infinite. The path has two sequences:

$$(S_0, S_0 2u, S_0 2^2 u^2, S_0 2^3 u^3, \dots) \quad and \quad (1, p, p^2, p^3 \dots)$$

It has two ordinary generating function(OGF):

$$\left(\sum_{k\geq 0} \hat{S}_k z^k, \sum_{k\geq 0}^{\infty} p^k w^k\right) = \left(\sum_{k\geq 0}^{\infty} S_0 (2u)^k z^k, \sum_{k\geq 0}^{\infty} p^k w^k\right) = \left(\frac{S_0}{1 - 2uz},\ \frac{1}{1 - pw}\right)$$

The lemma 11 of adding operation of two ordinary generating function is motivated by Flajolet and Sedgewick on page 94 in the book [14, Chapter 3.1].

Lemma 11 *(Addition)*
If $A(z) = \sum_{k\geq 0} a_k z^k$ is the OGF of $a_0, a_1, a_2, \dots, a_k, \dots$
and $B(z) = \sum_{k\geq 0} b_k z^k$ is the OGF of $b_0, b_1, b_2, \dots, b_k, \dots$
then $A(z) + B(z)$ is the OGF of $a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots, a_k + b_k, \dots$

In the next example, starting from the sequence $\{1,1,\ldots,1\}$ which is taken from the page 93 in the book [14, Chapter 3.1], and a sequence $\{1,2,4,8,16,32,\ldots,2^N \cdot 1,\ldots\}$, a transformed share price is calculated by adding operation on the corresponding two share share price ordinary generating functions.

Example 3.2.2 *Original sequence 1:*

$$1,1,1,\ldots,1; \quad here, \quad a_N = 1, \ N = 0,1,2,\ldots$$

from the prior examples in this section, it can corresponds to a share price path which has OGF:

$$\sum_{k\geq 0} S_0 u^k z^k = \frac{S_0}{1-uz}$$

Original sequence 2:

$$1,2,4,8,16,32,\ldots,2^N \cdot 1,\ldots; here, \quad b_N = 2^N, \ N = 0,1,2,\ldots$$

it can corresponds to a share price path which has OGF:

$$\sum_{k\geq 0} S_0(cu)^k z^k = \frac{S_0}{1-cuz}$$

Consider the correspondence $\{c_k\}_{k=0}^{\infty} \leftrightarrow \{\tilde{S}_k\}_{k=0}^{\infty}$;

$\{\tilde{S}_k\}_{k=0}^{\infty}$ *corresponds to a path which have the generating function* $\frac{S_0}{1-uz} + \frac{S_0}{1-cuz}$

considering the probability of $\frac{\tilde{S}_{k+1}}{\tilde{S}_k}$ *at each time step is* p, *and the number of step of the path is infinite. The path has two sequences:*

$$(\tilde{S}_0, \tilde{S}_1, \tilde{S}_2, \tilde{S}_3, \ldots) \quad and \quad (1, p, p^2, p^3 \ldots)$$

It has two ordinary generating function(OGF):

$$\left(\sum_{k\geq 0} \tilde{S}_k z^k, \sum_{k\geq 0}^{\infty} p^k w^k\right) = \left(\frac{S_0}{1-uz} + \frac{S_0}{1-cuz}, \ \frac{1}{1-pw}\right)$$

From the prior examples(scaling) in this section, the sequence corresponds to a share price path which can be decomposed into one share price path $\{S_k\}_{k=0}^{\infty} = \{S_0 u^k\}_{k=0}^{\infty}$ *and another share price path* $\{\hat{S}_k\}_{k=0}^{\infty} = \{S_0(2u)^k\}_{k=0}^{\infty}$.

The lemma 12 of Differentiation operation of ordinary generating function is motivated by Flajolet and Sedgewick on page 94 and theorem 3.1 in the book [14, Chapter 3.1].

Lemma 12 *(Differentiation)*
If $A(z) = \sum_{k \geq 0} a_k z^k$ is the OGF of $a_0, a_1, a_2, \ldots, a_k, \ldots$

then $zA'(z) = \sum_{k \geq 1} k a_k z^k$ is the OGF of $0, a_1, 2a_2, 3a_3, \ldots, k a_k, \ldots$

From the lemma of differentiation operation on generating function and we gave a definition 13 of differentiation path transformation as follows.

Definition 13 *(*Differentiation path transformation*).*
Let share price sequence/path be denoted by $\{S_i\}$, construct OGF: $S(z) = \sum_j S_j z^j$. Then, define a symbolic transform OGF, and called $\hat{S}(z) = S'(z)$ and this defines a differentiation transformed path.

In the next example, from the definition of differentiation operation on generating function, starting from the sequence $\{1, 1, \ldots, 1\}$ which is taken from the page 93 in the book [14, Chapter 3.1], we provide a way to use the sequence representing the share price and then derive the first order differentiation transformed share price path. Then, in the remarks after the example, we give the second order differentiation transformed share price starting from the same sequence $\{1, 1, \ldots, 1\}$.

Example 3.2.3 *Example 1:*
It is known that an ordinary generating function is:

$$\frac{1}{1-z} = \sum_{N \geq 0} z^N$$

its sequence:

$$1, 1, 1, \ldots, 1; \quad here, \quad a_N = 1$$

from the prior examples in this section, it can corresponds to a share price path which has OGF:

$$\sum_{k \geq 0} S_0 u^k z^k = \frac{S_0}{1 - uz}$$

and the share price path is

$$\{S_k\}_{k=0}^{\infty} = (S_0, S_0 u, S_0 u^2, S_0 u^3, \ldots)$$

Consider the correspondence $\{b_k\}_{k=0}^{\infty} \leftrightarrow \{\hat{S}_k\}_{k=0}^{\infty}$;

then, its ordinary generating function(OGF):

$$\sum_{k \geq 0} \hat{S}_k z^k = \frac{d(\frac{S_0}{1-uz})}{dz} = (\frac{S_0}{1-uz})' = \frac{S_0 u}{(1-uz)^2}$$

from the definition of differentiation of generating function, an OGF can be obtained as follows by differentiating the original OGF:

$$z(\frac{S_0 u}{(1 - uz)^2}) = \sum_{k \geq 0} k S_k z^k = \sum_{k \geq 0} k S_0 u^k z^k$$

It corresponds to the share price path:

$$(\ 0, \ \{S_k\}_{k=1}^{\infty} \ ) = (\ 0, \ S_0 u, \ S_0 u^2, \ \dots \ )$$

Then,

$$
\begin{aligned}
\frac{S_0 u}{(1 - uz)^2} &= \sum_{k \geq 0} k S_0 u^k z^{k-1} \\
&= \sum_{k \geq 1} k S_0 u^k z^{k-1} \quad [\textit{No considering the negative power of } z]
\end{aligned}
$$

Considering the probability of $\frac{\hat{S}_{k+1}}{\hat{S}_k}$ at each time step is p, and the number of step of the path is infinite. The path has two sequences:

$$(\hat{S}_0, \hat{S}_1, \hat{S}_2, \hat{S}_3, \dots) \quad and \quad (1, p, p^2, p^3 \dots)$$

It has two ordinary generating function(OGF):

$$(\sum_{k \geq 0} \hat{S}_k z^k, \sum_{k \geq 0}^{\infty} p^k w^k) = \left(\frac{S_0 u}{(1 - uz)^2}, \ \frac{1}{1 - pw} \right)$$

Compared to the original path $\{S_k\}_{k=0}^{\infty} = \{S_0 u^k\}_{k=0}^{\infty}$, its differentiation transformed path becomes $\{\hat{S}_k\}_{k=0}^{\infty} = \{S_0 k u^k\}_{k=1}^{\infty} = \{S_0 u, \ S_0 2u^2, \ \dots\}$, which has a jump $\frac{\hat{S}_{k+1}}{S_k} = \frac{k+1}{k} u$.

Example 3.2.4 *(how to use the differentiation operation on OGF):*

1.

$$(\frac{S_0}{1 - uz})' = (\sum_{k \geq 0} S_k z^k)' = \sum_{k \geq 0} k S_k z^{k-1} = \sum_{k \geq 1} k S_k z^{k-1}$$

$$(\frac{S_0}{1 - uz})' = \frac{S_0 u}{(1 - uz)^2}$$

then,

$$\frac{S_0 u}{(1 - uz)^2} = \sum_{k \geq 1} k S_k z^{k-1}$$

It corresponds to a share price path

$$\hat{S}_k = (\ \{k \cdot S_k\}_{k=1}^{\infty} \ ) = (\ \{k \cdot S_0 u^k\}_{k=1}^{\infty} \ )$$

2. *When proceeding the differentiation to the 2nd order,*

$$\left(\frac{S_0}{1-uz}\right)'' = \left(\sum_{k\geq 0} S_k z^k\right)'' = \left(\sum_{k\geq 0} k S_k z^{k-1}\right)'$$

$$= \sum_{k\geq 0} k(k-1) S_k z^{k-2}$$

$$= \sum_{k\geq 2} k(k-1) S_k z^{k-2}$$

$$\left(\frac{S_0}{1-uz}\right)'' = \left(\frac{S_0 u}{(1-uz)^2}\right)' = \frac{S_0 2u^2}{(1-uz)^3}$$

then,

$$\frac{S_0 2u^2}{(1-uz)^3} = \frac{1}{z^2} \sum_{k\geq 2} k(k-1) S_k z^k$$

So, the sequence of numbers for the ordinary generating function

$$z^2\left(\frac{S_0 u^2}{(1-uz)^3}\right) = \frac{1}{2} \sum_{k\geq 2} k(k-1) S_k z^k = \frac{1}{2} \sum_{k\geq 0} k(k-1) S_k z^k$$

corresponds to a share price path

$$c_k = \frac{k(k-1)\cdot S_k}{2} = \frac{k(k-1)}{2} S_k = \binom{k}{2} S_k, \quad k = 0, 1, 2, \ldots$$

which has a jump $\frac{\hat{S}_{k+1}}{S_k} = \frac{\binom{k+1}{2}}{\binom{k}{2}} u = \frac{k+1}{k-1} u$ *and has the first two steps share price of zero;*

$$(c_0 = 0, c_1 = 0, c_2 = S_0 u^2, c_3 = S_0 3u^3, c_4 = S_0 6u^4, \ldots)$$
$$= \left(0,\ 0,\ S_2,\ \binom{3}{2} S_3,\ \binom{4}{2} S_4, \ldots\right)$$

where, $\{S_k\}_{k=0}^{\infty}$ *is the original share price path* $\{S_0, S_0 u, S_0 u^2, \ldots\}$

Similar to 1st order differentiation of generating functions, if considering the differentiation to the 2nd order transformed share price path, it has the generating function

$$\frac{S_0 2u^2}{(1-uz)^3} = \sum_{k\geq 2} k(k-1) S_k z^{k-2}$$

It corresponds to the transformed share price path

$$\hat{S}_k = \left(\ \{k(k-1)\cdot S_k\}_{k=2}^{\infty}\ \right) = \left(\ \{k(k-1)\cdot S_0 u^k\}_{k=2}^{\infty}\ \right)$$

In general case, it is easy to get a common used sequence equation and we derive the general order differentiation, say M-order, transformed share price path in the following example.

Example 3.2.5 *When $M >= 1$, a M-order differentiation of the ordinary generating function is:*

$$\left(\frac{S_0}{1-uz}\right)^{(M)} = \frac{S_0 M! u^M}{(1-uz)^{M+1}} = \left(\sum_{k\geq 0} S_k z^k\right)^{(M)}$$

$$= \frac{1}{z^M}\sum_{k\geq M} k(k-1)\ldots(k-M+1)S_k z^k$$

Then,

$$z^M\left(\frac{S_0 u^M}{(1-uz)^{M+1}}\right) = \sum_{k\geq M}\frac{k(k-1)\ldots(k-M+1)}{M!}S_k z^k = \sum_{k\geq M}\binom{k}{M}S_k z^k$$

it corresponds to a share price path

$$c_k = \frac{k(k-1)\ldots(k-M+1)}{M!}S_k = \binom{k}{M}S_k, \quad k = 0,1,2,\ldots$$

which has a jump $\frac{\hat{S}_{k+1}}{S_k} = \frac{\binom{k+1}{M}}{\binom{k}{M}}u = \frac{k+1}{k-M+1}u$ *and has the first M steps share price of zero;*

$$(c_0 = 0,\ldots,c_{M-1} = 0, c_M = S_0 u^M, c_{M+1} = S_0(M+1)u^{M+1}, c_4 = S_0\frac{M+2}{2}u^{M+2},\ldots)$$

$$= \left(0,\ 0,\ \ldots,\ S_M,\ \binom{M+1}{M}S_{M+1},\ \binom{M+2}{M}S_{M+2},\ldots\right)$$

$$= \left(0,\ 0,\ \ldots,\ S_M,\ \binom{M+1}{1}S_{M+1},\ \binom{M+2}{2}S_{M+2},\ldots\right)$$

where, $\{S_k\}_{k=0}^{\infty}$ *is the original share price path* $\{S_0, S_0 u, S_0 u^2,\ldots\}$

if considering the differentiation to the M-th order transformed share price path, it has the generating function

$$\frac{S_0 M! u^M}{(1-uz)^{M+1}} = \sum_{k\geq M} k(k-1)\ldots(k-M+1)S_k z^{k-M}$$

It corresponds to the transformed share price path

$$\hat{S}_k = \left(\ \{k(k-1)\ldots(k-M+1)\cdot S_k\}_{k=1}^{\infty}\ \right) = \left(\ \{k(k-1)\ldots(k-M+1)\cdot S_0 u^k\}_{k=1}^{\infty}\ \right)$$

Remarks

1. It is easy to check the above ordinary generating function using a general fact about binomial coefficient formula

$$\binom{-k}{n} = (-1)^n \cdot \binom{k+n-1}{n};$$

that is,

$$
\begin{aligned}
S_0 u^M z^M (1 - uz)^{-(M+1)} &= S_0 u^M z^M \sum_{k \geq 0} \binom{-(M+1)}{k} (uz)^k \\
&= S_0 u^M z^M \sum_{k \geq 0} \binom{k + (M+1) - 1}{k} (uz)^k \\
&= S_0 u^M z^M \sum_{k \geq 0} \binom{k+M}{k} (uz)^k \\
&= S_0 \sum_{k \geq 0} \binom{k+M}{k} (uz)^{k+M} = S_0 \sum_{k \geq M} \binom{k}{k-M} (uz)^k \\
&= \sum_{k \geq M} \binom{k}{M} S_0 u^k z^k = \sum_{k \geq M} \binom{k}{M} S_k z^k
\end{aligned}
$$

The lemma 14 of integration operation of ordinary generating function is motivated by Flajolet and Sedgewick on page 94 and theorem 3.1 in the book [14, Chapter 3.1].

Lemma 14 *(Integration)*
If $A(z) = \sum_{k \geq 0} a_k z^k$ *is the OGF of* $a_0, a_1, a_2, \ldots, a_k, \ldots$

then $\displaystyle\int_0^z A(t)dt = \sum_{n \geq 1} \frac{a_{n-1}}{n} z^n$ *is the OGF of* $0, a_0, \dfrac{a_1}{2}, \dfrac{a_2}{3}, \ldots, \dfrac{a_{k-1}}{k}, \ldots$

In the next example, starting from the same example of sequence $\{1, 1, \ldots, 1\}$ which is taken from the page 93 in the book [14, Chapter 3.1], a transformed share price path by integration operation is derived.

Example 3.2.6 *It is known that an ordinary generating function is:*

$$\frac{1}{1-z} = \sum_{N \geq 0} z^N$$

its corresponding sequence of numbers:

$$1, 1, 1, \ldots, 1; \quad here, \quad a_N = 1, \ N = 0, 1, 2, \ldots$$

from the prior examples in this section, it can corresponds to a share price path which has OGF:

$$\sum_{k \geq 0} S_0 u^k z^k = \frac{S_0}{1 - uz}$$

and the share price path is

$$\{S_k\}_{k=0}^{\infty} = (S_0, S_0 u, S_0 u^2, S_0 u^3, \dots)$$

Consider the correspondence $\{b_k\}_{k=0}^{\infty} \leftrightarrow \{\hat{S}_k\}_{k=0}^{\infty}$;

Since

$$\frac{S_0}{u}(\ln \frac{1}{1-uz})' \;\; = \;\; \frac{S_0}{1-uz} = S_0 \sum_{k \geq 0} u^k z^k = \sum_{k \geq 0} S_k z^k$$

Integrating both side with respect to z, we can obtain OGF:

$$\frac{S_0}{u}(\ln \frac{1}{1-uz}) = \sum_{k \geq 0} \frac{S_k z^{k+1}}{k+1} = \sum_{k \geq 1} \frac{S_{k-1} z^k}{k}$$

When $k = 0$, the initial share price should be set as $\hat{S}_0 = 0$ in the path $\{\hat{S}_k\}_{k=0}^{\infty}$,

so, the transformed share price path is

$$\{\hat{S}_k\}_{k=0}^{\infty} = \left(0, \{\frac{S_{k-1}}{k}\}_{k=1}^{\infty} \right) = \left(0, \{\frac{S_0 u^{k-1}}{k}\}_{k=1}^{\infty} \right)$$

The lemma 15 of convolution operation of ordinary generating function is motivated by Flajolet and Sedgewick on page 95 in the book [14, Theorem 3.1].

Lemma 15 *(Convolution)*
If $A(z) = \sum_{k \geq 0} a_k z^k$ *is the OGF of* $a_0, a_1, a_2, \dots, a_k, \dots$
and $B(z) = \sum_{k \geq 0} b_k z^k$ *is the OGF of* $b_0, b_1, b_2, \dots, b_k, \dots$
then $A(z)B(z)$ *is the OGF of* $a_0 b_0, a_0 b_1 + a_1 b_0, \dots, \sum_{0 \leq k \leq n} a_k b_{n-k}, \dots$

Proof of the lemma:

$$A(z)B(z) \;\; = \;\; \sum_{k \geq 0} a_k z^k \sum_{n \geq 0} b_n z^n$$

$$\overset{\text{Distribute}}{=} \sum_{k \geq 0} \sum_{n \geq 0} a_k b_n z^{n+k}$$

$$\overset{\text{Change n to n-k}}{=} \sum_{k \geq 0} \sum_{n \geq k} a_k b_{n-k} z^n$$

$$\overset{\text{Switch order of summation}}{=} \sum_{n \geq 0} \sum_{0 \leq k \leq n} a_k b_{n-k} z^n$$

The proof of convolution operation can be referred to [14, Theorem 3.1].

From the definition of convolution operation on generating function and we gave a corollary 16 of convolution path transformation as follows, then, we gave the convolution share price path explanation.

Corollary 16 *(Convolution path transformation)*
Let two share price sequences/paths be denoted by $\{S_i^1\}$, $\{S_i^2\}$ *respectively, construct two OGF:* $S^1(z) = \sum_j S_j^1 z^j$, $S^2(z) = \sum_j S_j^2 z^j$. *Then, define a symbolic convolution transform OGF, and called* $\hat{S}(z) = S^1(z)S^2(z)$ *and this defines a convolution transformed path*

$$\{\hat{S}_k\}_{k=0}^{\infty} = \left(\sum_{0 \leq j \leq k} S_j^1 S_{k-j}^2 \right)$$

Example 3.2.7 *Convolution Path explanation:*

The convolution transformed share price can be constructed as follows,

firstly, construct two share price paths $\{S_j^1\}_{j=0}^{\infty}$, *and* $\{S_j^2\}_{j=0}^{\infty}$,

secondly, choose one share price paths, say, $\{S_j^2\}_{j=0}^{\infty}$, *truncate the path at a finite time* k $0 \leq k \leq \infty$

To find the convolution share price at time k, *construct the inversed time share price path* $\{\tilde{S}_j^2\}_{j=k}^{0}$,

Then,

$$\{\hat{S}_k\} = \sum_{0 \leq j \leq k} S_j^1 S_{k-j}^2$$

One common application is 'Partial sum'. The following example is taken from the page 94 in the book [14, Section 3.1].

Example 3.2.8 *If* $A(z) = \sum_{k \geq 0} a_k z^k$ *and* $\dfrac{1}{1-z} = \sum_{N \geq 0} z^N$,

then $\dfrac{1}{1-z} A(z) = \sum_{n \geq 0} (\sum_{0 \leq k \leq n} a_k) z^n$,

So, $[z^N] \dfrac{1}{1-z} A(z) = \sum_{0 \leq k \leq N} a_k$

3.3 Methods using generating functions

The generating functions(GFs) are a useful method to solve linear recurrence equations. It is also useful to count combinatorial classes with GFs.

In this section, we give the preliminary knowledges for using Flajolet symbolic method, which is motivated by Flajolet and Sedgewick in the book [14, Theorem 3.1], [14, Chapter 5.2] and [15, Chapter I. 2.1.], we presented the preliminary knowledge of symbolic method construction in a more compact way.

3.3.1 Solving recurrence equations

In this section, solving recurrence equations using generating function is stated which is motivated by Flajolet and Sedgewick in the book [14, Section 3.3]. It follows an example of solving recurrence equation which is taken from the page 104 in the book [14, Section 3.3].

Method description:

1. Create recurrence equation and make the recurrence equation valid for all positive integer n;

2. Multiply both sides of the recurrence equation by z^n and sum over on positive integer n;

3. Evaluate the sums to derive a functional equation satisfied by OGF.

4. Solve the equation to derive an explicit formula for the OGF.

5. Expand the OGF to obtain the coefficients of GF.

Example 3.3.1 *Given a recurrence equation:*

$$a_n = 5a_{n-1} - 6a_{n-2} \quad \forall n \geq 2 \quad with \ a_0 = 0 \ and \ a_1 = 1$$

Find: the explicit formula of $a_n, \ n \geq 0$.

Solution:

1. Make recurrence valid for all n, we add delta funtion on the RHS of recurrence equation

$$\delta_{n,1} = \begin{cases} 1, & \text{if} \quad n = 1 \\ 0, & \text{if} \quad n \neq 1 \end{cases}$$

we get:

$$a_n = 5a_{n-1} - 6a_{n-2} + \delta_{n,1}$$

2. Multiply both sides by z^n and sum over $n \geq 0$

$$A(z) = 5zA(z) - 6z^2 A(z) + z$$

3. Solve

$$A(z) = \frac{z}{1 - 5z + 6z^2}$$

4. Factor denominator $1 - 5z + 6z^2 = (1 - 3z)(1 - 2z)$
   Use partial fractions:

$$A(z) = \frac{c_0}{1 - 3z} + \frac{c_1}{1 - 2z}$$

5. Solve for coefficients:

$$\begin{cases} c_0 + c_1 = 0 \\ 2c_0 + 3c_1 = -1 \end{cases}$$

6. The coefficient

$$c_0 = 1 \quad c_1 = -1$$

$$A(z) = \frac{1}{1 - 3z} + \frac{1}{1 - 2z}$$

7. Use the related formulas in the section common operations on GF, it is easily to obtain that

$$[z^n]A(z) = 3^n - 2^n$$

3.3.2   Counting combinatorial classes(Symbolic Method)

In this section, the knowledge of symbolic method is stated which is motivated by Flajolet and Sedgewick in the book [14, Chapter 5.2] and the book [15, Chapter I. 2.1.].

Description of Symbolic Method:

The symbolic method is an important method in analytic combinatorial. It is an approach for translating formal definition of combinatorial objects into functional equations on generating function. The symbolic method is dissected as follows,

- Define a class of combinatorial objects, say, $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$

- Define a notion of size of an object in combinatorial class, say, for $\alpha \in \mathcal{A}$,   the size $|\alpha|$

- Define GF is sum over all members of the combinatorial class. say, $A(z) \equiv \sum_{\alpha \in \mathcal{A}} z^{|\alpha|}$
  Each term z^N in the GF corresponds to an object of size $|\alpha| = N$.

  Collect all the terms with the same size exponent N to expose counts that counting the number of objects with size N, say, denote by a_n.

- Define operations suitable for combinatorial constructions of different class of combinatorial objects.

  Common used operations for combinatorial construction is union(sum), product, and sequence, definition of the three operations will be given in the later section.

- Develop translations from combinatorial constructions to operations on GFs.

  In the later section, using GF counting unlabelled objects is introduced and corresponding transfer theorem *for translating combinatorial constructions into OGF* is introduced as propositions and proof is given later.

Firstly, Definitions of Combinatorial Classes is given as follows.

Definition 17 *(Combinatorial Classes) [14, Page 221]*
A combinatorial class $\mathcal{A}$ is a finite or countable set of combinatorial objects associated with a size function for each object $\|$ such that:

- $\forall\, \alpha \in \mathcal{A}, \quad 0 \leq |\alpha| < \infty$;

- *the number of elements of a given size is finite.*

Remark: Define a class of combinatorial objects with associated size function, by introducing symbolic method, we can get another view of generating function(GF).

Secondly, Definitions of Combinatorial Constructions is given as follows.

Definition 18 *(Cartesian product/product) [14, Page 223-225]*
Let $\mathcal{A}$ and $\mathcal{B}$ be two combinatorial classes, their product(Cartesian product) is the set of ordered pair of copies of objects, one copy from $\mathcal{A}$, and the other from $\mathcal{B}$, i.e. it is defined as $\mathcal{C} = \mathcal{A} \times \mathcal{B} = \{(\alpha, \beta) : \alpha \in \mathcal{A}, \beta \in \mathcal{B}\},$ for notational economy, write the product set as $\mathcal{C} = \mathcal{A} \times \mathcal{B} = \{\alpha\beta : \alpha \in \mathcal{A}, \beta \in \mathcal{B}\}$

Remark: For each object γ in the product class $\mathcal{C}$, $\gamma \in \mathcal{A} \times \mathcal{B}$, it has a size function $|\gamma| = |\alpha\beta| = |\alpha| + |\beta|$, because each object in the product set of $\mathcal{A}$ and $\mathcal{B}$ is one object from combinatorial class $\mathcal{A}$ followed by one object from another combinatorial class $\mathcal{B}$.

Definition 19 *(Combinatorial Sum/Disjoint Union) [14, Page 223-225]*
Let $\mathcal{A}$ and $\mathcal{B}$ be two combinatorial classes, their disjoint union(sum) is defined to be the union of disjoint combinatorial classes without disjointness condition imposed on combinatorial classes.
Using standard set-theoretic unions notations,the combinatorial sum of $\mathcal{A}$ and $\mathcal{B}$ is defined as:

- *1st, Using two distinct markers $\circ$ and $\diamond$ for the two combinatorial classes $\mathcal{A}$ and $\mathcal{B}$; $\circ$ is a marker for $\mathcal{A}$ and $\diamond$ for $\mathcal{B}$.*

- *Each marker has a size function of value zero; $|\circ| = 0$ and $|\diamond| = 0$.*

Then, the disjoint union(sum) is defined as $\mathcal{C} = \mathcal{A} + \mathcal{B} := (\{\circ\} \times \mathcal{A}) \cup (\{\diamond\} \times \mathcal{B})$.

Remark: Each object γ in the disjoint union class $\mathcal{C} = \mathcal{A} + \mathcal{B}$ is either $\gamma \in \{\circ\} \times \mathcal{A}$ or $\gamma \in \{\diamond\} \times \mathcal{B}$, the object is a copy of the same size from the original combinatorial classes $\mathcal{A}$ and $\mathcal{B}$; that is, $\gamma \in \{\circ\} \times \mathcal{A}$ is a copy $\alpha \in \mathcal{A}$ and $\gamma \in \{\diamond\} \times \mathcal{B}$ is a copy $\beta \in \mathcal{B}$. It is noted that $\gamma \in \mathcal{A} + \mathcal{B}$, $|\gamma| =$ the size it had originally.

Definition 20 *(Sequence Construction) [14, Page 223-225]*
Let $\mathcal{A}$ *be a combinatorial class, the sequence of* $\mathcal{A}$ *is a sequence of objects* $\mathcal{A}$ *, i.e.*
it is a class defined as an infinite sum by $SEQ(A) = \{\epsilon\} + \mathcal{A} + \mathcal{A} \times \mathcal{A} + \mathcal{A} \times \mathcal{A} \times \mathcal{A} + \ldots,$
in other words, $SEQ(A) = \{(\alpha_1, \ldots, \alpha_l) \mid l \geq 0, \ \alpha_j \in \mathcal{A}, \ 0 \leq l \leq l\}.$

Remark:

1. $\{\epsilon\}$ is a set containing one object of size $|\epsilon| = 0$, an object of size zero is called a neutral object; ϵ is a neutral object corresponding to $(\alpha_1, \ldots, \alpha_l)$, with $l = 0$.

2. To use symbolic method, the sequence of $\mathcal{A}$ defined above should also be a combinatorial class. By the definition of combinatorial class(Definition 17), we need to ensure the class satisfies (1) each object in the class has a size function; (2) the number of objects with a given size, say, $|\gamma| = n$ is finite, i.e. the coefficient of term z^n , say, a_n is finite.

   Question: How to make sure the sequence of a combinatorial class $\mathcal{A}$, $SEQ(\mathcal{A})$ is a combinatorial class?

   Answer:

   (a) It is obvious that the first condition satisfies, as each object $\gamma \in SEQ(\mathcal{A})$ is constructed from combinatorial class $\mathcal{A}$ in which each object has size function.

   (b) To count the number of objects of size $'n'$ in $SEQ(\mathcal{A})$, we need to know how many objects of size $'n'$ there are in $\mathcal{A}^j = \mathcal{A} \times \overset{j}{\cdots} \times \mathcal{A}$; by definition of product of combinatorial classes, the size of an object in $\mathcal{A}^j$ is the sum of the sizes of its j components.

   Aim: we would like to make sure the number of objects of size $'n'$ in $SEQ(\mathcal{A})$ is finite.

   If all objects in $\mathcal{A}$ have size at least one, then the objects in $\mathcal{A}^j$ have size at least j , then the terms of the form $\mathcal{A}^j$, for $j > n$ will never give rise to objects of size $'n'$, instead, if there are objects in $\mathcal{A}$ of size 0, then we may have objects of size $'n'$ in all terms of $\mathcal{A}^j$, including $j > n$, therefore, there will be an infinite number of objects in $SEQ(\mathcal{A})$ of size $'n'$, it contradicts the definition of combinatorial class.

   (c) In summary, when considering sequence operation of a combinatorial class, say, $SEQ(\mathcal{A})$, we only consider the sequence construction $SEQ(\mathcal{A})$ of combinatorial class $\mathcal{A}$ which does not have elements of size zero, that is, $\mathcal{A}$ contains no object of size zero.

Thirdly, Definition of Basic Construction Blocks is given as follows.

Definition 21 *(Basic building blocks) [14, Page 221]*
Let $\mathcal{A}$ *be a combinatorial class, let* α *be an object in the combinatorial class, each*
object $\alpha \in \mathcal{A}$ *has a size function. An object of size one is called an atom. It builds*

up an atomic class, denoted by Z. An object of size zero is called a neutral object, it builds up a neutral class, denoted by E. If a combinatorial class contains nothing and has no object, it is called empty class, denoted by $\emptyset$, when translating it into GF, its GF is '0', then there is no corresponding term $z^{|\alpha|}$.

Sequence construction: [15, Page 25]

1. If a set $\mathcal{B}$ is a combinatorial class,

   Then, the sequence class denoted by $SEQ(\mathcal{B})$, it is defined by an infinite sum $SEQ(\mathcal{B}) = \{\epsilon\} + \mathcal{B} + (\mathcal{B} \times \mathcal{B}) + (\mathcal{B} \times \mathcal{B} \times \mathcal{B}) + \dots$, where

   ϵ is a neutral object of size zero.

2. The sequence represents the combinatorial class $\mathcal{A}$, where $\mathcal{A} = \{(\beta_1, \dots, \beta_l) \mid l \geq 0, \beta_j \in \mathcal{B}\}$. The neutral object (structure) in $\mathcal{A}$ corresponds to when $l = 0$; if $l = 0$, the sequence is (β_1, β_0) and it is not valid. So, it represent the empty word ϵ which has size (here, length) zero.

3. The combinatorial class $\mathcal{B}$ must not contain object of size 0, otherwise, inside the class $\mathcal{A} = SEQ(\mathcal{B})$, fix size n, the number of elements of size n will be infinite.

4. From the definition of size for sums and products, for any element $\alpha = (\beta_1, \dots, \beta_l) \in \mathcal{A}$, the size of the element is the sum of the size of its components, that is, $|\alpha| = \sum_{i=1}^{l} |\beta_i|$

Multiset construction: [15, Page 26]

1. Multisets are finite sets but arbitrary repetition of elements are allowed.

2. If a combinatorial class $\mathcal{A}$ is obtained by all finite multisets of elements from $\mathcal{B}$, the set $\mathcal{A}$ is denoted by $\mathcal{A} = MSET(\mathcal{B})$.

3. The multiset constructed from the set $\mathcal{B}$ is defined by

$$MSET(\mathcal{B}) = SEQ(\mathcal{B})/R$$

   with the equivalence relation R is defined by $(\alpha_1, \dots, \alpha_r) \; R \; (\beta_1, \dots, \beta_r)$ iff there exists some arbitrary permutation σ of $[1, \dots, r]$ such that, for all $1 \leq j \leq r$, $\beta_j = \alpha_{\sigma(j)}$.

Cycle construction: [15, Page 26]

1. Notation: if $\mathcal{B}$ is a combinatorial class, then the cycle construction from the class $\mathcal{B}$ is denoted by $CYC(\mathcal{B})$.

2. Assume no empty cycle, $CYC(\mathcal{B})$ is defined by

$$CYC(\mathcal{B}) := (SEQ(\mathcal{B}) \setminus \{\epsilon\})/S$$

where for any two elements in the class $CYC(\mathcal{B})$, S is their equivalence relation defined by $(\beta_1, \ldots, \beta_r)$ S $(\beta'_1, \ldots, \beta'_r)$ iff there exists some circular shift τ of $[1, \ldots, r]$ such that, for all $1 \leq j \leq r$, $\beta'_j = \beta_{1+(j-1+d) \bmod r}$.

3. For example, $(\beta_1, \beta_2, \beta_3)$ S $(\beta_2, \beta_3, \beta_1)$ S $(\beta_3, \beta_1, \beta_2)$,

here, $r = 3$, $(\beta_2, \beta_3, \beta_1) = (\beta'_1, \beta'_2, \beta'_3)$, it circularly shift one element from $(\beta_1, \beta_2, \beta_3)$,

so, it shift $d = 1$, $\beta'_j = \beta_{1+(j-1+1) \bmod 3}$

$(\beta_3, \beta_1, \beta_2) = (\beta'_1, \beta'_2, \beta'_3)$, it circularly shift two element from $(\beta_1, \beta_2, \beta_3)$,

so, it shift $d = 2$, $\beta'_j = \beta_{1+(j-1+2) \bmod 3}$

It is noted that $1 \leq d \leq r$.

4. For example, assuming no empty sequence, the set of sequences formed from the class $\mathcal{B} = \{a, b\}$ is denoted by

$$\mathcal{A}^{(3)} = \{(\beta_1, \beta_2, \beta_3) | \beta_j \in \mathcal{B}\} = SEQ_3(\mathcal{B}) = \mathcal{B} \times \mathcal{B} \times \mathcal{B}$$

then, $\sharp \mathcal{A}^{(3)} = 2^3 = 8$

The class of cycles formed from the class $\mathcal{A}^{(3)}$ is $CYC^{(3)}(\mathcal{B}) = \mathcal{A}^{(3)}/S$, where S is defined by $(\beta_1, \beta_2, \beta_3)$ S $(\beta'_1, \beta'_2, \beta'_3)$ iff there exists some circular shift τ of $[1, 2, 3]$ such that, for all $1 \leq j \leq 3$, $\beta'_j = \beta_{1+(j-1+d) \bmod 3}$, with $1 \leq d \leq 3$, representing the number of positions shifted.

Then, sequences are grouped into equivalence classes according to the relation S, and $\sharp CYC^{(3)}(\mathcal{B}) = 4 = 2^{r-1}$, with $r = 3$.

Powerset construction: [15, Page 26]

1. Notation: $PSET(\mathcal{B})$, where $\mathcal{B}$ is a combinatorial class.

2. A powerset is a combinatorial class consisting of all finite subsets of the combinatorial class $\mathcal{B}$, if starts from a multiset construction $MSET(\mathcal{B})$ of $\mathcal{B}$, the powerset is formed from multisets that involve no repetitions.

3. Let $\tilde{\mathcal{A}} = \{(\beta_1, \ldots, \beta_l) \mid 0 \leq l \leq card(\mathcal{B}), \beta_i \neq \beta_j \text{ if } i \neq j, \beta_i \in \mathcal{B}\}$,

Then, $\mathcal{A} = PSET(\mathcal{B}) = \tilde{\mathcal{A}}/R$, with the equivalence relation R is defined by $(\alpha_1, \ldots, \alpha_r)$ R $(\beta_1, \ldots, \beta_r)$ iff there exists some arbitrary permutation σ of $[1, \ldots, r]$ such that, for all $1 \leq j \leq r$, $\beta_j = \alpha_{\sigma(j)}$.

4. For any element $\alpha = (\beta_1, \beta_2, \ldots, \beta_l) \in \mathcal{A} = PSET(\mathcal{B})$, define its size function using notation $|\cdot|$ for the size function, then, $|\alpha| = \sum_{i=1}^{l} |\beta_i|$

5. For example, represent natural numbers, Let $Z := \{\cdot\}$, with $\cdot$ an atom of size 1,

Then, $\mathcal{I} = SEQ(Z) \{\epsilon\} = \{(\beta_1, \beta_2, \ldots, \beta_l)|l \geq 1, \ \beta_i \in \mathcal{B}\}$

Three common operations on combinatorial classes (product, disjoint union and sequence) are very often be used to define a combinatorial class in terms of atoms and other combinatorial classes, when counting number of combinatorial objects in a combinatorial class. When using the three common operations get a definition of the combinatorial class, use transfer theorem to translate operations on combinatorial classes into operations on generating function. The transfer theorem are introduced as follows. Which is motivated by Flajolet and Sedgewick in the book [14, Page 225].

Proposition 22 *[14, Theorem 5.1]*
The generating function for the product of two combinatorial classes is the product of their generating functions.

Suppose: We are given generating functions $A(z)$ and $B(z)$ for $\mathcal{A}$ and $\mathcal{B}$,
Aim: Prove the generating function for $\mathcal{A} \times \mathcal{B}$ equals to $A(z)B(z)$.

Proof:
Following definition of product operation in the last section, if

$$\mathcal{C} = \mathcal{A} \times \mathcal{B} = \{\alpha\beta : \alpha \in \mathcal{A}, \beta \in \mathcal{B}\}$$

The key idea is that the sum over all product members $\alpha\beta \in \mathcal{A} \times \mathcal{B}$ equals to two steps: (1) firstly, for each member $\alpha \in \mathcal{A}$, sum over the α followed by the sum over all members $\beta \in \mathcal{B}$; (2) secondly, sum over all members $\alpha \in \mathcal{A}$, followed by the sum over all members $\beta \in \mathcal{B}$. that is,

$$\sum_{\alpha\beta \in \mathcal{A} \times \mathcal{B}} = \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{B}}$$

Besides, for each object in $\mathcal{A} \times \mathcal{B}$, its size function is $|\alpha\beta| = |\alpha| + |\beta|$
Therefore,

$$
\begin{aligned}
\sum_{\gamma \in \mathcal{A} \times \mathcal{B}} z^{|\gamma|} \quad &= \quad \sum_{\alpha\beta \in \mathcal{A} \times \mathcal{B}} z^{|\alpha\beta|} \\
&= \quad \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{B}} z^{\alpha|+|\beta|} \\
&\overset{distribute}{=} \quad \sum_{\alpha \in \mathcal{A}} z^{|\alpha|} \sum_{\beta \in \mathcal{B}} z^{|\beta|} \\
&= \quad A(z)B(z)
\end{aligned}
$$

Proposition 23 *[14, Theorem 5.1]*
The generating function for the sum(disjoint union) of two combinatorial classes(disjoint) is the sum of their corresponding generating functions.

Suppose: We are given generating functions $A(z)$ and $B(z)$ for $\mathcal{A}$ and $\mathcal{B}$,

Aim: Prove the generating function for $\mathcal{A} + \mathcal{B}$ equals to $A(z) + B(z)$

Proof: Following definition of disjoint union operation in the last section, if

$$\mathcal{C} = \mathcal{A} + \mathcal{B} = \{\gamma : \gamma \in \mathcal{A} \cup \mathcal{B}\}$$

Firstly, for each object, $\gamma \in \mathcal{A} \cup \mathcal{B}$, its size equals to the size it had originally. Secondly, as γ is copies from $\mathcal{A}$ or copies from $\mathcal{B}$, then the sum over all disjoint union members $\gamma \in \mathcal{A} \cup \mathcal{B}$, equals to that taking sum over members in each combinatorial class $\mathcal{A}$ and $\mathcal{B}$ respectively, then taking an addition of the two sum. That is,

$$\sum_{\gamma \in \mathcal{A} + \mathcal{B}} \overset{\substack{\text{if A and B} \\ \text{are disjoint set}}}{=} \sum_{\gamma \in \mathcal{A}} + \sum_{\gamma \in \mathcal{B}}$$

Therefore, we obtain

$$\begin{aligned}
\sum_{\gamma \in \mathcal{A} + \mathcal{B}} z^{|\gamma|} &= \left(\sum_{\gamma \in \mathcal{A}} + \sum_{\gamma \in \mathcal{B}} \right) z^{|\gamma|} \\
&= \sum_{\gamma \in \mathcal{A}} z^{|\gamma|} + \sum_{\gamma \in \mathcal{B}} z^{|\gamma|} \\
&= A(z) + B(z)
\end{aligned}$$

Following definition of Sequence operation in the last section, if

$$SEQ(\mathcal{A}) = \epsilon + \mathcal{A} + \mathcal{A} \times \mathcal{A} + \mathcal{A} \times \mathcal{A} \times \mathcal{A} + \ldots$$

we give the following proposition,

Proposition 24 *[14, Theorem 5.1]*
The generating function for the sequence of a combinatorial class whose generating function is $A(z)$ is $\frac{1}{1-A(z)}$.

Suppose: Given a generating function $A(z)$ for a combinatorial class $\mathcal{A}$,

Aim: Prove the Generating function for $SEQ(\mathcal{A})$ equals to $\frac{1}{1-A(z)}$

Proof:

If we assume there are no objects of size 0 in $\mathcal{A}$, then $SEQ(\mathcal{A})$ is also a combinatorial class.

Following definition of sequence operation in the last section, we know

$$SEQ(\mathcal{A}) = \epsilon + \mathcal{A} + \mathcal{A} \times \mathcal{A} + \mathcal{A} \times \mathcal{A} \times \mathcal{A} + \ldots$$

By applying the transfer theorem for disjoint union and product operations, we can translate the operations on combinatorial classes in the combinatorial class $SEQ(\mathcal{A})$ into corresponding generating function, denote generating function of $SEQ(\mathcal{A})$ by $S(z)$, then we obtain,

$$\begin{aligned}
S(z) &= 1 + A(z) + A(z)^2 + A(z)^3 + \cdots \\
&= \frac{1}{1 - A(z)}
\end{aligned}$$

Remark: we know $\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots$ the series converges at all values except for $x = 1$. We have assumed no object of size '0' in the combinatorial class $\mathcal{A}$, therefore, in its Generating function $A(z)$, there is no term $z^0 = 1$, therefore we can substitute $A(z)$ to x, we can obtain $\frac{1}{1-A(z)}$.

Next, from the definition of symbolic operation on generating functions, we gave a definition 25 of symbolic path transformation as follows.

Definition 25 *(Symbolic path transformation)*
Let share price sequence/path be denoted by $\{S_i\}$, construct OGF: $S(z) = \sum_j S_j z^j$.
Then, define a symbolic transform OGF, and called $\hat{S}(z) = \frac{1}{1-S(z)}$ and this defines a symbolic transformed path.

3.4   Examples of counting using symbolic method

We have discussed dissection of symbolic method in the Section 3.3, in this section, we will use symbolic method to count common fundamental combinatorial classes.

Firstly, count binary tree with two different definition of size function.

Secondly, count general trees and forests.

Thirdly, count unrestricted lattice paths in 1 dimension and 2 dimensions respectively.

Fourthly, count restricted lattice paths from (0,0) to (n,n) with the restriction that never going above the diagonal line.

3.4.1   Count Binary Trees with different size function

Motivated by Flajolet and Sedgewick, along the lines of description of the generating method of enumeration of binary trees in [14, Chapter 3.8], applying the symbolic construction of binary trees described on page 228 in [14, Chapter 5.2], we give the following Question1, Question2, and present the solution in an organised way.

Question1: How many binary trees with N internal nodes?

Answer:

1st,   Starting from a definition of a combinatorial class of binary trees.

> **Definition 26** *[14, Page 258] A binary tree is an external node or an internal node attached to an ordered pair of binary trees called the left subbinary tree and the right subbinary tree of that internal node. Denote the combinatorial class of binary trees by T.*

2nd,   Define a notion of size function for each object in the combinatorial class of binary trees T.

From the question, we give each binary tree $t \in T$ a size function,

$$|t| \equiv \text{ the number of } \textbf{internal nodes} \text{ in a binary tree } \quad t \in T$$

3rd, From the question, we would like to find the number of binary trees with N internal nodes, so, define

$$T_N \equiv \text{ the number of binary trees with } N \text{ internal nodes}$$

4th, By symbolic method, define a generating function(GF) for the combinatorial class of binary tree; its GF is the sum over all members of the class of binary trees; here, we use ordinary generating function(formal power series)

$$T(z) \stackrel{\text{define}}{\equiv} \sum_{t \in T} z^{|t|}$$

Each object of size N corresponds to each term z^N in the GF, collect all the terms with the same size exponent to expose counts T_N, that is, the number of binary trees with N internal nodes, therefore, the definition of GF for the combinatorial class is

$$T(z) \stackrel{\text{define}}{\equiv} \sum_{t \in T} z^{|t|}$$
$$= \sum_{N \geq 0} T_N z^N$$

5th, Define a combinatorial construction for the combinatorial class of binary trees.

The construction process can be done in two steps.

Firstly, define basic building blocks for the combinatorial class: (1) by the definition of binary trees class, there are two types of basic blocks; one is external node, denoted by '□', the other is internal node, denoted by '●'. (2) Based on the definition of size function,

$$|t| \equiv \text{the number of internal nodes in } t \in T,$$

we classify the two building blocks into atomic class and neutral class respectively; the atomic class contains an internal node of size 1, which contributes size '1' to the size exponent of z in the GF ($z^{|t|}$), denote the atomic class by $Z_\bullet$. Its generating function is $z^1 = z$. The neutral class, denoted by $Z_\square$, contains an external node of size '0', which has no contribution to the size of an object $t \in T$, and contributes size '0' to the size exponent of z, its generating function is $z^0 = 1$. Secondly, make the combinatorial construction of the combinatorial class.

Based on the definition of the combinatorial class, define the class in terms of basic building blocks and other combinatorial classes by using decomposition techniques and three common operations on combinatorial classes mentioned in the Section 3.4.2.

Here, by definition of the class, the class of binary trees can be a class which contains an external node or a class which contains an internal node followed by an ordered pair of two combinatorial class of binary trees.

After the two steps of the combinatorial construction, the construction can be obtained as follows,

$$T = Z_\square + T \times Z_\bullet \times T$$

6th, Apply the transfer theorem between operations on combinatorial classes and operations on generating functions. Here, we use ordinary generating function(OGF), we obtain the equation:

$$T(z) = 1 + zT(z)^2$$

$$\begin{array}{c} \text{solve} \\ \text{quadratic equation} \\ \Longrightarrow \end{array} \qquad T(z) = \frac{1 \pm \sqrt{1-4z}}{2z}$$

Question: we only need one generating function, how can we decide which one to choose?

Answer:

$$zT(z) = \frac{1 \pm \sqrt{1 - 4z}}{2}$$

take $z = 0$, then $LHS = 0$, RHS should also equal to 0, when $z = 0$.

So, we choose $T(z) = \frac{1-\sqrt{1-4z}}{2z}$.

7th, Finally, expand the coefficient of GF $T(z)$ to obtain the explicit formula for the number of binary trees with N internal nodes, T_N .

from

$$zT(z) = \frac{1 - \sqrt{1 - 4z}}{2},$$

we get,

$$[z^N]T(z) = [z^{N+1}]\left(\frac{1 - \sqrt{1 - 4z}}{2}\right)$$

Expand via generalized binomial theorem,

$$\begin{aligned} \frac{1 - \sqrt{1 - 4z}}{2} &= \frac{1}{2} - \frac{1}{2}(1 - 4z)^{\frac{1}{2}} \\ &= \frac{1}{2} - \frac{1}{2}\sum_{N \geq 0}\left(\binom{\frac{1}{2}}{N}(-4z)^N\right) \\ &= -\frac{1}{2}\sum_{N \geq 1}\left(\binom{\frac{1}{2}}{N}(-4)^N z^N\right) \end{aligned}$$

Therefore,

$$
T_N \quad = \quad [z^N]T(z) = [z^{N+1}]\left(\frac{1-\sqrt{1-4z}}{2}\right)
$$

$$
= \quad -\frac{1}{2}\binom{\frac{1}{2}}{N+1}(-4)^{N+1}
$$

$$
= \quad -\frac{1}{2}\cdot\frac{\frac{1}{2}(\frac{1}{2}-1)(\frac{1}{2}-1)\ldots(\frac{1}{2}-N)(-4)^N\cdot(-4)}{(N+1)!}
$$

$$
= \quad \frac{(-\frac{1}{2})(-\frac{3}{2})(-\frac{5}{2})\ldots(-\frac{2N-1}{2})(-4)^N}{(N+1)!}
$$

$$
\underset{\substack{\text{Distribute}_{(-2)^N}\\\text{among factors}}}{=} \quad \frac{1\cdot3\cdot5\cdots(2N-1)\cdot2^N}{(N+1)!}
$$

$$
= \quad \frac{1}{N+1}\cdot\frac{1\cdot3\cdot5\cdots(2N-1)}{N!}\cdot\frac{2\cdot4\cdot6\cdots(2N)}{1\cdot2\cdot3\cdots N}
$$

$$
= \quad \frac{1}{N+1}\binom{2N}{N}
$$

Question2: How many binary trees with N external nodes?

Answer:

1st, Starting from a definition of a combinatorial class of binary trees. The definition is the same as the 1st step in Question1. Again, denote the class of binary trees by T.

2nd, Define a notion of size function of each object in the combinatorial class of binary trees. To distinguish Question2's size function with Question1, from the Question2, we define,

$$\boxed{t} \equiv \text{ the number of external nodes in a binary tree } t \in T$$

3rd, From the question, we would like to find the number of binary trees with N external nodes. Define,

$$T_N^{\square} \equiv \text{the number of binary trees with } N \text{ external nodes}$$

4th, By symbolic method, define a GF for the combinatorial class of binary tree. to distinguish with Question1, sum over all members of the binary tree,and its GF is defined by

$$
T^{\square}(z) \quad \overset{\text{define}}{=} \quad \sum_{t\in T} z^{\boxed{t}}
$$

$$
\underset{\substack{\text{collect all terms with}\\\text{the same exponent}}}{=} \quad \sum_{N\geq 0} T_N^{\square}\cdot z^N
$$

5th, Combinatorial Construction; the construction process can be done in two steps.

Firstly, define basic building blocks for the combinatorial class: (1) the definition of the binary tree class is the same as Question1, there are also two types of basic blocks, that is, external node($\square$) and internal node($\bullet$) .

(2) based on the definition of size function in Question2,

$$\boxed{t} \equiv \text{ the number of external nodes in a binary tree } t \in T,$$

External node and internal node builds up atomic class and neutral class respectively.

Atomic class, denoted by $Z_\square$, contains an external node of size 1, which contributes size '1' to the exponent of z in the GF($z^{\boxed{t}}$), and its generating function is $z^1 = z$.

Neutral class, denoted by $Z_\bullet$, contains an internal node, which does not contribute to the size of an object $t \in T$ and contributes '0' to the exponent of 'z', its generating function is $z^0 = 1$.

Remark: No empty class, so, in the construction of the class, the building block definitely contribute a term in the GF. Neutral class which contains an object of size '0' , does not contribute to the size of an object $t \in T$ but do contribute a term in the GF, that is constant term $z^0 = 1$. However, empty class does not contribute to a term in the GF.

Secondly, make combinatorial construction by defining the class of binary tree in terms of basic building blocks and other combinatorial classes. Similarly, use decomposition techniques and three common operations for combinatorial classes.

In Question2, the class of binary tree can either be a class which contains an external node, or be a class which contains an internal node followed by an ordered pair of two combinatorial classes of binary trees.

After the two steps of making combinatorial construction, we obtain construction:

$$T = Z_\square + T \times Z_\bullet \times T$$

Remark: it is observed that the combinatorial construction is the same as the Question1 in this section.

6th, Apply the transfer theorem from operations on combinatorial classes to operations on generating functions. Here, we use OGF and obtain OGF equations as follows,

$$T^\square(z) = z + T^\square(z)^2$$

solve the quadratic,

$$T^\square(z) = \frac{1 \pm \sqrt{1 - 4z}}{2}$$

Recall in the Question1,

$$zT(z) = \frac{1 \pm \sqrt{1 - 4z}}{2}$$

Therefore, we can find the GF relationship between Question1 and Question2,

$$T^{\square}(z) = zT(z)$$

then,

$$
\begin{aligned}
[z^N]T^{\square}(z) &= [z^{N-1}]T(z) \\
&= \frac{1}{N}\binom{2(N-1)}{N-1}
\end{aligned}
$$

Remark: from the result of Question1, we know that

$$
\begin{array}{ccc}
\text{the number of binary trees with} & & \text{the number of binary trees with} \\
N \quad \text{external nodes} & = & N-1 \quad \text{internal nodes}
\end{array}
$$

It is noted that the above result is consistent with the following lemma that states the relationship between ($\sharp$ external nodes) and ($\sharp$ internal nodes) in a binary tree.

Lemma 27 *[14, Page 260] The number of external nodes in any binary tree is exactly one greater than the number of internal nodes.*

Proof: Let

$$
\begin{aligned}
e &\equiv \sharp \quad \text{of external nodes in a binary tree,} \\
i &\equiv \sharp \quad \text{of internal nodes in a binary tree}
\end{aligned}
$$

Key idea: count $\sharp$ of edges(links) in a binary tree in two different ways.

1st, By definition of a binary tree, each internal node has exactly two links 'from' it, so the number of links in the binary tree is '$2i$'.

2nd, Again, by definition of a binary tree, each node except for the root has exactly a link 'to' it, so, the number of links in the binary tree is '$i + e - 1$'.

3rd, The number of links in the binary tree by different counting method should obtain the same number. So, equating the two number we got,

$$2i = i + e - 1$$

$$\Rightarrow \qquad i = e - 1$$

3.4.2   Count Trees and Forests

In the last section, we considered a combinatorial class of binary trees; in a binary tree, no node have more than two children and can only have 0 child or 2 children. In this section, we consider general trees.

In this section, the definitions of trees, forests, and the enumeration question of forests and trees are motivated by Flajolet and Sedgewick in the book [14, Chapter 6.2], and we gave more detailed explanation using symbolic method in the solution of the enumeration question.

Definition 28 *[14, Page 261] A tree(also called a general tree) is a node(called the root) connected to a sequence of disjoint trees.*

Definition 29 *[14, Page 261] A forest is a sequence of disjoint trees.*

Remark: from the definition of a general tree and a forest, it can be observed that there is a one-to-one correspondence between the class of forests and the class of general trees. Each forest with N nodes corresponds to a general tree with $N + 1$ nodes by adding a root for the forest. That is, each tree is a root followed by a forest.

Question: [14, Theorem 6.2] How many forests and trees with N nodes respectively?

Answer:

1st,   Starting from definitions of two combinatorial classes.

Denote the class of general trees by 'G'.

Denote the class of forests by 'F'.

2nd,   Define a notion of size function for each object in the two combinatorial classes respectively.
From the question, we give each tree $g \in G$ a size function

$$|g| \equiv \text{ the number of nodes in a general tree } g \in G$$

and give each forest $f \in F$ a size function

$$|f| \equiv \text{ the number of nodes in a forest } f \in F$$

3rd,   From the question, we would like to find the number of forests and trees with N nodes, we define

$$
\begin{aligned}
G_N &\equiv & \text{the number of trees with } N \text{ nodes,} \\
F_N &\equiv & \text{the number of forests with } N \text{ nodes.}
\end{aligned}
$$

4th,   Apply symbolic method, define a GF for the two combinatorial class respectively. Their GF is the sum over all members of each combinatorial class respectively. Here, we use OGF.

$$G(z) \overset{\text{define}}{\equiv} \sum_{g \in G} z^{|g|}$$

Each object of size N corresponds to a term z^N in the GF, collect all the terms with the same size exponent to expose counts G_N, that is, the number of trees with N nodes, therefore,

$$G(z) \overset{\text{define}}{\equiv} \sum_{g \in G} z^{|g|}$$

$$= \sum_{N \geq 0} G_N z^N$$

Similarly,

$$F(z) \overset{\text{define}}{\equiv} \sum_{f \in F} z^{|f|}$$

Each object of size N corresponds to a term z^N in the GF, collect all the terms with the same size exponent to expose counts F_N, that is, the number of forests with N nodes, therefore,

$$F(z) \overset{\text{define}}{\equiv} \sum_{f \in F} z^{|f|}$$

$$= \sum_{N \geq 0} F_N z^N$$

5th, Make a combinatorial construction for the two combinatorial classes.

The construction process can be done in two steps.

Firstly, define basic building blocks. (1) By the definition of classes of forests and trees, there are only one type of basic block, that is, a node. (2) Based on the definition of their size functions, we classify 'node' into an atomic class, denoted by Z, because, for both combinatorial class(forests, trees), 'a node' contributes size '1' to their size functions. In each combinatorial class, its generating function is $z^1 = z$.

Secondly, Make the combinatorial construction for the two combinatorial classes respectively. Use the three common operations to define each combinatorial class in terms of basic building blocks and other combinatorial classes.

(1) By definition of the combinatorial class of forests, it is a sequence of disjoint (general) trees

$$F = SEQ(G)$$

Remark: F can be empty, that is, $F = E + G + G \times G + G \times G \times G \cdots$. If F is empty, G is only a node.

(2) By definition of the combinatorial class of general trees, as we have observed a one-to-one correspondence between a forest and a general tree, that is, a general tree is a node(called root) followed by a forest, we define,

$$G = Z \times F$$

6th, Apply the transfer theorem from operations on combinatorial classes to operations on generating functions. Here, we use OGF, and obtain OGF equation:

$$F(z) = \frac{1}{1 - G(z)} \qquad and \qquad G(z) = zF(z)$$

Together the two equations, obtain:

$$F(z) - zF(z)^2 = 1$$

$$\underset{\Longrightarrow}{\overset{\text{solve}}{\text{quadratic equation}}} \qquad F(z) = \tfrac{1 \pm \sqrt{1-4z}}{2z}$$

Similarly, to equate the equation, choose: $zF(z) = \tfrac{1-\sqrt{1-4z}}{2}$

Recall in Question1 of Section 3.5.1, the GF equation for binary trees with N internal nodes is

$$zT(z) = \frac{1 - \sqrt{1 - 4z}}{2}$$

Therefore, $F(Z) = T(z)$, recall the result in Question1 of Section 3.5.1, the number of forests with N nodes is,

$$\begin{aligned} F_N &= [z^N]F(z) = [z^N]T(z) = T_N \\ &= \frac{1}{N+1}\binom{2N}{N} \end{aligned}$$

Besides, we know $G(Z) = zF(z)$, so,the number of trees(general trees) with N nodes is,

$$\begin{aligned} G_N &= [z^N]G(z) = [z^{N-1}]F(z) = F_{N-1} \\ &= \frac{1}{N}\binom{2(N-1)}{N-1} \end{aligned}$$

3.4.3 Count general binary strings that contain consecutive 0's less than 3

In this section, we enumerate general binary strings with restriction that the maximum number of consecutive 0's is 2. The idea is motivated by Flajolet and Sedgewick in the book [14, Page 226-227] and take the calculation idea from [16], we gave the detailed explanation of solution using Flajolet symbolic method.

One Question
Notation:

  . '+': disjoint union for sets;

  . '×': Cartesian product for sets;

. $\mathbb{Z}_0$: the set $\{0\}$;

. $\mathbb{Z}_1$: the set $\{1\}$;

. $\mathbb{Z}_0 + \mathbb{Z}_1 = \{0\} \bigcup \{1\} = $ the set $\{0, 1\}$;

. Let A, B be two sets, $A \times B = \{(\alpha, \beta) : \alpha \in A, \beta \in B\} \overset{\text{denoted}}{\underset{by}{=}} \{\alpha\beta : \alpha \in A, \beta \in B\}$;

. Let A, B be two sets, assume A, B are disjoint, their disjoint union(sum) is defined as $A + B = $ the set of disjoint copies of objects in A and in $B = \{\gamma : \gamma \in \mathcal{A} \cup \mathcal{B}\}$; (Assume the two sets are disjoint sets, if not disjoint, work with $A \times \epsilon_1$ and $B \times \epsilon_2$)

. Let empty set denoted by $\emptyset, \emptyset = \{\epsilon\}$, $SEQ(A) = \emptyset + A + A \times A + A \times A \times A + \ldots$ Remark: The set A must not contain empty element, otherwise we cannot count the set $SEQ(A)$ by a specific characteristic, for example, if $\mathcal{A}$ has empty elements in the set $SEQ(A)$ that has length 'n', then, a_n would be infinite number, it is nonsense to count $SEQ(A)$ by length.

. $SEQ(\mathbb{Z}_0) = SEQ(\{0\}) = \{\epsilon, 0, 00, \ldots\} =$
$$= \emptyset + \{0\} + \{00\} + \ldots;$$

. $SEQ(\mathbb{Z}_1 \times SEQ(\mathbb{Z}_0)) = SEQ(\{1\} \times \{\epsilon, 0, 00, \ldots\}) = SEQ(\{1, 10, 100, \ldots\})$

Represent a set by generating function by counting the length of each element in the set.

In the following, length$(\alpha) \overset{\text{denoted}}{\underset{by}{=}} |\alpha|$

$$GF(A) = \sum_{\alpha \in A} a_{|\alpha|} \cdot z^{|\alpha|} = \sum_{n=0}^{\infty} a_n \cdot z^n$$

Where $a_{|\alpha|} = a_n = a_{\text{length}(\alpha)} = \sharp$ of elements in the set A that has length$(\alpha) = n$.

$$GF(B) = \sum_{\beta \in B} b_{|\beta|} \cdot z^{|\beta|} = \sum_{n=0}^{\infty} b_n \cdot z^n$$

Where $a_{|\beta|} = b_n = a_{\text{length}(\beta)} = \sharp$ of elements in the set B that has length$(\beta) = n$.

$$GF(A + B) = \sum_{\gamma \in (A+B)} c_{|\gamma|} \cdot z^{|\gamma|} = \sum_{\gamma \in A} c_{|\gamma|} \cdot z^{|\gamma|} + \sum_{\gamma \in B} c_{|\gamma|} \cdot z^{|\gamma|}$$
$$= \sum_{n=0}^{\infty} a_n \cdot z^n + \sum_{n=0}^{\infty} b_n \cdot z^n$$
$$= GF(A) + GF(B)$$

Where A+B are disjoint union of a set A and a set B; it is the set of disjoint copies of objects in A and in B. (Assume the two sets are disjoint sets, if not disjoint, work with $A \times \epsilon_1$ and $B \times \epsilon_2$).

$$
\begin{aligned}
GF(A \times B) = \sum_{\gamma \in A \times B} c_{|\gamma|} \cdot z^{|\gamma|} \quad &= \sum_{\alpha\beta \in A \times B} c_{|\alpha|+|\beta|} \cdot z^{|\alpha|+|\beta|} \\
&= \sum_{n=0}^{\infty} \Big(\sum_{k_1+k_2=n} c_{k_1+k_2} \Big) \cdot z^n \\
&= \sum_{n=0}^{\infty} \Big(\sum_{k_1+k_2=n} a_{k_1} \cdot b_{k_2} \Big) \cdot z^{k_1} \cdot z^{k_2} \\
&= \sum_{n=0}^{\infty} \Big(\sum_{k=0}^{n} a_k \cdot b_{n-k} \Big) \cdot z^k \cdot z^{n-k} = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} a_k z^k \cdot b_{n-k} z^{n-k} \\
&= \sum_{k=0}^{\infty} a_k z^k \cdot \sum_{n=k}^{\infty} b_{n-k} z^{n-k} = GF(A) \cdot GF(B)
\end{aligned}
$$

Then, use the above generating function for the sets by counting the length of each element in the sets, find the generating function for the set $SEQ(A) = \emptyset + A + A \times A + A \times A \times A + \ldots$, that is,

$$
\begin{aligned}
GF(SEQ(A)) &= GF(\emptyset + A + A \times A + A \times A \times A + \ldots) \\
&= GF(\emptyset) + GF(A) + (GF(A))^2 + (GF(A))^3 + \ldots
\end{aligned}
$$

calculate $GF(\emptyset)$, that is,
length$(\alpha) \overset{\text{denoted}}{\underset{by}{=}} |\alpha|$

$$
GF = \sum_{n=0}^{\infty} a_n \cdot z^n \overset{\text{empty set has no elements}}{\underset{length\ is\ 0,\ let\ a_0=1}{=}} 1
$$

$$
\begin{aligned}
\implies GF(SEQ(A)) &= 1 + GF(A) + (GF(A))^2 + (GF(A))^3 + \ldots \\
&= \frac{1}{1 - GF(A)}
\end{aligned}
$$

Then, use the above notation and formulas to solve the next question.

Question: Count general binary strings that contain consecutive 0's less than 3.

Idea: Characterize general bitstrings by grouping them into blocks staring with 1. A bitstring can be constructed from zero blocks starting with 1 or from more blocks starting with 1 followed by zero or more 0's. The bitstring may be preceded by zero or more 0's.
Firstly, use this characterize to construct general binary string and find its generating function. Secondly, construct the restricted binary string and use generating function to count the question.

Answer:

Let $\mathbb{Z}_0 = \{0\}$, $\mathbb{Z}_1 = \{1\}$, $\mathbb{Z}_0 + \mathbb{Z}_1 = \{0,1\}$
then, zero or more 0's is the set

$$SEQ(\mathbb{Z}_0) = \{\epsilon, 0, 00, 000, \ldots\}$$

block starting with 1 followed by zero or more 0's is the set

$$\begin{aligned}
\mathbb{Z}_1 \times SEQ(\mathbb{Z}_0) &= \{1\} \times \{\epsilon, 0, 00, 000, \ldots\} \\
&= \{1, 10, 100, 1000, \ldots\} \\
&\overset{\text{denoted}}{\underset{by}{=}} \mathbb{Z}_1 SEQ(\mathbb{Z}_0)
\end{aligned}$$

Remark: the block does not contain empty element, so we can count the set of zero or more such blocks, that is, the set $SEQ(\mathbb{Z}_1 \times SEQ(\mathbb{Z}_0))$.

Then, a general bitstring can be constructed by (zero or more 0's) followed by (zero or more blocks starting with 1 followed by zero or more 0's), that is ,

$$\begin{aligned}
B &= SEQ(\mathbb{Z}_0) \times SEQ(\mathbb{Z}_1 \times SEQ(\mathbb{Z}_0)) \\
&\overset{\text{denoted}}{\underset{by}{=}} SEQ(\mathbb{Z}_0) SEQ(\mathbb{Z}_1 SEQ(\mathbb{Z}_0))
\end{aligned}$$

Find the generating function for the sets $\mathbb{Z}_0$, $\mathbb{Z}_1$, $\mathbb{Z}_0 + \mathbb{Z}_1$ by counting the length of each element in these sets as follows,

$$GF(\mathbb{Z}_0) = GF(\{0\}) = 1 \cdot z^1 = z$$

$$GF(\mathbb{Z}_1) = GF(\{1\}) = 1 \cdot z^1 = z$$

$$GF(\mathbb{Z}_0 + \mathbb{Z}_1) = GF(\{0,1\}) = a_1 \cdot z^1 = 2 \cdot z,$$

$$\text{where, } a_1 = \sharp \text{ of elements of length 1}$$

Then, find the generating function for the sets $SEQ(\mathbb{Z}_0)$, $\mathbb{Z}_1 SEQ(\mathbb{Z}_0)$, $SEQ(\mathbb{Z}_1 SEQ(\mathbb{Z}_0))$ and $SEQ(\mathbb{Z}_0) SEQ(\mathbb{Z}_1 SEQ(\mathbb{Z}_0))$, that is,

$$GF(SEQ(\mathbb{Z}_0)) = \frac{1}{1 - GF(\mathbb{Z}_0)} = \frac{1}{1 - GF(\{0\})} = \frac{1}{1 - z},$$

$$GF(\mathbb{Z}_1 SEQ(\mathbb{Z}_0)) = GF(\mathbb{Z}_1) \cdot GF(SEQ(\mathbb{Z}_0)) = GF(\{1\}) \cdot GF(SEQ(\mathbb{Z}_0)) = \frac{z}{1 - z},$$

$$\begin{aligned}
GF(SEQ(\mathbb{Z}_1 SEQ(\mathbb{Z}_0))) &= \frac{1}{1 - GF(\mathbb{Z}_1 SEQ(\mathbb{Z}_0))} \\
&= \frac{1}{1 - \frac{z}{1-z}} = \frac{1 - z}{1 - 2z},
\end{aligned}$$

$$
\begin{aligned}
GF(SEQ(\mathbb{Z}_0)SEQ(\mathbb{Z}_1 SEQ(\mathbb{Z}_0))) &= GF(SEQ(\mathbb{Z}_0)) \cdot GF(SEQ(\mathbb{Z}_1 SEQ(\mathbb{Z}_0))) \\
&= \frac{1}{1-z} \cdot \frac{1-z}{1-2z} \\
&= \frac{1}{1-2z} = \sum_{N=0}^{\infty} (2z)^N
\end{aligned}
$$

Secondly, find restricted binary string with restriction containing consecutive 0's less than 3. It is obvious that the restriction has no effect on element 1, it only apply to the object (zero or more 0's), that is, the set $SEQ(\mathbb{Z}_0) = SEQ(\{0\})$, denote the set of such restricted binary string by $\mathcal{G}$.

Therefore, $\mathcal{G}$ can be constructed by:

$$
\mathcal{G} = (\emptyset + \mathbb{Z}_0 + \mathbb{Z}_0 \times \mathbb{Z}_0) \times SEQ(\mathbb{Z}_1 \times (\emptyset + \mathbb{Z}_0 + \mathbb{Z}_0 \times \mathbb{Z}_0))
$$

then, find the generating function for the set $(\emptyset + \mathbb{Z}_0 + \mathbb{Z}_0 \times \mathbb{Z}_0) = \{\epsilon, 0, 00\}$ and $\mathcal{G}$,

$$
\begin{aligned}
GF(\emptyset + \mathbb{Z}_0 + \mathbb{Z}_0 \times \mathbb{Z}_0) &= GF(\{\epsilon, 0, 00\}) \\
&= a_0 z^0 + a_1 z^1 + a_2 z^2 + a_3 z^3 + \ldots \\
&= 1 + 1 \cdot z^1 + 1 \cdot z^2 \\
&= 1 + z + z^2
\end{aligned}
$$

$$
\begin{aligned}
GF(\mathcal{G}) &= GF(\emptyset + \mathbb{Z}_0 + \mathbb{Z}_0 \times \mathbb{Z}_0) \cdot GF(SEQ(\mathbb{Z}_1 \times (\emptyset + \mathbb{Z}_0 + \mathbb{Z}_0 \times \mathbb{Z}_0))) \\
&= (1 + z + z^2) \cdot \frac{1}{1 - GF(\mathbb{Z}_1 \times (\emptyset + \mathbb{Z}_0 + \mathbb{Z}_0 \times \mathbb{Z}_0))} \\
&= (1 + z + z^2) \cdot \frac{1}{1 - GF(\mathbb{Z}_1) \times GF(\emptyset + \mathbb{Z}_0 + \mathbb{Z}_0 \times \mathbb{Z}_0))} \\
&= (1 + z + z^2) \cdot \frac{1}{1 - z \times (1 + z + z^2)} \\
&= \frac{1 + z + z^2}{1 - z \times (1 + z + z^2)} \\
&\overset{\underset{\text{calculation}}{\text{technique}}}{=} \frac{\frac{1-z^3}{1-z}}{1 - z \times \frac{1-z^3}{1-z}} \\
&= \frac{1 - z^3}{1 - 2z + z^4} \\
&= (1 - z^3) \cdot \sum_{k \geq 0} (2z - z^4)^k \\
&= (1 - z^3) \cdot \sum_{k \geq 0} z^k (2 - z^3)^k \\
&= (1 - z^3) \cdot \sum_{k \geq 0} z^k \sum_{l=0}^{k} \binom{k}{l} (-z^3)^l \cdot 2^{(k-l)}
\end{aligned}
$$

Then, take the coefficient of Z^N in the generating function $GF(\mathcal{G})$ as follows,

$$[z^N]GF(\mathcal{G}) = [z^N]G(z) = [z^N](1 - z^3) \cdot \sum_{k \geq 0} z^k \sum_{l=0}^{k} \binom{k}{l}(-z^3)^l \cdot 2^{(k-l)}$$

$$= ([z^N] - [z^{N-3}]) \sum_{k \geq 0} z^k \cdot \sum_{l=0}^{k} \binom{k}{l}(-z^3)^l \cdot 2^{(k-l)}$$

$$= \sum_{k \geq 0}^{N}([z^{N-k}] - [z^{N-3-k}]) \sum_{l=0}^{k} \binom{k}{l}(-z^3)^l \cdot 2^{(k-l)}l$$

$$\overset{\substack{N-k=3l \to l=\frac{N-k}{3},\text{(l is integer)} \\ N-3-k=3l \to l=\frac{N-k}{3}-1,\text{(l is integer)}}}{=\!=\!=} \sum_{k \geq 0, N \equiv k(mod3)}^{N} \binom{k}{\frac{N-k}{3}}(-1)^{\frac{N-k}{3}} \cdot 2^{k-\frac{N-k}{3}}$$

$$- \sum_{k \geq 0, N \equiv k(mod3)}^{N} \binom{k}{\frac{N-k}{3}-1}(-1)^{\frac{N-k}{3}-1} \cdot 2^{k-\frac{N-k}{3}+1}$$

$$= \sum_{k \geq 0, N \equiv k(mod3)}^{N} \left(\binom{k}{\frac{N-k}{3}} + 2 \cdot \binom{k}{\frac{N-k}{3}-1}\right)(-1)^{\frac{N-k}{3}} \cdot 2^{k-\frac{N-k}{3}}$$

Remark:

1. $N \equiv k(mod3) \Longleftrightarrow N - k = l \cdot 3$, where l, N, k are integers

$$\Longleftrightarrow \quad \begin{cases} N = p \cdot 3 + r, \\ k = q \cdot 3 + r, \end{cases}$$

2. Recall: definition of congruent modulo(modulus of the congruence).

3. As $N \equiv k(mod3)$ means $N - k = l \cdot 3, \ 0 \leq l \leq N$,
   the set of possible values for k; $0 \leq k \leq N$ in the above example $= \{k, k + 3, k + 2 \cdot 3, k + 3 \cdot 3, \ldots\}$, where $0 \leq k \leq N$;

4.

$$l = \frac{N - k}{3} - 1 = N - k = 3 \cdot (l + 1)$$
$$= N \equiv k(mod3)$$
$$= l = \frac{N - k}{3}$$

this is consistent with the calculation in the example that if we want the power of z be integer ,we must take $\frac{N-k}{3}$ be integer,that makes $l = \frac{N-k}{3}$ and $l = \frac{N-k}{3} - 1$ be integers.

5. The method can be extended into the similar question as counting the number of binary strings of fixed length 'n' that contains number of consecutive 0's less than any number 'm'.
   It can be solved using the same method as in the above example.

3.4.4   Count using generating functions

In this section, we continue to use the Flajolet symbolic method to construct the generating function to solve some enumeration problems. The symbolic method is motivated by Flajolet and Sedgewick in the book [14] and [15]. The enumeration problem is taken from the page 3-5 in the paper [42].

We use the calculation methods taken from [42, Problem 2] and [42, Problem 3], however, we provide more detailed knowledge of the explanation of the calculation method, in addition, in the example2, we added the derivation of generating function using Flajolet symbolic method. The calculation idea of example2 can also be referred to [43].

Example 3.4.1 *One variable [42, Problem 2]*

Question: Given a generating function $\mathcal{F}(x) = \sum_{k=0}^{\infty} f_k \cdot x^k$ if it is a polynomial, then $f_k = 0$,   when $k > deg(\mathcal{F}(x))$,

Find the sum

$$A = f_0 + f_n + f_{2n} + \cdots = \sum_{\substack{k=0 \\ n|k}}^{\infty} f_k$$

Remark:

1. $n|k \iff \exists\, l \in \mathbb{Z},\ k = l \cdot n$
   $n \nmid k \iff \exists\, l, m \in \mathbb{Z},\ k = l \cdot n + m\ \ (\, k \equiv m\ (mod\ n)\,)$

2. Define $w = \exp(i \cdot \frac{2\pi}{n})$ as one of the n-th roots of unity and obviously $w \neq 1$, then

$$1 + w + w^2 + \ldots + w^{n-1} = \frac{1 - w^n}{1 - w}$$

   Note: then, $w^k,\ k = 1, 2, \ldots, n - 1$ is also one of the n-th roots of unity and $w^k \neq 1$, thus,

$$1 + w^k + (w^k)^2 + \ldots + (w^k)^{n-1} = \frac{1 - (w^k)^n}{1 - w^k} = 0$$

Answer:

Let $w = \exp(i \cdot \frac{2\pi}{n})$ be one of the n-th roots of unity, then

$$\mathcal{F}(1) + \mathcal{F}(w) + \ldots + \mathcal{F}(w^{n-1}) = \sum_{k=0}^{\infty} f_k + \sum_{k=0}^{\infty} f_k w^k + \ldots + \sum_{k=0}^{\infty} f_k (w^{n-1})^k$$

$$= \sum_{k=0}^{\infty} f_k (1 + w^k + (w^k)^2 + \ldots + (w^k)^{n-1})$$

$$= (\sum_{\substack{k=0 \\ n|k}}^{\infty} + \sum_{\substack{k=0 \\ n \nmid k}}^{\infty}) f_k (1 + w^k + (w^k)^2 + \ldots + (w^k)^{n-1})$$

$$= \sum_{m=0}^{\infty} f_{n \cdot m} (1 + w^{n \cdot m} + (w^2)^{n \cdot m} + \ldots + (w^{n-1})^{n \cdot m}))$$

$$+ \sum_{m=0}^{\infty} \sum_{r=1}^{n-1} f_{n \cdot m + r} (1 + w^{n \cdot m} w^r + (w^2)^{n \cdot m} w^r + \ldots + (w^{n-1})^{n \cdot m} w^r)$$

$$= \sum_{m=0}^{\infty} n \cdot f_{n \cdot m} + \sum_{m=0}^{\infty} \sum_{r=1}^{n-1} f_{n \cdot m + r} (1 + w^r + (w^r)^2 + \ldots + (w^r)^{n-1})$$

$$\overset{w^r \neq 1}{=\!=\!=} \sum_{m=0}^{\infty} n \cdot f_{n \cdot m} + \sum_{m=0}^{\infty} \sum_{r=1}^{n-1} f_{n \cdot m + r} (\frac{1 - (w^r)^n}{1 - w^r})$$

$$= \sum_{m=0}^{\infty} n \cdot f_{n \cdot m} + 0 = \sum_{\substack{k=0 \\ n|k}}^{\infty} n \cdot f_k = n \cdot \sum_{\substack{k=0 \\ n|k}}^{\infty} f_k$$

so,

$$\sum_{\substack{k=0 \\ n|k}}^{\infty} f_k = \frac{1}{n}(\mathcal{F}(1) + \mathcal{F}(w) + \ldots + \mathcal{F}(w^{n-1}))$$

Example 3.4.2 *Two variables [42, Problem 3]*

Question:
Let p be an odd prime number, Find the $\sharp$ of subsets A of the set $\{1, 2, \ldots, 2p\}$, that satisfying 1) the subset A has p elements and the sum of all elements in A is divisible by p.

Idea:
If use generating function method, Let $g_{n,k} = \sharp$ of subsets of the set $\{1, 2, \ldots, 2p\}$ that have size $'k'$ and sum $'n'$. Define the generating function

$$\mathcal{G}(x, y) = \sum_{n,k=0}^{\infty} g_{n,k} \cdot x^n y^k$$

Note: p is an odd prime number, so $p \geq 1$.
then, $g_{0,p} = $ the number of subsets of size $'p'$ and sum 0 $= 0$

Then, the question is to find

$$g_{0,p} + g_{p,p} + g_{2p,p} + \ldots + g_{np,p} + \ldots = g_{p,p} + g_{2p,p} + \ldots + g_{np,p} + \ldots$$

Answer:

Firstly, to find the generating function $\mathcal{G}(x, y)$.

From the set $\{1, 2, \ldots, 2p\}$ to construct a subset A, for any element $m \in \{1, 2, \ldots, 2p\}$,

$$A = \{1, \epsilon\} \times \{2, \epsilon\} \times \{3, \epsilon\} \times \ldots \times \{2p, \epsilon\}$$

where $\epsilon = $ nothing $=$ empty element.

For any element $m \in \{1, 2, \ldots, 2p\}$, $'m'$ affect a subset by increasing the sum by m and the size by 1. So,

$$GF(\{m\}) = 1 \cdot x^m y^1$$

For empty element ϵ, ϵ affect a subset by increasing the sum by 0 and the size by 0. So,

$$GF(\{\epsilon\}) = 1 \cdot x^0 y^0 = 1$$

Thus,

$$GF(\{m, \epsilon\}) = GF(\{m\} + \{\epsilon\}) = GF(\{m\}) + GF(\{\epsilon\}) = x^m y + 1$$

where $'+'$ is disjoint union notation.

Then,

$$\begin{aligned} GF(A) &= GF(\{1, \epsilon\}) \cdot GF(\{2, \epsilon\}) \ldots GF(\{2p, \epsilon\}) \\ &= (1 + xy) \cdot (1 + x^2 y) \cdot (1 + x^3 y) \ldots (1 + x^{2p} y) = \mathcal{G}(x, y) \end{aligned}$$

Secondly, from the generating function $\mathcal{G}(x, y)$ to extract the coefficients

$$g_{0,p} + g_{p,p} + g_{2p,p} + g_{3p,p} + \ldots$$

which is the coefficients of y^p and $x^{l \cdot p}$, $l \in \mathbb{Z}^+$ in the function $\mathcal{G}(x, y)$. where,

$$\begin{aligned} \mathcal{G}(x, y) &= \sum_{n,k=0}^{\infty} g_{n,k} \cdot x^n y^k = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} g_{n,k} \cdot x^n y^k \\ &= \sum_{k=0}^{\infty} \mathcal{G}_k(x) \cdot y^k, \quad \text{where} \quad \mathcal{G}_k(x) = \sum_{n=0}^{\infty} g_{n,k} x^n \end{aligned}$$

Idea:　To extract coefficients, firstly, fix k, try to extract

$$g_{0,k} + g_{p,k} + g_{2p,k} + g_{3p,k} + \ldots$$

Then, find the generating function with variable y, that is,

$$\sum_{k=0}^{\infty} (g_{0,k} + g_{p,k} + g_{2p,k} + g_{3p,k} + \ldots) y^k$$

after that, then, extract the coefficient of terms containing y^p from the above generating function, thus, we get the coefficients

$$g_{0,p} + g_{p,p} + g_{2p,p} + g_{3p,p} + \ldots$$

The calculation procedure are as follows,

Let $w = e^{i \cdot \frac{2\pi}{p}}$ be one of the p-th roots of unity, where $p \geq 1$, it is obviously, $w \neq 1$, then,

$$g_{0,k} + g_{p,k} + g_{2p,k} + g_{3p,k} + \ldots = \sum_{\substack{n=0 \\ p|n}}^{\infty} g_{n,k}$$

$$= \frac{1}{p}(\mathcal{G}_k(1) + \mathcal{G}_k(w) + \mathcal{G}_k(w^2) + \ldots + \mathcal{G}_k(w^{p-1}))$$

Therefore, the generating function with variable y is,

$$\sum_{k=0}^{\infty}(g_{0,k} + g_{p,k} + g_{2p,k} + g_{3p,k} + \ldots)y^k = \frac{1}{p}(\mathcal{G}_k(1) + \mathcal{G}_k(w) + \mathcal{G}_k(w^2) + \ldots + \mathcal{G}_k(w^{p-1}))y^k$$

$$= \frac{1}{p}(\mathcal{G}(1,y) + \mathcal{G}(w,y) + \mathcal{G}(w^2,y) + \ldots + \mathcal{G}(w^{p-1},y))$$

Then, the question is left to calculate : $\mathcal{G}(w^l, y)$, where $0 \leq l \leq p-1$

From
$$\mathcal{G}(x, y) = (1 + xy) \cdot (1 + x^2 y) \cdot (1 + x^3 y) \ldots (1 + x^{2p} y),$$

we get,

$$\mathcal{G}(w^l, y) = (1 + w^l y) \cdot (1 + (w^l)^2 y) \cdot (1 + (w^l)^3 y) \ldots (1 + (w^l)^{2p} y)$$

$$\underset{\substack{w=e^{i \cdot \frac{2\pi}{p}} \\ w^p=1 \\ w^{lp}=1}}{=\!=\!=} (1 + w^l y) \cdot (1 + (w^l)^2 y) \cdot (1 + (w^l)^3 y) \ldots (1 + (w^l)^p y) \cdot$$

$$(1 + w^{lp+l} y) \cdot (1 + w^{lp+2l} y) \cdot (1 + w^{lp+3l} y) \ldots (1 + w^{lp} \cdot w^{lp} y)$$

$$=\!= \left((1 + w^l y) \cdot (1 + (w^l)^2 y) \cdot (1 + (w^l)^3 y) \ldots (1 + (w^l)^p y) \right)^2$$

$$\underset{w^{lp}=1}{=\!=} \left((1 + y) \cdot (1 + w^l y) \cdot (1 + (w^l)^2 y) \ldots (1 + (w^l)^{p-1} y) \right)^2$$

Remark:

1. If $l = 1, 2, \ldots, p-1$, as $w = e^{i \cdot \frac{2\pi}{p}} \neq 1$, then, $w^l \neq 1$, then, $1, w^l, (w^l)^2, \ldots, (w^l)^{p-1}$ are p different p-th roots of unity.

   When $w = e^{i \cdot \frac{2\pi}{p}} \neq 1$, it is obvious that $1, w, w^2, \ldots, w^{p-1}$ are p different p-th roots of unity.

2. Let $x^n - y^n$ be a polynomial of x with degree $'n'$, if there are $'n'$ different roots $x_1, x_2, \ldots, x_n$ of the polynomial, Then,

   $$x^n - y^n = const(x - x_1)(x - x_2) \ldots (x - x_n) = (x - x_1)(x - x_2) \ldots (x - x_n),$$

   where, the constant number is the coefficient of x^n.

The n-th roots of the polynomial can be found as $x = (y^n)^{\frac{1}{n}} \cdot e^{i \cdot \frac{0+2k\pi}{n}}$, where $k = 0, 1, 2, \ldots, n-1$. If define $w = e^{i \cdot \frac{2\pi}{n}} \neq 1$, the n-th different roots of the polynomial $x^n - y^n$ are $y, wy, w^2 y, \ldots, w^{n-1}y$ When $l = 1, 2, \ldots, p-1$,the n-th different roots of the polynomial $x^n - y^n$ can also be $y, w^l y, (w^l)^2 y, \ldots, (w^l)^{n-1} y$, where $l = 1, 2, \ldots, n-1$. So,

$$
\begin{aligned}
x^n - y^n &= (x-y)(x-w^l y)(x-(w^l)^2 y)\ldots(x-(w^l)^{n-1}y) \\
&= (x-y)(x-wy)(x-w^2 y)\ldots(x-w^{n-1}y)
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
x^n + y^n &= (x+y)(x+w^l y)(x+(w^l)^2 y)\ldots(x+(w^l)^{n-1}y) \\
&= (x+y)(x+wy)(x+w^2 y)\ldots(x+w^{n-1}y) \\
&= (x+wy)(x+w^2 y)\ldots(x+w^{n-1}y)(x+w^n y)
\end{aligned}
$$

Then, come back to the equations for $\mathcal{G}(w^l, y)$, it equals,

$$
\mathcal{G}(w^l, y) = \left((1+w^l y) \cdot (1+(w^l)^2 y) \cdot (1+(w^l)^3 y)\ldots(1+(w^l)^p y) \right)^2
$$

$$
\underset{0 \leq l \leq p-1}{\overset{w^{lp}=1}{=\!=\!=\!=}} \left((1+y) \cdot (1+w^l y) \cdot (1+(w^l)^2 y)\ldots(1+(w^l)^{p-1}y) \right)^2
$$

when $l = 0$,

$$
\mathcal{G}(w^l, y) = (1+y)^{2p}
$$

when $1 \leq l \leq p-1$,

$$
\mathcal{G}(w^l, y) = (1^p + y^p)^2 = (1 + y^p)^2
$$

Back to the generating function with variable y is

$$
\sum_{k=0}^{\infty} (g_{0,k} + g_{p,k} + g_{2p,k} + g_{3p,k} + \ldots)y^k = \frac{1}{p}\left((1+y)^{2p} + (p-1)(1+y^p)^2 \right)
$$

Then, take the coefficients of y^p from

$$
\sum_{k=0}^{\infty} \sum_{\substack{n=0 \\ p|n}}^{\infty} g_{n,k} \cdot y^k = \frac{1}{p}\left((1+y)^{2p} + (p-1)(1+y^p)^2 \right)
$$

thus,

$$
\begin{aligned}
[y^p]\frac{1}{p}\left((1+y)^{2p} + (p-1)(1+y^p)^2 \right) &= \frac{1}{p}\left(\binom{2p}{p} + (p-1)[y^p]\binom{2}{n}(y^p)^n \right) \\
&= \frac{1}{p}\left(\binom{2p}{p} + (p-1)\binom{2}{1} \right) \\
&= \frac{1}{p}\left(\binom{2p}{p} + 2(p-1) \right)
\end{aligned}
$$

3.5 Counting permutations and statistics of permutations

3.5.1 Permutations

In enumerative combinatorics, permutations of sets and multi-sets are interesting. There are many ways to represent a set permutation combinatorially.[21, page 29]

Firstly, a set permutation can be represented as a function as follows,

For any set S, $w : [n] \rightarrow S$, is defined by $w[i] = w_i$, here, $[n] = \{1, 2, \ldots n\}$

Secondly, a set permutation can be represented as a word, that is, $w = w_1 w_2 w_3 \ldots w_n$.

For instance, there is a permutation of the set $1, 2, 3, 4$, which is $w = 3214$ (word representation). If writing the set permutation $w = 3214$ as a function, it is $w : [4] \rightarrow S$, is defined by $w[1] = w_1 = 3, w[2] = w_2 = 2, w[3] = w_3 = 1, w[4] = w_4 = 4$.

Thirdly, a set permutation can be represented as a disjoint union of its distinct cycles, and written as a disjoint cycle notation.

Definition 30 *(Disjoint cycle notation of a set permuation) [21, page 29]*

Consider a set permutation $w : S \rightarrow S$, since the set permutation is a bijection and S is a finite set, for each $x \in S$, the sequence $x, w(x), w^2(x), \ldots$ will eventually return to the element x, thus, for some unique $l \geq 1$, $w^l(x) = x$ and the elements $x, w(x), w^2(x), \ldots, w^{l-1}(x)$ are distinct. The sequence $(x, w(x), w^2(x), \ldots, w^{l-1}(x))$ is called a cycle of the permutation w of length l. Suppose the permutation w has k cycles, the permutation is represented in disjoint cycle notations as $w = C_1 C_2 \ldots C_k$.

Remarks:

1. Since the set S is a finite set, there must **exist** some integer $l \geq 1$, such that $w^l(x) = x$ and $x, w(x), \ldots, w^{l-1}(x)$ are distinct elements.

2. Since the set permutation w is a bijection, such integer l must be unique, otherwise, if the integer is not unique, to make elements in the sequence $(x, w(x), \ldots, w^l(x))$ be distinct and $w^l(x) = x$, there must exist an element $x \in S$ such that w(x) has two values, it contradicts that the set permutation w is a function.

3. Every element of the set S appears in a unique cycle of the set permutation w.

4. The cyclically shift of a cycle in the set permutation w produce a same set permutation w, that is, cyclically shift of a cycle represent the same cycle in the set permutation w. For instance, the cycles $(x, w(x), \ldots, w^{l-1}(x))$ and $(w^i(x), w^{i+1}(x) \ldots, w^{l-1}(x), x, \ldots, w^{i-1}(x))$ are considered the same.

5. The set permutation w is represented in a disjoint cycle notation as $w = C_1 \ldots C_k$, Changing the order of $C_1, C_2, \ldots C_k$ produce a same set permutation w.

For instance, if a set permutation $w : [7] \to [7]$ is defined as a function by $w(1) = 4, w(2) = 2, w(3) = 7, w(4) = 1, w(5) = 3, w(6) = 6, w(7) = 5$,
or $w = 4271365$ (as a word)

then, the set permutation $w = (14)(2)(375)(6)$ (as a disjoint cycle notation).

Let $\mathfrak{S}_n$ be a set of permutations of the set $[n] = \{1, 2, \ldots, n\}$, that is, $\mathfrak{S}_n = \{w : [n] \to [n]\}$, it is noted that the representation of a set permutation $w \in \mathfrak{S}_n$ in disjoint cycle notation; $w = C_1 \ldots C_k$ is not unique. From the items (4) and (5) in the above remarks, changing the order of $C_1, C_2, \ldots C_k$ in a set permutation w or cyclically shift of any cycle in the set permutation will produce a same permutation w. Therefore, a standard cycle notation representation is defined as follows.

Definition 31 *(Standard form of disjoint cycle notation) [21, page 30]*

Consider a set permutation has k disjoint cycles and written as $w = C_1 \ldots C_k$, a standard representation of the disjoint cycle decomposition can be obtained by requiring that

(i) Each cycle is written with its largest number first.

(ii) The distinct cycles $C_1, C_2, \ldots C_k$ are written in increasing order of their largest numbers.

For instance, the standard cycle representation of the set permutation $w = (14)(2)(375)(6)$ is $w = (2)(41)(6)(753)$.

Remarks:

1. It is noted that $(14)(2)(375)(6)$ and $(14)(2)(357)(6)$ represent different permutations, since the cycle (357) is not the cyclical shift of the cycle (375).

Fourthly, a set permutation can be represented geometrically as a digraph (a finite directed graph) as follows.

Motivated by Stanley from the discussion on page 29, 30 in the book [21, Section 1.3], we provide a definition 32 of digraph form of set permutation as follows. Besides,

using the same example on the page 30 in the book [21, Section 1.3], we give detailed algorithm how disjoint union of directed cycles is obtained from a disjoint cycle notation of a permutation.

Definition 32 *(Digraph form of Set Permutation)*
Consider a set permutation $w : S \to S$, the digraph D_w of the set permutation w is defined to be the directed graph with the vertex set $S = \{x_1, x_2, \ldots, x_n\}$ and the edge set $V = \{(x_i, x_j) : x_j = w(x_i), 1 \leq i, j \leq n\}$. For every vertex x, for $1 \leq i \leq n$, there is an edge from x_i to $w(x_i)$, the edge $(x_i, w(x_i))$ can be represented as an arrow from x_i to $w(x_i)$.

Remarks:

1. Since the set permutation is a bijection function, every vertex only has one edge pointing out and one edge pointing in.

2. The disjoint cycle decomposition of a set permutation guarantees that its directed graph D_w will be a disjoint union of directed cycles.

Example 3.5.1 *the digraph D_w of the permutation $w = (2)(41)(6)(753)$ is obtained as follows.*

Going through all distinct cycles $C_1, C_2, \ldots, C_k$ of the set permutation w with $k = 4$,

for cycle $C_1 = (2)$,

$$w(2) = 2 \iff edge: (2, 2) \iff edge: \ 2 \to 2$$

for cycle $C_2 = (41)$,

$$w(4) = 1 \iff edge: (4, 1) \iff edge: \ 4 \to 1$$
$$w(1) = 4 \iff edge: (1, 4) \iff edge: \ 1 \to 4$$

for cycle $C_3 = (6)$,

$$w(6) = 6 \iff edge: (6, 6) \iff edge: \ 6 \to 6$$

for cycle $C_4 = (753)$,

$$w(7) = 5 \iff edge: (7, 5) \iff edge: \ 7 \to 5$$
$$w(5) = 3 \iff edge: (5, 3) \iff edge: \ 5 \to 3$$
$$w(3) = 7 \iff edge: (3, 7) \iff edge: \ 3 \to 7$$

then, the digraph D_w of the permutation $w = (2)(41)(6)(753)$ is as follows,
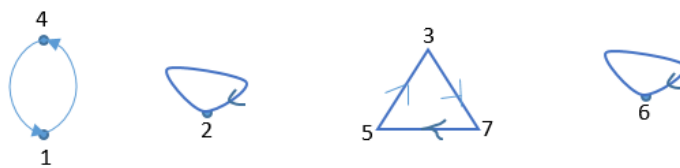
Figure 3.1: the digraph D_w of the permutation $w = (2)(41)(6)(753)$

Definition 33 *(Bracket Erasing Operation on Permutation) [21, Page 30]*
Consider a set permutation $w \in \mathfrak{S}_n$, written as its standard form of disjoint cycle notation; $w = C_1 \ldots C_k$, each cycle C_i (for $1 \leq i \leq k$), is marked off by a pair of parentheses. By bracket erasing operation, a permutation denoted by $\hat{w}$ can be obtained after erasing the parentheses of each cycle in the set permutation w. The bracket erasing operation is denoted by $\wedge$.

Proposition 34 *[21, Proposition 1.3.1] Let $\mathfrak{S}_n = \{w : [n] \to [n]\}$ be a set of permutations of the set $[n] = \{1, 2, \ldots, n\}$. Denote the bracket erasing operation as above by $\wedge$.*

(i) *The map $\mathfrak{S}_n \xrightarrow{\wedge} \mathfrak{S}_n$ is a bijection from $\mathfrak{S}_n$ to itself.*

(ii) *The number of left-to-right maximum in a permutation equals the number of cycles of the permutations.*

According to the discussion on page 30 in the book [21], we gave more clear and readable proof as follows.

Proof:

Suppose a permutation $w \in \mathfrak{S}_n$ has k cycles, let $w = C_1 C_2 \ldots C_k$, where C_i with $1 \leq i \leq k$ is a cycle of the permutation w.

1. From the permutation $w \in \mathfrak{S}_n$, with $w = C_1 C_2 \ldots C_k$, after erasing the parentheses of each cycle in the permutation, a unique permutation $\hat{w} \in \mathfrak{S}_n$ with no brackets can be obtained.

2. It remains to prove that from a permutation with no brackets $\hat{w} \in \mathfrak{S}_n$, a permutation $w \in \mathfrak{S}_n$ can be recovered uniquely.

   Let $\hat{w} = a_1 a_2 \ldots a_n \in \mathfrak{S}_n$,

   Firstly, going through $\hat{w} = a_1 a_2 \ldots a_n$ from left to right, insert a left parenthesis in the permutation $\hat{w}$ preceding every left-to-right maximum, which is an outstanding element a_i, for $1 \leq i \leq n$, such that $a_i > a_j$, for every $1 \leq j < i$.

It is noted that when $i = 1$, there is no a_j such that $1 \leq j < 1$, here, by convenience, the element $a_1 \in [n]$ is regarded as an outstanding element(a left-to-right maximum), that satisfies $a_i > a_j$, for every $1 \leq j < i$. Besides, Each left-to-right maxima is a largest number in each cycle of the permutation, the left parenthesis is inserted before the largest number, so, the number of left-to-right maximum in a permutation equals to the number of cycles of the permutations.

Secondly, Insert the right parentheses before every internal brackets already inserted in the permutation $\hat{w}$, and insert the right parentheses at the right end of the permutation.

Then, a unique permutation $w \in \mathfrak{S}_n$ written as the standard form of disjoint cycle notations is obtained.

Therefore, the defined map $\mathfrak{S}_n \overset{\wedge}{\longrightarrow} \mathfrak{S}_n$ is a bijection.

Remarks:

1. For instance, for a permutation w of $[7]$, with $w = (2)(41)(6)(753)$, after erasing the parentheses, it obtain $\hat{w} = 2416753$. Conversely, for a permutation $\hat{w} = 2416753$, the left-to-right maximum is $2, 4, 6, 7$ from left to right in the permutation $\hat{w}$. Insert a left parenthesis preceding every left-to-right maximum in $\hat{w}$; it obtain $(2(41(6(753$. Insert the right parentheses on appropriate positions, then, it obtain $w = (2)(41)(6)(753)$.

2. If represent the set permutation $\hat{w} = 2416753$ by a sequence of up and down steps, denote up step by $+1$ and down step by -1, then, a sequence is obtained from the standard form of the set permutation as $(+1, -1, +1, +1, -1, -1)$.

3.5.2   Statistics of Permutations

The presentation and definition 35 in this section is motivated from the discussion of cycle structures of permutations on page 29-30 in the book [21].

Let $\mathfrak{S}_n = \{w : [n] \rightarrow [n]\}$ be a set of permutations of the set $[n] = \{1, 2, \ldots, n\}$. A statistic on a set $\mathfrak{S}_n$ of permutations of $[n]$ is a function $f : \mathfrak{S}_n \rightarrow S$, where S is any set and is often taken to be natural numbers $\mathbb{N}$. For any permutation $w \in \mathfrak{S}_n$, $f(w)$ captures some interesting features of the permutation w.

Let S be a finite set, if $w \in \mathfrak{S}_S$, with $\sharp S = n$, then,
let $c_i = c_i(w) = $ the number of cycles of w of length i, the total number of cycles of w is denoted by $c(w)$, and $c(w) = \sum_{i=1}^{\infty} c_i(w)$.

Remarks:

1. Since the number of the finite set S equal to n, $\sharp S = n$, c_i is the number of cycles of the permutation $w \in \mathfrak{S}_n$ of length i, it is obvious that $n = $

$\sum_{i=1}^{n} i \cdot c_i$. For instance, if $S = [n] = \{1, 2, \ldots 7\}$, $w = (2)(41)(6)(753)$ is a permutation of $[7]$, it has $c_1 = 2$, $c_2 = 1$, $c_3 = 1$, $c_i = 0$, for $i \neq 1, 2, 3$, $1 \leq i \leq 7$.

2. Suppose a set permutation $w \in \mathfrak{S}_S$ has k cycles, it is noted that the notation for one cycle in the permutation is denoted by C_i, $1 \leq i \leq k$. However, the notation for the number of cycles of length i in the permutation is denoted by $c_i = c_i(w)$, which satisfies $k = c_1 + c_2 + \ldots + c_n = \sum_{i=1}^{\infty} c_i(w) = c(w)$. If $\sharp S = n$, then $n = \sum_{i=1}^{n} i \cdot c_i$.

Definition 35 *(Cycle structure statistics on permutation)*

Let S be a finite set, if $w \in \mathfrak{S}_S$, with $\sharp S = n$, then,
let $c_i = c_i(w) = $ the number of cycles of w of length i. The type of the permutation $w \in \mathfrak{S}_S$, denoted by $type(w)$, is defined to be the sequence $(c_1, \ldots, c_n)$. That is, $type(w) := (c_1, \ldots, c_n)$.

3.5.3   Count cycle structure statistics on permutations

Question: [21, Proposition 1.3.2] Count the number of permutations $w \in \mathfrak{S}_S$ of the type $(c_1, \ldots, c_n)$.

Answer:
Let $\mathfrak{S}_S = \{u : S \to S\}$ be a set of permutations of a finite set S, with $\sharp S = n$.
Let $\mathfrak{S}_S^{\mathcal{C}} = \{u \in \mathfrak{S}_S : type(w) = (c_1, \ldots, c_n)\}$ be the set of all permutations $u \in \mathfrak{S}_S$ of the type $\mathcal{C} = (c_1, \ldots, c_n)$.

Define a map $\Phi : \mathfrak{S}_S \to \mathfrak{S}_S^{\mathcal{C}}$, where $u \in \mathfrak{S}_S^{\mathcal{C}}$ is obtained from $w \in \mathfrak{S}_S$ by inserting parentheses in the way such that the disjoint cycle notation of the obtained permutation $u \in \mathfrak{S}_S^{\mathcal{C}}$ is written in increasing order of the lengths of the distinct cycles. That is, parenthesize the word $w \in \mathfrak{S}_S$ so that the first c_1 cycles have length 1, the second c_2 cycles have length 2, and so on.

Since $\sum_{i=1}^{n} i \cdot c_i = n$, starting from $w = w_1 w_2 \ldots w_n \in \mathfrak{S}_S$, after parenthesize the permutation word in the above manner, a unique $u \in \mathfrak{S}_S^{\mathcal{C}}$ will be obtained.

Conversely, from a given $u \in \mathfrak{S}_S^{\mathcal{C}}$, if u is written in disjoint cycle notations such that the cycle lengths are weakly increasing from left to right, that is, suppose $u \in \mathfrak{S}_S^{\mathcal{C}}$ has $\sum_{i=1}^{n} c_i = k$ cycles, it is written as $C_1 C_2 \ldots C_k$ with $\sharp C_1 \leq \sharp C_2 \leq \ldots \leq \sharp C_k$.

It is noted that a permutation $u \in \mathfrak{S}_S^{\mathcal{C}}$ may not be standard form of its disjoint cycle notations. For instance, given a type $\mathcal{C} = (c_1, c_2, \ldots, c_n) = (2, 1, 1, 0, \ldots, 0)$ we start from $u = (2)(6)(41)(753) \in \mathfrak{S}_S^{\mathcal{C}}$, but not from $u = (2)(41)(6)(753)$, which is the standard form of disjoint cycle notations, but it is not an element in $\mathfrak{S}_S^{\mathcal{C}}$.

Since $\mathfrak{S}_S^{\mathcal{C}}$ contains permutations which are written as a union of distinct cycles in weakly increasing order of the cycle lengths.

Then, given a permutation $u \in \mathfrak{S}_S^{\mathcal{C}}$,

Firstly, order the distinct cycles of length i will get $i!$ different permutations in $\mathfrak{S}_S$ with the same type $C = (c_1, c_2, \ldots, c_n)$.

Secondly, since the circular shift of each cycle represent the same cycle, circularly shifting one cycle of length i will get $\binom{i}{1} = i$ different permutations in $\mathfrak{S}_S$ with the same type $C = (c_1, c_2, \ldots, c_n)$.

Thus, circularly shifting the distinct c_i cycles of length i will get $\binom{i}{1}^{c_i}$ different permutations in $\mathfrak{S}_S$ with the same type $C = (c_1, c_2, \ldots, c_n)$.

Therefore, from a permutation $u \in \mathfrak{S}_S^{\mathcal{C}}$ with the type $C = (c_1, c_2, \ldots, c_n)$, there are

$$\binom{1}{1}^{c_1} c_1! \cdot \binom{2}{1}^{c_2} c_2! \ldots \binom{n}{1}^{c_n} \cdot c_n! = 1^{c_1} c_1! 2^{c_2} c_2! \ldots n^{c_n} c_n!$$

different permutations $w \in \mathfrak{S}_S$ that corresponds to $u \in \mathfrak{S}_S^{\mathcal{C}}$.

Thus, the total number of permutations in the set $\mathfrak{S}_S$ equals to

$$1^{c_1} c_1! 2^{c_2} c_2! \ldots n^{c_n} c_n! \cdot (\sharp \mathfrak{S}_S^{\mathcal{C}})$$

It is already known that the number of permutations $w \in \mathfrak{S}_S$ is $n!$, therefore,

$$\sharp \mathfrak{S}_S^{\mathcal{C}} = \frac{n!}{1^{c_1} c_1! 2^{c_2} c_2! \ldots n^{c_n} c_n!}$$

The application of the above counts is very useful when representing the counts using generating functions.

Fix n, using symbolic method, the exponential generating function for the count of permutations $w \in \mathfrak{S}_n$ with a given type $(c_1, \ldots, c_n)$ is defined as follows,

$$Z_n = \frac{1}{n} \sum_{w \in \mathfrak{S}_n} t^{type(w)}, \quad \text{where} \quad type(w) := (c_1, \ldots, c_n)$$

If using n variables in the polynomial, write

$$t^{type(w)} = t_1^{c_1} t_2^{c_2} \ldots t_n^{c_n},$$

$$Z_n = Z_n(t_1, \ldots, t_n) = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} t^{type(w)}$$

then, when using Z_n as an coefficient, the generating function $\sum_{n \geq 0} Z_n x^n$ will have a nice form, which gives the following theorem.

Theorem 36 *[21, Theorem 1.3.3]* $\sum_{n \geq 0} Z_n x^n = \exp(t_1 x + t_2 \frac{x^2}{2} + t_3 \frac{x^3}{3} + \ldots)$

Proof:

The right-hand side(RHS) is

$$\exp(\sum_{i \geq 1} t_i \frac{x^i}{i}) = \prod_{i=1}^{\infty} \exp(t_i \frac{x^i}{i}),$$

That is,

$$\text{RHS} = (\sum_{n=0}^{\infty} \frac{t_1^n x^n}{n!}) \cdot (\sum_{n=0}^{\infty} \frac{t_2^n (\frac{x^2}{2})^n}{n!})) \cdot (\sum_{n=0}^{\infty} \frac{t_3^n (\frac{x^3}{3})^n}{n!})) \ldots$$

Taking the coefficient of x^n in the RHS, that is, taking the coefficient of $x^{k_1 + 2k_2 + 3k_3 + \cdots} = x^n$, here, $(k_1, k_2, k_3, \ldots)$ is a sequence with each element greater than or equal to zero.

For fixed $n \geq 0$, the only way to get x^n is when $k_1 + 2k_2 + 3k_3 + \ldots + nk_n = n$, here, set $(k_1, k_2, \ldots, k_n) = (c_1, c_2, \ldots, c_n)$ and the coefficient of $t_1^{c_1} t_2^{c_2} \ldots t_n^{c_n} x^n$ is $\frac{1}{c_1! 2^{c_2} c_2! \ldots n^{c_n} c_n!}$, That is, the coefficient of $t_1^{c_1} t_2^{c_2} \ldots t_n^{c_n} x^n$ is

$$\frac{1}{c_1! 2^{c_2} c_2! \ldots n^{c_n} c_n!} = \frac{1}{n!} \frac{n}{c_1! 2^{c_2} c_2! \ldots n^{c_n} c_n!},$$

that is, the coefficient of the RHS equals to the coefficient of x^n in the generating function $\sum_{n \geq 0} Z_n x^n$.

An example of applying the above theorem is given as follows.

Question: [21, Example 1.3.5]
Find the expected number of k-cycles in a permutation $w \in \mathfrak{S}_n$, where k-cycles are cycles of length k in the permutation w.

Answer: Let $E_k(n)$ denotes the expected number of k-cycles in a permutation $w \in \mathfrak{S}_n$.

If the expectation is taken with respect to the uniform distribution on $\mathfrak{S}_n$, since $\sharp \mathfrak{S}_n = \sharp$ of permutations of $[n] = n!$, the uniform distribution on $\mathfrak{S}_n$ gives

$$\text{prob (one permutation } w \in \mathfrak{S}_n) = \frac{1}{n!}$$

Let $c_k(w) = $ the number of cycles of length k in the permutation $w \in \mathfrak{S}_n$.

so,

$$E_k(n) = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} c_k(w)$$

By definition,

$$Z_n = Z_n(t_1, \ldots, t_n) = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} t^{type(w)} = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} t_1^{c_1(w)} t_2^{c_2(w)} \ldots t_n^{c_n(w)}$$

from the above definition of Z_n, to get $\frac{1}{n!} c_k(w)$, for one permutation $w \in \mathfrak{S}_n$, differentiate one term corresponding to the permutation $w \in \mathfrak{S}_n$ in the polynomial Z_n with respect to t_k, and then, in that term, set $t_i = 1$, for $1 \le i \le n$.

Therefore, to get the $\frac{1}{n!} \sum_{w \in \mathfrak{S}_n} c_k(w)$, differentiate polynomial Z_n with respect to t_k, and then, set $t_i = 1$, for $1 \le i \le n$.

That is,

$$E_k(n) = \frac{\partial}{\partial t_k} Z_n(t_1, \ldots, t_n)|_{t_i=1, \text{ for } 1 \le i \le n}$$

Since it is already known that,

$$\sum_{n \ge 0} Z_n x^n = \exp\left(\sum_{i \ge 1} t_i \frac{x^i}{i}\right)$$

Therefore,

$$
\begin{aligned}
\sum_{n \ge 0} E_k(n) x^n &= \frac{\partial}{\partial t_k}\left(\sum_{n \ge 0} Z_n x^n\right)\Bigg|_{t_i=1, \text{ for } 1 \le i \le n} \\
&= \frac{\partial}{\partial t_k} \exp\left(\sum_{i \ge 1} t_i \frac{x^i}{i}\right)\Bigg|_{t_i=1, \text{ for } 1 \le i \le n} \\
&= \frac{x^k}{k} \exp\left(\sum_{i \ge 1} t_i \frac{x^i}{i}\right)\Bigg|_{t_i=1, \text{ for } 1 \le i \le n} \\
&= \frac{x^k}{k} \exp\left(\sum_{i \ge 1} \frac{x^i}{i}\right) \\
&= \frac{x^k}{k} \exp\left(\log(1-x)^{-1}\right) \\
&= \frac{x^k}{k} \frac{1}{1-x} \\
&= \frac{x^k}{k} \sum_{n \ge 0} x^n
\end{aligned}
$$

Note: the first equation follows because differentiation is linear.

To find $E_k(n)$, extract the coefficient of x^n on both sides of the equation in the above equation, then,

$$
\begin{aligned}
E_k(n) &= [x^n] \sum_{n \geq 0} \frac{1}{k} x^{n+k} \\
&= [x^n] \sum_{n-k \geq 0} \frac{1}{k} x^n \\
&= \frac{1}{k}, \text{ for } n \geq k
\end{aligned}
$$

Note: in the first equation, the limit notation n in the summation is not the same as the n in the extract notation $[n]$.

In the above example, the number of permutation $w \in \mathfrak{S}_n$ with a given cycle type $(c_1, c_2, \ldots, c_n)$, suppose the permutation has k cycles, it is noted that $\sum_{1 \leq i \leq n} = k$. For a fixed permutation cycle type $(c_1, c_2, \ldots, c_n)$, a unique number of cycles in the permutation can be obtained as $\sum_{1 \leq i \leq n}$, say, the number of cycles equals to $k = \sum_{1 \leq i \leq n}$.

Suppose the number of cycles in a permutation $w \in \mathfrak{S}_n$ is k,
In the next example, counting the number of permutation with exactly k cycles.

Question: [21, Lemma 1.3.6] Count the number of permutation $w \in \mathfrak{S}_n$ with exactly k cycles.

Answer: let $c(n, k) :=$ the number of permutations $w \in \mathfrak{S}_n$ with exactly k cycles.

Firstly, find the recurrence equation for $c(n, k)$;

Starting from a permutation $w \in \mathfrak{S}_{n-1}$ with m cycles, when constructing a permutation $\tilde{w} \in \mathfrak{S}_n$ by adding the number n into w, the number of cycles in the permutation $\tilde{w} \in \mathfrak{S}_n$ can only be m or $m+1$.

Therefore, to construct a permutation $\tilde{w} \in \mathfrak{S}_n$ with k cycles, there are two possibilities; one is from a permutation $w \in \mathfrak{S}_{n-1}$ with k cycles, another way is from a permutation $w \in \mathfrak{S}_{n-1}$ with $k-1$ cycles.

So, firstly, there are $c(n-1, k)$ ways to choose a permutation $w \in \mathfrak{S}_{n-1}$ with k cycles.

Secondly, then, there are $n-1$ ways to insert the number n after any of the numbers $1, 2, \ldots, n-1$ in the disjoint cycle decomposition of w. That is, in the new permutation $\tilde{w} \in \mathfrak{S}_n$, the number n appears in a cycle of length at least 2.

So, there are $(n-1)c(n-1, k)$ permutations $\tilde{w} \in \mathfrak{S}_n$ with k cycles, where there is no cycle of length 1 in the permutation $\tilde{w}$, that is, $\tilde{w}(i) \neq i$, for $i = 1, 2, \ldots, n$.

On the other hand,

1st, there are $c(n-1, k-1)$ ways to choose a permutation z $w \in \mathfrak{S}_{n-1}$ with $k-1$ cycles.

2nd, then, since changing the order of distinct cycles produces the same permutation, there are one way to insert the number n as a new cycle into the permutation $w \in \mathfrak{S}_{n-1}$ with $k-1$ cycles such that the new permutation $\tilde{w} \in \mathfrak{S}_n$ has k cycles.

So, there are $c(n-1, k-1) \cdot 1$ permutations $\hat{w} \in \mathfrak{S}_n$ with k cycles, where the minimum length of cycles in the permutation $\tilde{w}$ is one.

Then, from a permutation $w \in \mathfrak{S}_{n-1}$ with $k-1$ cycles, the new permutation $\tilde{w} \in \mathfrak{S}_n$ is obtained as follows,

$$\begin{cases} \tilde{w}(i) = n, & \text{if } i = n. \\ \tilde{w}(i) = w(i), & \text{if } i = 1, 2, \ldots, n-1 \in [n-1]. \end{cases}$$

Since the two possibilities are disjoint ways to construct $\tilde{w} \in \mathfrak{S}_n$ with k cycles,

$$c(n, k) = (n-1)c(n-1, k) + c(n-1, k-1), \quad \text{for } n, k \geq 1$$

Besides,

1. if $n = 0$, $\sharp$ of permutations $\tilde{w} \in \mathfrak{S}_n$ has k cycles with $k = 0$, set $c(0, 0) = 1$

2. if $n \geq 1$, $\sharp$ of permutations $\tilde{w} \in \mathfrak{S}_n$ must have at least $k = 1$ cycle, so, when $k = 0, n \geq 1$, set $c(n, k) = 0$.

3. since $\sharp$ of permutations $\tilde{w} \in \mathfrak{S}_n$ have at most n cycles, if $n < k$, set $c(n, k) = 0$.

The generating function version for the counts of permutations $w \in \mathfrak{S}_n$ with exactly k cycles can be obtained using symbolic method. It is given as a proposition as follows,

Proposition 37 *[21, Proposition 1.3.7] Let t be an indeterminate variable and fix $n \geq 0$, let $c(n, k)$ be the number of permutations $w \in \mathfrak{S}_n$ with exactly k cycles. Then, the generating function of $c(n, k)$ for fixed $n \geq 0$ is*

$$\sum_{k=0}^{n} c(n, k) t^k = t(t+1)(t+2) \ldots (t+n-1)$$

The following proof idea is mainly follow the second proof of the proposition 1.3.7 in the book [21], but we added the derivation of the counts $c(n, k)$ using Flajolet symbolic method.

Proof: Let $c_i(w) = c_i =$ the number of cycles of length i in a permutation $w \in \mathfrak{S}_n$, for $1 \leq i \leq n$.

for fixed $n \geq 0$,

$$
\begin{aligned}
\sum_{k=0}^{n} c(n,k) t^k &= \sum_{k=0}^{n} c(n,k) t^{\sum_{i=1}^{n} c_i(w)} \\
&= \sum_{w \in \mathfrak{S}_n} t^{\sum_{i=1}^{n} c_i(w)} \\
&= \sum_{w \in \mathfrak{S}_n} t^{c_1} t^{c_2} \dots t^{c_n} \\
&= \sum_{w \in \mathfrak{S}_n} t_1^{c_1} t_2^{c_2} \dots t_n^{c_n} \quad (\text{with } t_i = t, \text{ for } i = 1, 2, \dots n)
\end{aligned}
$$

It is noted that the generating function of counts of permutation $w \in \mathfrak{S}_n$ with a given cycle type $type(w) = (c_1, c_2, \dots, c_n)$ is known as

$$
Z_n = Z_n(t_1, \dots, t_n) = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} t^{type(w)} = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} t_1^{c_1(w)} t_2^{c_2(w)} \dots t_n^{c_n(w)}
$$

Thus,

$$
\sum_{k=0}^{n} c(n,k) t^k = n! \cdot Z_n(t, t, t, \dots)
$$

It is also known that the generating function of the sequence $\{Z_n\}_{n=0}^{\infty}$ is

$$
\sum_{n \geq 0} Z_n x^n = \exp(t_1 x + t_2 \frac{x^2}{2} + t_3 \frac{x^3}{3} + \dots)
$$

then, taking the generating function of $\dfrac{1}{n!} \displaystyle\sum_{k=0}^{n} c(n,k) t^k$, for $n \geq 0$ gives

$$
\begin{aligned}
\sum_{n \geq 0} \left(\sum_{k=0}^{n} c(n,k) t^k \right) \frac{x^n}{n!} &= \exp(t(x + \frac{x^2}{2} + \frac{x^3}{3} + \dots)) \\
&= \exp(t(\log(1-x)^{-1})) \\
&= (1-x)^{-t} \\
&= \sum_{n \geq 0} (-1)^n \binom{-t}{n} x^n \\
&= \sum_{n \geq 0} t(t+1) \dots (t+n-1) \frac{x^n}{n!}
\end{aligned}
$$

Therefore,

$$
\sum_{k=0}^{n} c(n,k) t^k = t(t+1) \dots (t+n-1)
$$

3.5.4   Count Inversion Structure Statistics on Permutations

A permutation $w \in \mathfrak{S}_n$ can be associated with an integer sequence $(a_1, a_2, \ldots, a_n)$, with $0 \leq a_i \leq n - i$ by the following method.

The method is motivated from the discussion of associating a permutation with a given integer sequence $(a_1, a_2, \ldots, a_n)$, with $0 \leq a_i \leq n - i$ on page 35-36 in the book [21], we gave a detailed and applicable algorithm of the natural correspondence from the given integer sequence to a permutation and also give the argument to justify the algorithm is reasonable by analysing the relation between the possible values of a_{n-i} and the number of positions in the permutation u_i. The method is stated as follows.

Given an integer sequence $(a_1, \ldots, a_n)$ with $0 \leq a_i \leq n-i$, a permutation $w \in \mathfrak{S}_n$ is built up as follows,

starting from $w =$ empty word,

1. Insert the number n into w, write the new permutation word as $u_1 = w_1 = n$.

2. Insert the number $n - 1$ into u_1 with a position after a_{n-1} elements in the old permutation word u_1 which is built up in the last step.

   For instance, if $a_{n-1} = 0$, then, the new permutation is $u_2 = w_1 w_2 = n-1, n$, thus, the number $n - 1$ is inserted into a position after 0 element in the last step permutation $u_1 = n$, so, the number $n - 1$ is inserted before the number n in the last step permutation.

   If $a_{n-1} = 1$, then, the new permutation is $u_2 = w_1 w_2 = n, n - 1$, that is, insert $n - 1$ into a position after 1 elements in the last step permutation $u_1 = n$, so, the number $n - 1$ is inserted after the number n in the last step permutation.

3. Proceed to insert the numbers $n - i$, for $i = 2, 3, \ldots, n - 1$ into a position after a_{n-i} elements in their last step permutations $u_i = w_1 w_2 \ldots w_i$, which is a permutation of the original inserted numbers $n, n - 1, \ldots, n - i + 1$.

4. The procedure ends up until $i = n - 1$, that is, insert the number 1 in a position after $a_{n-(n-1)} = a_1$ elements in the sequence $u_{n-1} = w_1 w_2 \ldots w_{n-1}$, which is the permutation of the original inserted numbers $n, n - 1, \ldots, 2$ and obtain $u_n = w_1 w_2 \ldots w_n$, which is a permutation $w \in \mathfrak{S}_n$.

   Remarks: It can be proved that it is always valid to insert a number in a position which is after a_{n-i} elements from left to right in the originally inserted numbers $u_i = w_1 w_2 \ldots w_i$.

   Since in the permutation $u_i = w_1 w_2 \ldots w_i$, there are already i numbers, and then $i + 1$ available position except the ending position. It is consistent with the given condition that a_{n-i} is an integer satisfying $0 \leq a_{n-i} \leq i$.

Conversely, given a permutation $w \in \mathfrak{S}_n$, an integer sequence $(a_1, a_2, \ldots, a_n)$ with $0 \le a_i \le n - i$ can also be uniquely built up.

Before doing the converse building procedure, the definition of inversions is given as follows.

Definition 38 *(Definition of Inversions) [21, Page 36]*

Let $\mathfrak{S}_n$ be a set of permutations of the set $[n] = \{1, 2, \ldots, n\}$.
Consider a set permutation $w = w_1 w_2 \ldots w_n \in \mathfrak{S}_n$, a pair (w_i, w_j) in the permutation is called an inversion of the permutation if and only if the pair (w_i, w_j) satisfies both $i < j$ and $w_i > w_j$.

Then, a relation between any number i with $1 \le i \le n$ in a permutation $w \in \mathfrak{S}_n$ and any integer a_i, for $1 \le i \le n$ in the integer sequence $(a_1, a_2, \ldots, a_n)$ with $0 \le a_i \le n - i$ can be given as proposition as follows.

Following the definitions 38 of inversions, motivated from the discussion regarding the integer sequence $(a_1, a_2, \ldots, a_n)$, with $0 \le a_i \le n - i$ on the page [21, page 34-36], we derive a proposition that claims each index i, with $0 \le i \le n$ in a permutaiton $w \in \mathfrak{S}_n$ can be uniquely characterized by an integer a_i in the integer sequence $(a_1, a_2, \ldots, a_n)$, with $0 \le a_i \le n - i$.

Proposition 39 *Given a set permutation $w \in \mathfrak{S}_n$, and any number $1 \le i \le n$. Fix the number i, the number of inversions for the fixed number i in the permutation w is defined as*

$$a_i := \sharp \left\{ (k, i), 1 \le k \le n :\ (k, i)\ \text{is an inversion of the given set permutation } w \right\}$$

then, the number i, with $1 \le i \le n$ in the given permutation $w \in \mathfrak{S}_n$ can be uniquely characterized by an integer a_i in the integer sequence $(a_1, a_2, \ldots, a_n)$ with $0 \le a_i \le n - i$.

Proof: Recalling the procedure of constructing a set permutation $w \in \mathfrak{S}_n$ from a given integer sequence $(a_1, a_2, \ldots, a_n)$, it is noted that a number i is inserted after inserting a number $i + 1$, for any number $1 \le i \le n - 1$. That is, numbers in the set $\{1, 2, \ldots, n\}$ are inserted in decreasing order.

When inserting any number i, for $1 \le i \le n - 1$ into the last step permutation of numbers $n, n - 1, \ldots, i + 1$, the number i is less than all numbers already inserted in the last step permutation. If the number i is inserted after a_i, with $0 \le a_i \le n - i$ elements from left to right in the last step permutation, then, the number of inversion in the new permutation of numbers $n, n - 1, \ldots, i + 1, i$ increased by a_i inversions, which is the number of inversions for the fixed number i in the new permutation.

Besides, in the new permutation of numbers $n, n - 1, \ldots, i + 1, i$, the number of inversions a_j, for any fixed number j, with $i + 1 \le j \le n$ will not be changed. It can be proved as follows.

Inserting any number i, $1 \le i \le n-1$ into the last step permutation of numbers $n, n-1, \ldots, i+1$ results in two possible position relation between the number i and any number j, with $i+1 \le j \le n$;

1. when the new number i is inserted before j, with $i+1 \le j \le n$, the position of the number i is before the position of the number j. Since $i < j$, from the definition of a_j, the number of inversions for any fixed number j, with $i+1 \le j \le n$ will not change and it still be a_j.

2. when the new number i is inserted after j, with $i+1 \le j \le n$, the position of the number i is after the position of the number j. From the definition of a_j in the new permutation of the numbers $n, n-1, \ldots, i+1, i$,

$$a_j = \sharp \{(k,j), i \le k \le n : \text{ the position of } k \text{ is before } j \text{ and } k > j\}$$

the number of inversions for fixed number j, with $i+1 \le j \le n$, will not change, since there is no new element inserted before the number j in the last step permutation of numbers $n, n-1, \ldots, i+1$.

Therefore, if the number i is inserted into the last step permutation of numbers $n, n-1, \ldots, i+1$, the number of inversion in the new permutation of numbers $n, n-1, \ldots, i+1, i$ increased by a_i inversions, and in the new permutation, the number of inversions a_j, for any fixed number j, with $i+1 \le j \le n$ will not be changed.

Thus, an integer a_i in the integer sequence $(a_1, a_2, \ldots, a_n)$ with $0 \le a_i \le n-i$ can uniquely characterize a number i, with $1 \le i \le n$ in a permutation $w \in \mathfrak{S}_n$.

Since the above proposition, given a permutation $w \in \mathfrak{S}_n$,
an integer sequence $(a_1, a_2, \ldots, a_n)$ with $0 \le a_i \le n-i$ can be built up as follows,

Let $a_i = a_i(w) := $ the number of inversions for fixed number i in the given permutation $w \in \mathfrak{S}_n$,

Since a_i uniquely characterize the number i in the permutation w, go through i from the number $i = 1$ to $i = n$ in the permutation $w = w_1 w_2 \ldots w_n \in \mathfrak{S}_n$,

find the number a_i by counting:

$$a_i = \sharp \{(k,i), 1 \le k \le n : k \text{ is to the left of } i \text{ in the permutation } w \text{ and } k > i\}$$

for any fixed number $i \in [n] = \{1, 2, \ldots, n\}$

then, an integer sequence $(a_1, a_2, \ldots, a_n)$ with $0 \le a_i \le n-i$ is obtained from a given permutation $w \in \mathfrak{S}_n$.

Definition 40 *(Inversion table) [21, Page 36]*

Let $\mathfrak{S}_n$ be a set of permutations of the set $[n] = \{1, 2, \ldots, n\}$.
Consider a set permutation $w = w_1 w_2 \ldots w_n \in \mathfrak{S}_n$, an integer sequence $(a_1, a_2, \ldots, a_n)$
with $0 \le a_i \le n - i$ can be obtained from the permutation $w \in \mathfrak{S}_n$. Then, the
integer sequence is called the inversion table of the permutation $w \in \mathfrak{S}_n$, denoted
by $I(w) = (a_1, a_2, \ldots, a_n)$.

Remarks:
In the procedure from a set permutation $w \in \mathfrak{S}_n$ to an integer sequence $(a_1, a_2, \ldots, a_n)$,
with $1 \le a_i \le n - i$, it is noticed that in the permutation $u_{n-i} = w_1 w_2 \ldots w_{n-i}$ of
the numbers $i+1, i+2, \ldots, n$, when inserting a number i into the position after a_i
elements in the permutation u_{n-i}, the number of inversions in the new permutation
u_{n-i+1} are increased by a_i compared to the last step permutation u_{n-i}, where
$u_{n-i+1} = w_1 w_2 \ldots w_{n-i+1}$ is a permutation of the numbers $i, i + 1, \ldots, n$.

Proposition 41 *(a bijection) [21, Proposition 1.3.13]*

Fix an integer number n,
Let $\mathfrak{S}_n = \{\, w : w$ is a permutation of $[n] = \{1, 2, \ldots, n\}\}$,
Let $\mathcal{T}_n = \{\, (a_1, \ldots, a_n) : 0 \le a_i \le n - i, \ a_i \in \mathbb{Z}^+\} = \{\, I(w) : w \in \mathfrak{S}_n \,\}$.

Then, the map $I : \mathfrak{S}_n \to \mathcal{T}_n$ that sends each permutation $w \in \mathfrak{S}_n$ to its inversion
table $I(w) \in \mathcal{T}_n$ is a bijection.

Proof: The proof can be done from the above discussion.

It is known that an integer a_i in the integer sequence $(a_1, a_2, \ldots, a_n)$ with $0 \le a_i \le n - i$ can uniquely characterize a number i, with $1 \le i \le n$ in a permutation $w \in \mathfrak{S}_n$,

therefore, the inversion table $I(w) = (a_1, a_2, \ldots, a_n)$ can uniquely determine a permutation $w \in \mathfrak{S}_n$.

Corollary 42 *[21, Corollary 1.3.13]*
(GF for the counts of permutations by inversions)

Fix an integer number n,

Let $\mathfrak{S}_n = \{\, w : w$ is a permutation of $[n] = \{1, 2, \ldots, n\}\}$,

Let $inv(w) :=$ the number of inversions of the permutation $w \in \mathfrak{S}_n$,

Then, the generating function for the counts of permutations by the number of
inversions is

$$\sum_{w \in \mathfrak{S}_n} q^{inv(w)} = (1 + q)(1 + q + q^2) \ldots (1 + q + q^2 + \ldots + q^{n-1})$$

In the following proof, we added the derivation of the derivation of $\displaystyle\sum_{w \in \mathfrak{S}_n} q^{inv(w)} =$

$$\sum_{(a_1, a_2, \ldots, a_n) \in \mathcal{T}_n} q^{a_1 + a_2 + \ldots + a_n}$$ using Flajolet symbolic method, say Cartesian product symbolic method, which makes the proof more readable.

Proof: Let $\mathcal{T}_n = \{\ (a_1, \ldots, a_n) : 0 \leq a_i \leq n - i, \ \ a_i \in \mathbb{Z}^+\} = \{\ I(w) : w \in \mathfrak{S}_n\ \}$,

Let $I(w) = (a_1, a_2, \ldots, a_n) \in \mathcal{T}_n$, it is noted that $inv(w) = a_1 + a_2 + \ldots + a_n$.

Since $w \in \mathfrak{S}_n \mapsto I(w) \in \mathcal{T}_n$ is a bijection,

$$
\begin{aligned}
\sum_{w \in \mathfrak{S}_n} q^{inv(w)} &= \sum_{I(w) \in \mathcal{T}_n} q^{inv(w)} \\
&= \sum_{(a_1, a_2, \ldots, a_n) \in \mathcal{T}_n} q^{a_1 + a_2 + \ldots + a_n}
\end{aligned}
$$

Since $\mathcal{T}_n = \{\ (a_1, \ldots, a_n) : 0 \leq a_i \leq n - i, \ \ a_i \in \mathbb{Z}^+\}$, it follows that,

$$
\begin{aligned}
\sum_{w \in \mathfrak{S}_n} q^{inv(w)} &= \sum_{a_1 = 0}^{n-1} \sum_{a_2 = 0}^{n-2} \cdots \sum_{a_n = 0}^{0} q^{a_1} q^{a_2} \ldots q^{a_n} \\
&= (\sum_{a_1 = 0}^{n-1} q^{a_1})(\sum_{a_2 = 0}^{n-2} q^{a_2}) \ldots (\sum_{a_n = 0}^{0} q^{a_n}) \\
&= (1 + q + q^2 + \ldots + q^{n-1}) \ldots (1 + q + q^2)(1 + q)
\end{aligned}
$$

Definition 43 *(The q-analogue of n!) [21, Page 37]*

In general, q-analogue of a mathematical object is an object depending on the variable q that reduces to the original object when setting $q = 1$. If the reduced mathematical object is $n!$, then the polynomial $(1+q)(1+q+q^2) \ldots (1+q+q^2+\ldots+q^{n-1})$ is called "the q-analogue of n!", denoted by $(n)!$. Each term in the polynomial is called a q-analogue of a corresponding number; the polynomial $(1+q+q^2+\ldots+q^{n-1}) = \frac{1-q^n}{1-q}$ is called "the q-analogue of n".

3.5.5   Count Descents Statistics on Permutations

Recall the definition of a statistic on a set of permutation, a statistic on a set $\mathfrak{S}_n$ of permutations of $[n]$ is a function $f : \mathfrak{S}_n \mapsto S$, where S is any set (often taken to be an non-negative integer $\mathbb{N}$ or positive integer $\mathbb{Z}^+$).

Two statistics $type(w)$ and $inv(w)$ on permutations $w \in \mathfrak{S}_n$ has already discussed in the above two sections. In this section, another fundamental statistic associated with permutations will be discussed. That is, define $des(w) =$ the number of descents of a permutation $w \in \mathfrak{S}_n$, then, the statistic is $des : \mathfrak{S}_n \mapsto S$, with $S = des(w) : w \in s_n$. Definitions of descents, descent set are given as follows.

Definition 44 *(Descents) [21, Page 38]*

Let $\mathfrak{S}_n$ be a set of permutations of the set $[n] = \{1, 2, \ldots, n\}$.
Consider a set permutation $w = w_1 w_2 \ldots w_n \in \mathfrak{S}_n$,

then, the index $1 \leq i \leq n-1$ is called a descent of the permutation w if $w_i > w_{i+1}$.

Definition 45 *(Descent set) [21, Page 38]*

Let $\mathfrak{S}_n$ be a set of permutations of the set $[n] = \{1, 2, \ldots, n\}$.
Consider a set permutation $w = w_1 w_2 \ldots w_n \in \mathfrak{S}_n$,

then, the descent set $D(w)$ of the permutation w is defined by $D(w) = \{i : w_i > w_{i+1}, \text{ for } 1 \leq i \leq n - 1\}$. It is obvious that maximum value of descents of a permutation is $n - 1$, therefore, $D(w) \subseteq [n - 1] = \{1, 2, \ldots, n - 1\}$.

Remarks: Given a permutation $w = w_1 w_2 \ldots w_n \in \mathfrak{S}_n$, its descent set $D(w)$ can be found by going through letters of the permutation word from w_1 to w_n, and then finding i such that $w_i > w_{i+1}$, for $\leq i \leq n - 1$.

Based on the descent set of a permutation of distinct elements, a permutation statistic is defined as follows.

Definition 46 *[21, Page 39]*
(A statistic based on descent set)

Let $\mathfrak{S}_n$ be a set of permutations of the set $[n] = \{1, 2, \ldots, n\}$.
Consider a set permutation $w = w_1 w_2 \ldots w_n \in \mathfrak{S}_n$,
Let $D(w)$ be the descent set of the permutation $w \in \mathfrak{S}_n$,

then, the number of descents of the permutation w denoted by $des(w)$ is defined to be $des(w) = \sharp \{i : w_i > w_{i+1}, \text{ for } 1 \leq i \leq n - 1\} = \sharp \, D(w)$

Remarks:

1. What should be distinguished is definitions of permutation statistics $des(w)$ and $inv(w)$. The statistic $inv(w)$ is

$$
\begin{aligned}
inv(w) \;=\; & \sharp \{(w_i, w_j) : \; i < j \text{ and } w_i > w_j, \text{ for } 1 \leq i, j \leq n\} \\
=\; & \text{ the sum of elements in the inversion table } I(w)
\end{aligned}
$$

   The statistic $des(w)$ is

$$
\begin{aligned}
des(w) \;=\; & \sharp \{i : w_i > w_{i+1}, \text{ for } 1 \leq i \leq n - 1\} \\
=\; & \sharp \, D(w)
\end{aligned}
$$

2. It is noted that one permutation $w \in \mathfrak{S}_n$ corresponds to one inversion table $I(w) \in \mathcal{T}_n$, where $\mathcal{T}_n = \{ (a_1, \ldots, a_n) : 0 \leq a_i \leq n - i, \quad a_i \in \mathbb{Z}^+\} = \{ I(w) : w \in \mathfrak{S}_n \}$. However, the descent set $D(w)$ of a permutation and the permutation $w \in \mathfrak{S}_n$ do not have one-to-one correspondence; a decent set $D(w)$ might corresponds to many permutations. For instance, given a descent set $D(w) = \{1, 3\}$, when corresponding permutations $w \in \mathfrak{S}_6$, the descent set can corresponds to a permutation $w = 214356$ or a permutation 324156

Suppose a finite set $S \subseteq [n-1]$ is given, another two statistics based on descent sets are introduced as follows.

Definition 47 *[21, Page 38]*
(Two statistics based on descent set)

Let $\mathfrak{S}_n$ be a set of permutations of the set $[n] = \{1, 2, \ldots, n\}$,
For any permutation $w \in \mathfrak{S}_n$,
Let $S = \{s_1, \ldots, s_k\}_< \subseteq [n-1]$ be a finite set written in increasing order,

Then, define $\alpha(S) :=$ the number of permutations $w \in \mathfrak{S}_n$ whose descent set is contained in the set $S := \sharp\{w \in \mathfrak{S}_n : D(w) \subseteq S\}$.

Define $\beta(S) :=$ the number of permutations $w \in \mathfrak{S}_n$ whose descent set is equal to the set $S := \sharp\{w \in \mathfrak{S}_n : D(w) = S\}$.

Remarks: Since all subsets of the set S is descent sets in the definition of $\alpha(S)$, by definition of βS, a relation between $\alpha(S)$ and $\beta(S)$ is

$$\alpha(S) = \sum_{T \subseteq S} \beta(T) = \sum_{T \in 2^S} \beta(T)$$

Proposition 48 *[21, Proposition 1.4.1]*
(Find $\alpha(S)$ with given set S)

Let $S = \{s_1, \ldots, s_k\}$ with $s_1 < s_2 \ldots < s_k$, denoted by $\{s_1, \ldots, s_k\}_< \subseteq [n-1]$

For any permutation $w \in \mathfrak{S}_n$,

Let $D(w)$ be the descent set of the permutation $w \in \mathfrak{S}_n$,

Let $\alpha(S) := \sharp\{w \in \mathfrak{S}_n : D(w) \subseteq S\}$

Then,

$$\alpha(S) = \binom{n}{s_1, s_2 - s_1, s_3 - s_2, \ldots, n - s_k}$$

In the following proof of the proposition, we added more readable argument of the combinatorial method.

Proof: From the definition of $\alpha(S)$,

for any permutation $w \in \mathfrak{S}_n$,

to obtain a permutation $w = w_1 w_2 \ldots w_n \in \mathfrak{S}_n$ with its descent set $D(w) \subseteq S$, construct $D(w)$ from the elements of the given set $S = \{s_1, s_2, \ldots, s_k\}$.

Since the elements $s_1 < s_2 < \ldots < s_k \in S$ are indexes of letters in the permutation word, for any s_i, with $1 \le i \le k$, it may be a descent of the permutation $w \in \mathfrak{S}_n$, or it may not be a descent of the permutation $w \in \mathfrak{S}_n$.

However, the remaining indexes of letters in the permutation word w must not be descents of the permutation w, that is,

$$
\begin{aligned}
w_1 \quad &< w_2 < w_3 < \ldots < w_{s_1}, \\
w_{s_1+1} \quad &< w_{s_1+2} < w_{s_1+3} < \ldots < w_{s_2}, \\
&\vdots \\
w_{s_{k-1}+1} \quad &< w_{s_{k-1}+2} < w_{s_{k-1}+3} < \ldots < w_{s_k}, \\
w_{s_k+1} \quad &< w_{s_k+2} < w_{s_k+3} < \ldots < w_n
\end{aligned}
$$

Therefore, choose the first s_1 elements from $[n] = \{1, 2, \ldots, n\}$ for the letters $w_1, w_2, w_3 \ldots, w_{s_1}$ in the permutation word w. There are $\binom{n}{s_1}$ ways.

Then, choose the second $s_2 - s_1$ elements from the remaining $n - s_1$ elements for the letters $w_{s_1}, w_{s_1+1}, w_{s_1+2}, \ldots, w_{s_2}$. There are $\binom{n-s_1}{s_2-s_1}$ ways.

Proceed the choosing procedure until the final two sequence of permutation letters. That is, choose $s_k - s_{k-1}$ elements from the remaining $n - s_{k-1}$ elements for the letters $w_{s_{k-1}+1}, w_{s_{k-1}+2}, w_{s_{k-1}+3}, \ldots, w_{s_k}$; there are $\binom{n-s_{k-1}}{s_k-s_{k-1}}$ ways. Lastly, it remains to choose $n - s_k$ elements from the remaining $n - s_k$ elements; there is only one way.

Thus, given a set $S = \{s_1, s_2 \ldots, s_k\}_<$, the number of permutations $w \in \mathfrak{S}_n$ with its descent set $D(w) \subseteq S$ is equal to

$$
\begin{aligned}
\alpha(S) \;=\; & \binom{n}{s_1}\binom{n-s_1}{s_2-s_1}\binom{n-s_2}{s_3-s_2} \cdots \binom{n-s_{k-1}}{s_k-s_{k-1}}\binom{n-s_k}{n-s_k} \\
=\; & \binom{n}{s_1}\binom{n-s_1}{s_2-s_1}\binom{n-s_2}{s_3-s_2} \cdots \binom{n-s_{k-1}}{s_k-s_{k-1}} \\
=\; & \binom{n}{s_1, s_2 - s_1, s_3 - s_2, \ldots, n - s_k}
\end{aligned}
$$

Example 1: [21, Example 1.4.2]
Let a finite set $S = \{3, 8\}$, let a fixed number $n \geq 9$.

From the definition of $\alpha_n(S) = \sharp\{w \in \mathfrak{S}_n : D(w) \subseteq S\}$,

then,

$$
\begin{aligned}
\beta_n(3, 8) \;=\; & \sharp\{w \in \mathfrak{S}_n : D(w) = \{3, 8\}\} \\
=\; & \alpha_n(3, 8) - \alpha_n(3) - \alpha_n(8) + \alpha_n(\emptyset)
\end{aligned}
$$

Here,

$$
\begin{aligned}
\alpha_n(\emptyset) \;=\; & \text{the number of permutations } w \in \mathfrak{S}_n \text{ with } D(w) \subseteq \emptyset \\
=\; & \sharp \, of \, permutations \; w \in \mathfrak{S}_n \text{ with } D(w) = \emptyset
\end{aligned}
$$

Since the number of descents is $des(w) = \sharp D(w) = 0$, and only one permutation $w = w_1 w_2 \ldots w_n$, with $w_1 < w_2 < \ldots < w_n$ in $\mathfrak{S}_n$ has 0 descents, it follows that $\alpha_n(\emptyset) = 1$.

Besides, from the above proposition regarding $\alpha_n(S)$, with a given finite set S,

$$
\begin{aligned}
\alpha_n(3,8) &= \sharp\{\, w \in \mathfrak{S}_n \ : \ D(w) \subseteq \{3\} \,\} \\
&= \binom{n}{3,\, 8-3,\, n-8} = \binom{n}{3}\binom{n-3}{5}\binom{n-8}{n-8} \\
&= \binom{n}{3}\binom{n-3}{5}
\end{aligned}
$$

$$
\begin{aligned}
\alpha_n(3) &= \sharp\{\, w \in \mathfrak{S}_n \ : \ D(w) \subseteq \{3\} \,\} \\
&= \binom{n}{3,\, n-3} = \binom{n}{3}
\end{aligned}
$$

Similarly,

$$
\alpha_n(8) = \sharp\{\, w \in \mathfrak{S}_n \ : \ D(w) \subseteq \{8\} \,\} = \binom{n}{8}
$$

Thus,

$$
\begin{aligned}
\beta_n(3,8) &= \alpha_n(3,8) - \alpha_n(3) - \alpha_n(8) + \alpha_n(\emptyset) \\
&= \binom{n}{3,\, 5,\, n-8} - \binom{n}{3} - \binom{n}{8} + 1
\end{aligned}
$$

Remarks: Recall that the number of elements in a permutation descent set $D(w) = \sharp\, D(w) = des(w)$ ranges from 0 to $n-1$.

Example2: [21, Page 39]

If a permutation $w \in \mathfrak{S}_n$ has the descent set $D(w) = \{1,3,5,\ldots\} \cap [n-1]$,

then, the permutation is called an alternating permutation, or called down-up, since the permutation is $w = w_1 w_2 \ldots w_n$, with $w_1 > w_2 < w_3 > w_4 < \ldots$

For a given descent set $D(w) = \{1,3,5,\ldots\} \cap [n-1]$, there are many alternating permutations which correspond to the descent set.

Remarks: Similarly, if a permutation $w \in \mathfrak{S}_n$ has the descent set $D(w) = \{2,4,6,\ldots\} \cap [n-1]$, then, the permutation is reverse alternating permutation, or up-down, since the permutation is $w = w_1 w_2 \ldots w_n$, with $w_1 > w_2 < w_3 > w_4 < \ldots$

Next, using the same notation and definition of $A(d,k)$ taken from [21, Page 39], motivated from the example 1.4.2 on [21, Page 38] and the first few examples of Eulerian polynomial stated [21, Page 39], we derive a formula $A(n,k)$ for counting the number of permutations $w \in \mathfrak{S}_n$ with a fixed number of descents $k-1$.

Theorem 49 *The number of permutations* $w \in \mathfrak{S}_n$ *with exactly* $k - 1$ *descents is calculated using the formula*

$$
\begin{aligned}
A(n, k) &= \sharp\{\, w \in \mathfrak{S}_4 \; : \; w \text{ has } k - 1 \text{ descents} \,\} \\
&= \underbrace{\sum_{1 \le i_1 < i_2 < \ldots < i_{k-1} \le n-1}}_{} \beta_n(i_1, i_2, \ldots, i_{k-1}), \\
&\qquad\qquad where \; \beta_n(i_1, i_2, \ldots, i_{k-1}) = \sharp\{ w \in \mathfrak{S}_n : D(w) = \{i_1, i_2, \ldots, i_{k-1}\}\}
\end{aligned}
$$

Question:

Fix an integer number n,

Let $\mathfrak{S}_n$ be a set of permutations of the set $[n] = \{1, 2, \ldots, n\}$.

Find the formula for the number of permutations $w \in \mathfrak{S}_n$ with exactly $k - 1$ descents.

Remarks: It is noted that for any permutation $w \in \mathfrak{S}_n$, the maximum number of its descents is $n - 1$, the minimum number of its descents is 0, so, consider the number of descents is denoted by $k - 1$, thus, when write the counts in generating function, the powers of the polynomial can be from 1 to k.

Solution:

Define $A(n, k) := \sharp$ of permutations $w \in \mathfrak{S}_n$ with exactly $k - 1$ descents.

Fix an integer number n, define the generating function for the sequence $\{A(n, k)\}_{k=1}^n$ as

$$
\begin{aligned}
A_n(x) &= \sum_{w \in \mathfrak{S}_n} x^{1 + des(w)} \\
&= \sum_{k=1}^n A(n, k) x^k
\end{aligned}
$$

By the definition of $A(n, k)$,

$A(0, k) = \sharp$ of permutations $w \in \mathfrak{S}_n$ with exactly -1 descents.

Set $A(0, k) = 1$, thus, $A_0(x) = 1$.

Next, to derive the formula of $A(n, k)$, start from an example $A(4, k)$, find the coefficients of $A_4(x)$, then, generalize to obtain the formula of $A(n, k)$.

Then, find the expansion of $A_4(x) = \sum_{k=1}^4 A(4, k) x^k$,

where $A(4, k) = \sharp$ of permutations $w = w_1 w_2 w_3 w_4 \in \mathfrak{S}_4$ with $k - 1$ descents, for $1 \le k \le 4$.

1st, find $A(4, 4)$, which is $\sharp$ of permutations $w \in \mathfrak{S}_4$ with 3 descents. That is, considering any permutation $w = w_1 w_2 w_3 w_4$, with $w_1 > w_2 > w_3 > w_4$. There is only one possible permutation. The descent set of this permutation is $D(w) = \{1, 2, 3\}$ and $A(4, 4) = 1$. The corresponding term in $A_4(x)$ is $A(4, 4) x^4 = x^4$.

2nd, find $A(4, 3)$, which is $\sharp$ of permutations $w \in \mathfrak{S}_4$ with 2 descents.

Suppose any set permutation $w = w_1 w_2 w_3 w_4 \in \mathfrak{S}_4$, there are 3 gaps between letters in the permutation.

Since the permutation has 2 descents, choosing any 2 gaps to put the greater symbol $>$, there are $\binom{3}{2}$ ways to choose 2 gaps from 3 gaps, that is, $(1,2), (1,3), (2,3)$. Different ways to put $>$ corresponds to different descent sets, which are integer sets $S = \{i,j\}_<$, for $1 \leq i, j \leq 3$.

Thus, find $\sharp\ \{w \in \mathfrak{S}_4\ :\ D(w) = S\}$, which by definition is $\beta_n(S)$, where $S = \{i,j\}$, for $1 \leq i < j \leq 3$.

That is, find $\beta_4(1,2),\quad \beta_4(1,3),\quad \beta_4(2,3)$.

For any finite integer set $S = \{s_1, s_2, \ldots, s_k\}_< \subseteq [n-1]$,

from the definition of $\alpha_n(S) = \sharp\{w \in \mathfrak{S}_n : D(w) \subseteq S\}$,
$\beta_n(S) = \sharp\{w \in \mathfrak{S}_n : D(w) = S\}$,

and from the above proposition of finding $\alpha_n(S)$ by

$$\alpha(S) = \binom{n}{s_1,\, s_2 - s_1,\, s_3 - s_2,\, \ldots,\, n - s_k}$$

Thus, if $S = \{i,j\}$, for $1 \leq i < j \leq 3$, which means descents sets containing 2 descents,

then, $\beta_4(i,j) = \alpha_4(i,j) - \alpha_4(i) - \alpha_4(j) + \alpha_4(\emptyset)$

It is already known from Example 1 that

$$\begin{aligned}
\alpha_n(\emptyset) &= \quad \sharp \text{ of permutations } w \in \mathfrak{S}_n \text{ with } D(w) = \emptyset \\
&= \quad 1
\end{aligned}$$

from the above proposition for finding $\alpha_n(S)$,

$$\begin{aligned}
\alpha_4(i) &= \sharp\{\ w \in \mathfrak{S}_4\ :\ D(w) \subseteq \{i\}\ \} \\
&= \binom{4}{i,\, 4-i} = \binom{4}{i}
\end{aligned}$$

Similarly,

$$\alpha_4(j) \quad = \quad \binom{4}{j}$$

and

$$\alpha_4(i,j) \quad = \quad \binom{4}{i,\, j-i,\, 4-j}, \text{ for } 1 \leq i < j \leq 3$$

Then, sum up $\beta_4(i,j)$, for $1 \le i < j \le 3$ to obtain the coefficient $A(4,3)$, that is,

$$A(4,3) \;=\; \sum_{1 \le i < j \le 3} \beta_4(i,j), \quad \text{where} \quad \beta_4(i,j) = \sharp\{w \in \mathfrak{S}_4 : D(w) = \{i,j\}\}$$

It can be generalized to the formula:

$$
\begin{aligned}
A(n,3) \;&=\; \sharp\{\, w \in \mathfrak{S}_4 \;:\; w \text{ has 2 descents }\} \\
&=\; \sum_{1 \le i < j \le n-1} \beta_n(i,j), \text{ where } \beta_n(i,j) = \sharp\{w \in \mathfrak{S}_n : D(w) = \{i,j\}\} \\
&=\; \sum_{1 \le i < j \le n-1} (\alpha_n(i,j) - \alpha_n(i) - \alpha_n(j) + \alpha_n(\emptyset))
\end{aligned}
$$

Then, it can be generalized to the formula:

$$
\begin{aligned}
A(n,k) \;&=\; \sharp\{\, w \in \mathfrak{S}_4 \;:\; w \text{ has } k-1 \text{ descents }\} \\
&=\; \sum_{1 \le i_1 < i_2 < \ldots < i_{k-1} \le n-1} \beta_n(i_1, i_2, \ldots, i_{k-1}), \\
&\qquad \text{where } \beta_n(i_1, i_2, \ldots, i_{k-1}) = \sharp\{w \in \mathfrak{S}_n : D(w) = \{i_1, i_2, \ldots, i_{k-1}\}\}
\end{aligned}
$$

3.6 How to extract the leading terms of a generating function

Generating function is a way to represent a sequence of numbers, it is useful to study the way to extract coefficients from generating functions. In this section, one method of extracting coefficients from the relationship of several generating functions instead of one generating function is given, the method idea is taken from [44].

Idea:

Use generating function to get recurrence relation, Then, compute with the recurrence.

If $b(z) = a(z) \cdot \frac{1}{1-z^M}$,

that is, $\sum_{n \ge 0} b_n z^n - \sum_{n \ge 0} b_n z^n z^M = \sum_{n \ge 0} a_n z^n$

Then, the recurrence relation is taken by extracting coefficients of x^n on both sides of the above equation,

so, $b_n - b_{n-M} = a_n$ for $n \ge M$

Then, give an easy way to compute the leading terms by hand. If we know $[z^n]\frac{1}{1-z}$ and we want to find $[z^n]\frac{1}{1-z} \cdot \frac{1}{1-z^5}$

Firstly,

$$a_n = [z^n]\frac{1}{1-z} = 1 + z + z^2 + z^3 + \ldots$$

and

$$\{a_n\}_{n=0}^{\infty} = 1, 1, 1, \ldots$$

Secondly, $b_n = [z^n]\frac{1}{1-z} \cdot \frac{1}{1-z^5}$, it is known that $b_n = b_{n-M} + a_n$, here, $M = 5$ So, $b_n = b_{n-5} + a_n$

That is, $b_0 = b_{0-5} + a_0 = a_0, \quad b_1 = b_{1-5} + a_1 = a_1$

$\ldots\ldots$

$b_5 = b_0 + a_5 = a_0 + a_5, \ b_5 = b_1 + a_5 = a_1 + a_5,$

therefore, $[z^n]\frac{1}{1-z} = a_0 \ldots a_5 \ldots a_{10} \ldots a_{15} \ldots = 1 \ldots 1 \ldots 1 \ldots 1 \ldots$

The corresponding sequence $\{b_n\}_{n=0}^{\infty}$ is $[z^n]\frac{1}{1-z}\frac{1}{1-z^5} = b_0 \ldots b_5 \ldots b_{10} \ldots b_{15} \ldots = 1, \ldots, 2, \ldots, 3, \ldots, 4, \ldots$, that is, each number appears $M = 5$ times.

It can be generalized that the leading coefficients of the generating function $\frac{1}{1-z}\frac{1}{1-z^M}$ is obtained from $bn = b_{n-M} + a_n$, that is,

$$[z^n]\frac{1}{1-z}\frac{1}{1-z^M} = b_0 \ldots b_M \ldots b_{2M} \ldots b_{3M} = 1, \ldots, 2, \ldots, 3, \ldots, 4, \ldots$$

Where each number appears M times.

3.7   Case study/path transform example

Case I. We start by treating a simple case of restricted paths with no jumps down. More exactly, we assume a sequence or path $S_j = S_0 u^j$, compute

$$\begin{aligned} S(z) &= \sum_{j=0}^{\infty} S_j z^j \\ &= \sum_{j=0}^{\infty} S_0 u^j z^j \\ &= \frac{S_0}{1 - uz} \end{aligned}$$

The methodology consists in

(i) computing the transformed generating function;
(ii) and thus computing the associated path.

Example 3.7.1 *Symbolic sequence transformation. In this case*

$$
\begin{aligned}
\hat{S}(z) &= \frac{1}{1 - S(Z)} \\
&= \frac{1}{1 - \frac{S_0}{1 - uz}} \\
&= \frac{1 - uz - S_0 + S_0}{1 - uz - S_0} \\
&= 1 + \frac{S_0}{1 - uz - S_0} \\
&= 1 + \frac{S_0}{1 - S_0} \frac{1}{1 - z\frac{u}{1 - S_0}} \\
&= 1 + \frac{S_0}{1 - S_0} \sum_{j=0}^{\infty} \left(z \frac{u}{1 - S_0} \right)^j
\end{aligned}
$$

The answer is that the symbolic transformed share price is derived from the addition of two generating functions(addition property),

One generating function is $A(Z) \equiv 1$ so, the sequence is $(1, 0, 0, \ldots)$.

Another generating function is

$$
\frac{S_0}{1 - S_0} \frac{1}{1 - z\frac{u}{1 - S_0}}
$$

which generates the sequence

$$
\frac{S_0}{1 - S_0} \left(z \frac{u}{1 - S_0} \right)^j .
$$

This is the upper share price path with initial share value $\hat{S}_0 = \frac{S_0}{1 - S_0}$ and the symbolic transformed u jump $\hat{u} = \frac{u}{1 - S_0}$

Similar computation for this example
Differentiation path transform;

Example 3.7.2 *Differentiation path transformation. In this case*

$$
\begin{aligned}
\hat{S}(z) &= S'(z) \\
&= \frac{S_0 u}{(1 - uz)^2} \\
&= S_0 u \sum_{j \geq 0} \left(\sum_{k \geq 0}^{j} u^k u^{n-k} \right) z^j \\
&= S_0 u \sum_{j \geq 0} \left((j+1) u^j \right) z^j
\end{aligned}
$$

The answer is that the differentiation transformed share price is derived from a generating function,

the generating function is

$$S_0 u \sum_{j \geq 0} \left((j+1) u^j \right) z^j$$

which generates the sequence

$$S_0 u \left((j+1) u^j \right) z^j \ .$$

This is the upper share price path with initial share value $\hat{S}_0 = S_0 u$ and the differentiation transformed u jump to $\hat{u} = \left(\frac{t+1}{t} \right) u = u + \frac{u}{t}$, with $t \geq 1$.

Remark. Notice that in the symbolic path transformation, if we allow exactly one jump down, the calculations can become more involved.

Example 3.7.3 *Assume that $S_1 = S_0 d, S_{j+1} = S_j u, \ j = 1, 2, \ldots$*

Consider a new path with initial share value $\hat{S}_0 = \frac{S_0 d}{u}$ and $S_j = S_0 d u^{j-1}, \ j = 0, 1, 2, \ldots$, its path generating function has a constant difference $\hat{S}_0 - S_0$ compared to the original exactly one jump down path happening in the first step, since S_0 is a fixed value.

Then, the path generating function for the first step jump down path is

$$
\begin{aligned}
S(z) &= \sum_{j=0}^{\infty} S_j z^j - \left(\hat{S}_0 - S_0 \right) \\
&= \frac{S_0 d}{u} \sum_{j=0}^{\infty} u^j z^j - \left(\hat{S}_0 - S_0 \right) \\
&= \frac{S_0 d}{u} \left(\frac{1}{1 - uz} \right) - \left(\hat{S}_0 - S_0 \right)
\end{aligned}
$$

So, let the constant $\hat{S}_0 - S_0 = D$, the symbolic transformed generating function is

$$
\begin{aligned}
\hat{S}(z) \;&=\; \frac{1}{1 - S(z)} \\[2mm]
&=\; \frac{1}{1 - D - \frac{S_0 d}{u}\left(\frac{1}{1-uz}\right)} \\[2mm]
&\qquad \text{Let a constant } C = \frac{S_0 d}{u} \\[1mm]
&\qquad \text{Let a constant } m = 1 - D \\[2mm]
&=\; \frac{1}{m - \frac{C}{1-uz}} \\[2mm]
&=\; \frac{1}{m}\left(\frac{m(1-uz) - C + C}{m(1-uz) - C}\right) \\[2mm]
&=\; \frac{1}{m}\left(1 + \frac{C}{m}\left(\frac{1}{1 - uz - C/m}\right)\right) \\[2mm]
&=\; \frac{1}{m}\left(1 + \frac{C/m}{1 - C/m}\frac{1}{1 - z\frac{u}{1-C/m}}\right) \\[2mm]
&=\; \frac{1}{m} + \frac{C}{1 - C/m}\sum_{j=0}^{\infty}\left(z\frac{u}{1-C/m}\right)^{j}.
\end{aligned}
$$

The answer is that the symbolic transformed share price is derived from the addition of two generating functions(addition property),

One generating function is $A(Z) \equiv \frac{1}{m}$, with constant $m = 1 - D = 1 - \frac{S_0 d}{u} + S_0$ so, the sequence is $(\frac{1}{m}, 0, 0, \ldots)$.

Another generating function is

$$
\frac{C}{1 - C/m}\frac{1}{1 - z\frac{u}{1-C/m}}
$$

which generates the sequence

$$
\frac{C}{1 - C/m}\left(z\frac{u}{1-C/m}\right)^{j}.
$$

This is the upper share price path with initial share value $\tilde{S}_0 = \frac{C}{1-C/m}$, where $C = \frac{S_0 d}{u}$, $m = 1 - D = 1 - \left(\frac{S_0 d}{u} - S_0\right)$, D is the difference between the original first down path initial state S_0 and the changed upper path initial state $C = \frac{S_0 d}{u}$.

the final symbolic transformed path is an upper share price path with u changed to $\hat{u} = \frac{u}{1-C/m}$

Remark. Similarly, if the exact one down step happens on a time step j, that is,

$$
S_{j-1} \to S_j = down\ step
$$

we can separate the path to two paths; the first path is a sequence $(S_0, S_1, S_2, \ldots, S_{j-1})$ *and the second path is* $(S_j, S_{j+1}, S_{j+2}, \ldots)$.

Then, we can use the same methodology in the example 3, *which the exact one down step happens on a time step* 1. *It can be see that the calculation can become more involved, but we might only need change the corresponding constants compared to the first step down step symbolic transformation path.*

Case II. Consider discrete binomial market with jumps u and d, an interest rate r. Assuming no arbitrage condition holds, it is $d \leq 1 + r \leq u$. Fix k and define the option price sequences

$$x_n = \mathrm{OP}(S_n^k).$$

where $S_n = S_0 Y_1 \ldots Y_n$, Y_i are i.i.d with $Q(Y = u) = q_u$ $Q(Y = d) = q_d$,

$$q_u = \frac{1 + r - d}{u - d}, q_d = 1 - q_u$$

Then,

$$
\begin{aligned}
x_n &= \mathrm{OP}(S_n^k) \\
&= \frac{E_Q S_n^k}{(1 + r)^n} \\
&= S_0^k \left(\frac{E_Q Y^k}{1 + r} \right)^n \\
&= Ba^n = g_n
\end{aligned}
$$

where,

$$B = S_0^k, \ a = a(r, u, d) = \frac{u^k q_u + d^k q_d}{1 + r}$$

Then, Greeks are derived from the scaled derivative of generating function $x(z)$. because,

$$\frac{\partial g_n}{\partial \theta} = nBa^{n-1} a_\theta'.$$

Chapter 4

Path Calculations/Interpretations

4.1 Count unrestricted lattice paths with two directions

In this section, the enumeration of unrestricted lattice paths is discussed. In question 1, counting lattice paths containing 'up' and 'down' steps of fixed length N is like one dimension counting problem. The problem is solved in three ways. Firstly, a simple combinatorial method is used to solve the question by constructing a recurrence equation. Then, we use the symbolic method which is motivated by Flajolet and Sedgewick in the book [14, Chapter 3.8] [14, Chapter 5.2], solve the question by constructing lattice path in two ways, the construction method is motivated from the bitstring construction which is stated on page 226 in the book [14, Chapter 5.2]. The method of counting unrestrited lattice path of fixed length can also be referred to [23, Example 2.5].

Assuming each step in a lattice path has two step choices {up, right}.

[1-Dimension]

Question1: How many lattice paths of length N ?

Remark: A lattice path of length N is a sequence of 'up' and 'right' steps with length 'N'. It can be solved easily by combinatorial knowledge and get result 2^N, or we can decuce a recursive definition for the number of sought lattice paths, that is,

$$a_n = (up)a_{n-1} + (right)a_{n-1},$$

and get

$$a_n = 2a_{n-1},$$

solve the recurrence equation, we can get $a_n = 2^n$.

Here, we would like to use symbolic method to directly make a construction for the combinatorial classes, and then apply transfer theorem to get equations for generating functions. Two method of construction for the combinatorial class will be given in the answer.

Answer:

1st, Starting from a definition of the combinatorial class.

Denote the class of lattice paths employing two steps {up, right} by '$\mathcal{W}$'.

2nd, Define a notion of size function for the combinatorial class '$\mathcal{W}$'.

From the question, define

$$|w| \equiv \text{ the number of bits in a lattice path } w \in \mathcal{W}$$

3rd, From the question, define sought number,

$$W_N \equiv \text{ the number of lattice paths of length } N \ (N \text{ bits})$$

4th, Use symbolic method, define a GF for the combinatorial class of the lattice paths, its GF is the sum over all members of the class '$\mathcal{W}$',

$$W(z) \quad \overset{\text{define}}{=\!=} \quad \sum_{w \in \mathcal{W}} z^{|w|}$$

$$\overset{\substack{\text{collect all terms with} \\ \text{the same size exponent}}}{=\!=} \quad \sum_{N \geq 0} W_N \cdot z^N$$

5th, Define a combinatorial construction of the combinatorial class '$\mathcal{W}$'.

Firstly, define basic building blocks. (1) By definition, there are two types of bits; one is 'up' bit, the other is 'right' bit.
(2) Based on the definition of size function of the combinatorial class, each bit builds up an atomic class;
'up' bit builds up a class denoted by 'Z_{up}', and contributes size '1' to the number of bits in a lattice path $w \in \mathcal{W}$, its GF is $z^1 = z$.
'Right' bit builds up a class, denoted by 'Z_{right}', and contributes size '1' to the number of bits in a lattice path $w \in \mathcal{W}$, its GF is $z^1 = z$. **Secondly,** Define a combinatorial construction for the combinatorial class '$\mathcal{W}$'. There are two methods of construction.

Method1: [14, Chapter 3.8, 5.2]

Definition 50 *A lattice path $w \in \mathcal{W}$ is a sequence of 'up' steps and 'right' steps.*

Construction:

$$\mathcal{W} = \text{SEQ}(Z_{\text{up}} + Z_{\text{right}})$$

Apply transfer theorem:

$$W(z) = \frac{1}{1 - 2z}$$

Expand coefficient:

$$
\begin{aligned}
W_N &= [z^N]W(z) = [z^N]\left(\frac{1}{1 - 2z}\right) \\
&= [z^N]\sum_{k \geq 0} 2^k z^k \\
&= 2^N
\end{aligned}
$$

<u>Method2</u>: [14, Chapter 3.8, 5.2]

Definition 51 *A lattice path $w \in \mathcal{W}$ is either a path of zero length or a path of non-zero length.*

(1) If assuming a path of non-zero length has a length 'N', using first passage decomposition technique, it can be separated into two disjoint classes; one class is 'up' bit followed by a path of non-zero length 'N − 1', the other is 'right' bit followed by a path of non-zero length 'N − 1'.

(2) 'Up' bit contributes size '1' to the size function and constructs an atomic class, denoted by 'Z_{up}', its GF is $z^1 = z$.
'Right' bit contributes size '1' to the size function and constructs an atomic class, denoted by 'Z_{right}', its GF is also $z^1 = z$.

(3) as 'N', 'N-1' is not a fixed number, the paths of a length 'N' and 'N − 1' construct the same combinatorial class '$\mathcal{W}$'.

(4) Besides, a path of zero length construct a neutral class, which contains a neutral object that contributes size(length) zero('0') to the number of bits in a path $w \in \mathcal{W}$, its GF is $z^0 = 1$

Therefore, obtain construction:

$$\mathcal{W} = E + Z_{\mathrm{up}} \times \mathcal{W} + Z_{\mathrm{right}} \times \mathcal{W}$$

Apply transfer theorem:

$$W(z) = 1 + zW(z) + zW(z)$$

$$\implies W(z) = 1 + 2zW(z)$$

$$\implies W(z) = \frac{1}{1 - 2z}$$

Expand coefficient:

$$
\begin{aligned}
W_N &= [z^N]W(z) = [z^N]\left(\frac{1}{1 - 2z}\right) \\
&= [z^N]\sum_{k \geq 0} 2^k z^k \\
&= 2^N
\end{aligned}
$$

In question 2, counting lattice paths containing 'up' and 'down' steps of fixed length N and fixed number of up steps k is like two dimensions counting problem. The prolem is solved by applying the symbolic method for Parameters which is motivated by Flajolet and Sedgewick on page 127 and page 243 in the book [14, Chapter 5.4].

[2-Dimension]

Question2: How many lattice paths of length N have exactly k bits that are 'up' steps?

Answer:

1st, Starting from a definition of the combinatorial class.

Denote the class of lattice paths employing two steps {up, right} of length N by '$\mathcal{W}_N$'.

2nd, Define a notion of size function for each object 'w' in the combinatorial class '$\mathcal{W}_N$', $w \in \mathcal{W}_N$. Define size,

$$|w_N| \equiv \text{ the number of 'up' bits in a lattice path } w_N \in \mathcal{W}_N$$

3rd, Define the question, From the question, define

$$W_{Nk} \equiv \text{ the number of lattice paths of length } N \text{ with } k \text{ 'up' bits}$$

4th, Use symbolic method, define a GF for the combinatorial class $\mathcal{W}_N$, its GF is the sum over all members of the class $\mathcal{W}_N$, Here, we use OGF.

$$W_N(z) \stackrel{\text{define}}{\equiv} \sum_{w_N \in \mathcal{W}_N} z^{|w_N|}$$

Each object of size k corresponds to a term z^k in the GF, collect all the terms with the same size exponent to expose counts W_{Nk}, that is, the number of lattice paths of length N with k 'up' bits, therefore,

$$W_N(z) \stackrel{\text{define}}{\equiv} \sum_{w_N \in W_N} z^{|w_N|}$$
$$= \sum_{0 \leq k \leq N} W_{Nk} z^k$$

As no object of size($\sharp$ up bits) greater than N in a path $w_N \in \mathcal{W}_N$, in the GF, the corresponding term z^k, *when $k > N$* will be zero. We have defined $W_{Nk} \equiv$ the number of lattice paths of length N with k 'up' bits, , that is, when $k > N$, in the GF, $W_{Nk} = 0$, so, we can take $k \geq 0$ in the GF.

$$W_N(z) \stackrel{\text{define}}{\equiv} \sum_{w_N \in W_N} z^{|w_N|}$$
$$= \sum_{0 \leq k \leq N} W_{Nk} z^k$$
$$= \sum_{k \geq 0} W_{Nk} z^k$$

5th, Define a combinatorial construction for the combinatorial class '$\mathcal{W}_N$' Firstly, define building blocks. By definition, there are two types of bits; one is 'up' bit, the other is 'right' bit.

The construction process can be done in two steps.

Firstly, define basic building blocks for the combinatorial class. (1) By definition, there are two types of bits; one is 'up' bit, the other is 'right' bit. (2) Based on the size function of the combinatorial class $\mathcal{W}_N$,

$$|w_N| \equiv \sharp \text{ of 'up' bits in a lattice path } w_N \in \mathcal{W}_N$$

(3) 'up' bit builds up an atomic class, which contributes size '1' to the number of 'up' bits in a path $w_N \in \mathcal{W}_N$, denote the atomic class by 'Z_{up}', its GF is $z^1 = z$.

(4) 'right' bit builds up a neutral class, which contributes size '0' to the number of 'up' bits in a path $w_N \in \mathcal{W}_N$, denote the neutral class by 'Z_{right}', its GF is $z^0 = 1$.

Secondly, Define a combinatorial construction for the combinatorial class '$\mathcal{W}_N$'.

Definition 52 *A lattice path $w_N \in \mathcal{W}_N$ is a path of non-zero length N.*

Remark: Here, the non-zero length N is a fixed number, therefore, paths of length 'N' and paths of length '$N-1$' would construct different combinatorial classes, that is, constructing the classes $\mathcal{W}_N$ and $\mathcal{W}_{N-1}$ respectively.

(1) If assume a path of length N has exactly 'k' 'up' bits, using first passage decomposition technique, it can be either a 'up' bit followed by a path of length '$N-1$' with '$k-1$' 'up' bits or a 'right' bit followed by a path of length '$N-1$' with 'k' 'up' bits.

(2) A 'up' bit constructs an atomic class 'Z_{up}', its GF is $z^1 = z$.

(3) Paths of length '$N-1$' with '$k-1$' 'up' bits build up a combinatorial class $\mathcal{W}_{N-1}$, here, '$N-1$' is a fixed number, but '$k-1$' is not a fixed number.

(4) A 'right' bit constructs a neutral class 'Z_{right}', its GF is $z^0 = 1$.

(5) Paths of length '$N-1$' with 'k' 'up' bits build up a combinatorial class $\mathcal{W}_{N-1}$, here, '$N-1$' is a fixed number, but 'k' is not a fixed number.

Therefore, using three common operations, we obtain construction:

$$\mathcal{W}_N = Z_{up} \times \mathcal{W}_{N-1} + Z_{right} \times \mathcal{W}_{N-1}$$

Apply transfer theorem of operations from combinatorial classes to GFs,

$$W_N(z) = zW_{N-1}(z) + 1 \cdot W_{N-1}(z)$$

then,

$$W_N(z) = (1+z)W_{N-1}(z)$$

By induction,

$$W_N(z) = (1+z)^N W_0(z)$$

As in the combinatorial class $\mathcal{W}_0$, the lattice of length '0' will have '0' 'up' bits, so, it contributes '0' for the size function($\sharp$ of up bits) of the combinatorial class $\mathcal{W}_0$, its GF is $z^0 = 1$, that is, $\mathcal{W}_0(z) = 1$.

Therefore,

$$W_N(z) = (1+z)^N$$

Expand this function with the binomial theorem,

$$(1+z)^N \qquad = \qquad \sum_{0 \le k \le N} \binom{N}{k} z^k$$

$$\text{when } k \ge N, \ \binom{N}{k} = 0 \quad \sum_{k \ge 0} \binom{N}{k} z^k$$

Therefore,

$$[z^N](1+z)^N = \binom{N}{k}$$

Remark: (1) If the question is to count the number of paths from $(0,0)$ to (j,k), as the length of each path is $N = j + k$, the answer will be $\binom{j+k}{k}$.

(2) In the similar method, the symbolic method can be generalized to dimensions greater than 2. For example, to count the number of self-avoiding paths from $(0,0,0)$ to (i,j,k) with three choices up, right, front.

4.2   Count restricted lattice paths with two directions

In this section, a question of counting the number of restricted lattice paths advised by supervisor is solved. The question solution and definition 53 of dyck path are motivated by Wallner in the thesis [23, Example 2.7].

Question: How many lattice paths from $(0,0)$ to (N,N) employing two steps {up, right} and not going above the diagonal line?

Remark: The combinatorial class contains restricted lattice paths. It is not useful to define the restriction as a size function, so, we still define the length of the path as its size function, from another perspective, it is a famous class of Dyck paths. Use first and last passage decomposition technique to define the particular class.

Answer:

1st, Starting from a definition of the combinatorial class. Denote the class of lattice paths employing two steps {up, right} and not going above diagonal by 'D'.

2nd, Define a notion of size function for the combinatorial class 'D'. As we talked in the Remark, define the length of path as its size function, for each object $d \in D$,

$$\begin{aligned} |d| &\equiv \text{ the number of bits in a lattice path } \quad d \in D \\ &\equiv \text{ the length of a lattice path } \quad d \in D \end{aligned}$$

3rd, Define sought variable in the question, from the question, the Dyck path can only be of even number length, therefore, define,

$$D_{2N} \equiv \text{ the number of Dyck paths in } D \text{ of length } 2N \text{ } (2N\text{bits})$$

4th, Use symbolic method , define a GF for the combinatorial class D, its GF is the sum over all members of the class D.

$$D(z) \stackrel{\text{define}}{\equiv} \sum_{d \in D} z^{|d|}$$

Each object of size $2N$ corresponds to a term z^{2N} in the GF, collect all the terms with the same size exponent to expose counts D_{2N}, that is, the number of Dyck paths of length $2N$, therefore,

$$D(z) \stackrel{\text{define}}{\equiv} \sum_{d \in D} z^{|d|}$$
$$= \sum_{N \geq 0} D_{2N} z^{2N}$$

5th, Define a combinatorial construction for the combinatorial class of Dyck paths D.

Firstly, define basic building blocks for the combinatorial class. (1) By definition, there are two types of bits in a dych path; one is 'up' bit, the other is 'right' bit.
(2) Based on the size function of the combinatorial class D.

$$
\begin{aligned}
|d| &\equiv \text{ the number of bits in a lattice path } d \in D \\
&\equiv \text{ the length of a lattice path } d \in D
\end{aligned}
$$

(3) 'up' bit builds up an atomic class, which contributes size '1' to the number of 'up' bits in a Dyck path $d \in D$, denote the atomic class by 'Z_{up}', its GF is $z^1 = z$.

(4) 'right' bit builds up a atomic class, which contributes size '1' to the number of bits in a lattice path $d \in D$, denote the atomic class by 'Z_{right}', its GF is $z^1 = z$.

Secondly, Define a combinatorial construction for the combinatorial class 'D' using first passage and last passage decomposition technique.

Definition 53 *A Dyck path $d \in D$ is either a path of zero length or a path of non-zero length.*

(1) If assume a Dyck path of non-zero length $2N$, starting from $(0,0)$, the Dyck path can touch the diagonal at several times. Suppose the first touching diagonal point is (k,k) and the Dyck path can be decomposed into two part paths. The first part path is starting from $(0,0)$ to the first touching point

(k, k), it is a path of length $2k$. The second part path is starting from the first touching point (k, k) to the end point (N, N), it is a path of length $2(N - k)$.

(2) Next, we construct the two parts paths. For the first part path, it will never touch diagonal except for the starting point $(0, 0)$ and the first touching point (k, k). It is obvious that its first step must be 'Right' bit, and its last step must be 'Up' bit. The first part path of length '2k' without the initial step 'Right' and its final step 'Up' bit construct again a Dyck path of length '$2(k - 1)$'.

For the second-part path starting from the first touching point (k, k) to the end point (N, N) is a Dyck path of length $2(N - k)$

(3) Use decomposition technique, a Dyck path of non-zero length '2N' is a 'Right' bit followed by a dyck path of length '$2(k - 1)$', followed by a 'Up' bit and followed by a Dyck path of length '$2(N - k)$'.

(4) '$k - 1$' and '$N - k$' is not fixed numbers, it is obvious that Dyck paths of length '$2(k - 1)$' and length '$2(N - k)$' build up the same combinatorial class D. Also, Dyck paths of length 'N' build up a combinatorial class 'D'.

(5) A 'up' bit constructs an atomic class 'Z_{up}', its GF is $z^1 = z$.

(6) A 'right' bit constructs a atomic class 'Z_{right}', its GF is $z^1 = z$.

(7) Besides, a Dyck path of zero length constructs a neutral class which contains a neutral object that contributes size(length) '0' to the number of bits in a Dyck path $d \in D$, its GF is $z^0 = 1$

Therefore, we obtain combinatorial construction:

$$D = E + Z_{right} \times D \times Z_{up} \times D$$

6th, Apply transfer theorem of operations from combinatorial classes to GFs,

$$D(z) = 1 + zD(z)zD(z)$$

then,

$$D(z) = 1 + z^2 D(z)^2$$

By quadratic formula,

$$D(z) = \frac{1 \pm \sqrt{1 - 4z^2}}{2z^2}$$

$$z^2 D(z) = \frac{1 \pm \sqrt{1 - 4z^2}}{2}$$

take $z = 0$, then $LHS = 0$, RHS should also equal to 0 when $z = 0$.

So, we choose $z^2 D(z) = \frac{1 - \sqrt{1 - 4z^2}}{2}$.

7th, Expand the coefficient of GF $D(z)$ to obtain the explicit formula for the number of Dyck paths of length 2N, D_{2N}.
Note: the Dyck path can only be of even length.

Expand via generalized binomial theorem,

$$
\begin{aligned}
z^2 D(z) &= \frac{1}{2} - \frac{1}{2}\left(1 - 4z^2\right)^{\frac{1}{2}} \\
&= \frac{1}{2} - \frac{1}{2}\sum_{N \geq 0}\left(\binom{\frac{1}{2}}{N}(-4z^2)^N\right) \\
&= -\frac{1}{2}\sum_{N \geq 1}\binom{\frac{1}{2}}{N}(-4)^N z^{2N}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
D_{2N} &= [z^{2N}]D(z) = [z^{2N+2}]\left(\frac{1 - \sqrt{1 - 4z^2}}{2}\right) \\
&= [z^{2N+2}]\left(-\frac{1}{2}\sum_{N \geq 1}\binom{\frac{1}{2}}{N}(-4)^N z^{2N}\right) \\
&= -\frac{1}{2}\binom{\frac{1}{2}}{N+1}(-4)^{N+1} \\
&= -\frac{1}{2}\cdot\frac{\frac{1}{2}(\frac{1}{2}-1)(\frac{1}{2}-2)\ldots(\frac{1}{2}-N)(-4)^N\cdot(-4)}{(N+1)!} \\
&= \frac{(-\frac{1}{2})(-\frac{3}{2})(-\frac{5}{2})\ldots(-\frac{2N-1}{2})(-4)^N}{(N+1)!} \\
\underset{\substack{\text{Distribute}(-2)^N \\ \text{among factors}}}{=} & \frac{1\cdot3\cdot5\cdots(2N-1)\cdot2^N}{(N+1)!} \\
&= \frac{1}{N+1}\cdot\frac{1\cdot3\cdot5\cdots(2N-1)}{N!}\cdot\frac{2\cdot4\cdot6\cdots(2N)}{1\cdot2\cdot3\cdots N} \\
&= \frac{1}{N+1}\binom{2N}{N}
\end{aligned}
$$

4.3   Count self-avoiding walks with three directions

In this section, an enumeration of self-avoiding walks with three directions are presented. The enumeration method and presentation are mainly taken from the section 7 and section 10 in the paper "Enumerating up-side self-avoiding walks on integer lattices" by Williams [24]. In addition, we added the detailed explanation of derivation of the general recursive method before deriving the two variable generating function $\mathcal{G}(t, v)$ and extracting the coefficient $g(n, m)$.

Question: On square lattice, count the number of $N-$ step self-avoiding walks of fixed height $'m'$, and starting from the original point $(0, 0)$, which has directions {up, right, left}?

Answer: Assume the height 'm' corresponds to Y-coordinate of the ending point of the self-avoiding walks. Use USSAWS as the abbreviation of up-side self-avoiding walks in the following answer.

1. Divide N-step paths of fixed height 'm' into disjoint set of paths according to the last step of all paths, that is,

   Let

   $$
   \begin{aligned}
   U_{[n,m]} &= \text{N-step USSAWS ending with a step 'Up' and a height 'm'} \\
   \overrightarrow{S}_{[n,m]} &= \text{N-step USSAWS ending with a step 'Right' and a height 'm'} \\
   \overleftarrow{S}_{[n,m]} &= \text{N-step USSAWS ending with a step 'Left' and a height 'm'}
   \end{aligned}
   $$

2. Define counts for each disjoint sets.

   Let

   $$
   \begin{aligned}
   u(n,m) &= \text{the number of } U_{[n,m]}s \\
   \overrightarrow{s}(n,m) &= \text{the number of } \overrightarrow{S}_{[n,m]}s \\
   \overleftarrow{s}(n,m) &= \text{the number of } \overleftarrow{S}_{[n,m]}s \\
   s(n,m) &= \text{the number of N-step USSAWS ending with sideways} \\
   &\quad \text{and with height 'm'} \\
   g(n,m) &= \text{the number of N-step USSAWS ending with fixed height 'm'}
   \end{aligned}
   $$

   $$
   \begin{aligned}
   So, \quad s(n,m) &= \overrightarrow{s}(n,m) + \overleftarrow{s}(n,m) \\
   g(n,m) &= s(n,m) + u(n,m)
   \end{aligned}
   $$

3. Using GF for two variables, define two variable GFs for each counts sequence $\{u(n,m)\}_{n\geq 0,m\geq 0}$, $\{s(n,m)\}_{n\geq 0,m\geq 0}$, $\{g(n,m)\}_{n\geq 0,m\geq 0}$ respectively. That is,

$$
\mathcal{U}(t,v) = \sum_{m\geq 0,n\geq 0}^{\infty} u(n,m)t^m v^n
$$

$$
\mathcal{S}(t,v) = \sum_{m\geq 0,n\geq 0}^{\infty} s(n,m)t^m v^n
$$

$$
\mathcal{G}(t,v) = \sum_{m\geq 0,n\geq 0}^{\infty} g(n,m)t^m v^n
$$

4. Construct diagram for the combinatorial set on the square lattice by adding one step to an walk in the set $U_{[n,m]}$, $\overrightarrow{S}_{[n,m]}$, $\overleftarrow{S}_{[n,m]}$ respectively, we get,

$$
U_{[n,m]} \longrightarrow
\begin{array}{l}
\nearrow \quad U_{[n+1,m+1]} \\
\overrightarrow{S}_{[n+1,m]} \\
\searrow \quad \overleftarrow{S}_{[n+1,m]}
\end{array}
$$

$$\nearrow \quad U_{[n+1,m+1]}$$
$$\overrightarrow{S}_{[n,m]} \quad \longrightarrow \quad \overrightarrow{S}_{[n+1,m]}$$

$$\nearrow \quad U_{[n+1,m+1]}$$
$$\overleftarrow{S}_{[n,m]} \quad \longrightarrow \quad \overleftarrow{S}_{[n+1,m]}$$

5. Combinatorial construction using common operation for combinatorial class, that is, product $\times$ and disjoint union $+$. From the diagram constructed in step 4, it can be seen that for every $U_{[n+1,m+1]}$ is an $U_{[n,m]}$ or $\overrightarrow{S}_{[n,m]}$, or $\overleftarrow{S}_{[n,m]}$ followed by an extra Up step. For every $\overrightarrow{S}_{[n+1,m]}$ is an $U_{[n,m]}$ or $\overrightarrow{S}_{[n,m]}$, followed by an extra Right step. For every $\overleftarrow{S}_{[n+1,m]}$ is an $U_{[n,m]}$ or $\overleftarrow{S}_{[n,m]}$, followed by an extra Left step. That is,

$$U_{[n,m]} = U_{[n,m]} \times Up + \overrightarrow{S}_{[n,m]} \times Up + \overleftarrow{S}_{[n,m]} \times Up$$
$$\overrightarrow{S}_{[n,m]} = U_{[n,m]} \times Right + \overrightarrow{S}_{[n,m]} \times Right$$
$$\overleftarrow{S}_{[n,m]} = U_{[n,m]} \times Left + \overleftarrow{S}_{[n,m]} \times Left$$

Therefore, when doing counting for the combinatorial classes, we obtain recursive equations,

$$\begin{cases} u(n,m) = u(n-1,m-1) + \overrightarrow{s}(n-1,m-1) + \overleftarrow{s}(n-1,m-1) \\ \overrightarrow{s}(n,m) = u(n-1,m) + \overrightarrow{s}(n-1,m) \\ \overleftarrow{s}(n,m) = u(n-1,m) + \overleftarrow{s}(n-1,m) \end{cases}$$

to reduce the calculation of recursive equations, we simplify the equation using $s(n,m) = \overrightarrow{s}(n,m) + \overleftarrow{s}(n,m)$, then, we obtain,

$$\begin{cases} u(n,m) = u(n-1,m-1) + s(n-1,m-1) \\ s(n,m) = 2u(n-1,m) + s(n-1,m) \end{cases}$$

6. Now, solve the recursive equations using 2-variable generating functions. From the first equation, we get

$$s(n,m) = u(n+1,m+1) - u(n,m)$$

substitute it into the second equation, we get,

$$u(n+1,m+1) - u(n,m) = 2u(n-1,m) + u(n,m+1) - u(n-1,m)$$

change $n-1$ to n, the original n becomes $n+1$, then simplify it and arrange it,

$$u(n+2,m+1) - u(n+1,m) - u(n,m) - u(n+1,m+1) = 0$$

7. Use 2-variable GF to find $u(n,m)$, then find $s(n,m)$, then use the $g(n,m) = u(n,m) + s(n,m)$. Firstly, multiplying by $t^m v^n$ on both side of the equation, then take sum for m, n from 0 to ∞, to simplify the notation, in the next calculation, use '$\sum$' represent $\sum\limits_{m\geq 0, n\geq 0}^{\infty}$, then we get

$$\frac{\sum u(n+2, m+1)t^{m+1}v^{n+2}}{tv^2} - \frac{\sum u(n+1, m)t^m v^{n+1}}{v}$$
$$- \sum u(n, m)t^m v^n - \frac{\sum u(n+1, m+1)t^{m+1}v^{n+1}}{tv} = 0 \quad (4.1)$$

the trick here is tocalculate

$$\sum u(n+2, m+1)t^{m+1}v^{n+2}, \ \sum u(n+1, m)t^m v^{n+1}, \ \sum u(n+1, m+1)t^{m+1}v^{n+1}.$$
The calculation of the three sums is in the following calculation.

Note: to avoid subtract repeated terms, we take different sum limit for n and m as follows.

$$\sum u(n+2, m+1)t^{m+1}v^{n+2} = \mathcal{U}(t,v) - \text{ terms where the exponent of 't' is '0'}$$
$$\text{or exponent of 'v' is '0' or '1'}$$
$$= \mathcal{U}(t,v) - \sum_{n\geq 2} u(n,0)t^0 v^n$$
$$- \sum_{m\geq 0} u(0,m)t^m v^0 - \sum_{m\geq 0} u(1,m)t^m v^1$$
$$= \mathcal{U}(t,v) - 0 - u(0,0)t^0 v^0 - u(1,1)t^1 v^1$$
$$= \mathcal{U}(t,v) - 1 - tv$$

Remark: in the above calculation, when calculate the sum $\sum_{n\geq 2} u(n,0)t^0 v^n$, we first notice that the coefficient

$u(n,0) = $ the number of n-step USSAWS ending with 'Up' step and the height '0'

As all the USSAWS start from the original point $(0,0)$, when the walks ending with Up step has length $n \geq 1$, as the downward step is not allowed, the walk must have minimum height 1, that is, the walks ending with Up step has length $n \geq 1$, will never has height 0, therefore, $u(n,0) = 0$, $\forall n \geq 1$.

Therefore, when you take the sum, $\sum_{n\geq 2} u(n,0)t^0 v^n = 0 \cdot t^0 v^n = 0$

Secondly, when we calculate the sum $\sum_{m\geq 0} u(0,m)t^m v^0$, notice that the coefficient

$u(0,m) = $ the number of 0-step USSAWS ending with 'Up' step and the height 'm'

, when the walk starting from (0,0) and of length 0, the walk can only end with the height '0', that is, $u(0,m) = 0$, $\forall m \geq 1$. By convention, $u(0,0) = 1$.

therefore, when you take the sum, $\sum_{m\geq 0} u(0,m)t^m v^0 = u(0,0) \cdot t^0 v^0 = 1$

Lastly, when we calculate the sum $\sum_{m\geq 0} u(1,m)t^m v^1$, notice that the coefficient

$u(1,m)$ = the number of 1-step USSAWS ending with 'Up' step and the height 'm'

, when one walk starting from $(0,0)$ and has length '1', to make sure it is ending with Up step, the walk can only take the height $m = 1$, that is, $u(1,m) = 0$, $\forall m \neq 1$,

therefore, when you take the sum,

$$\sum_{m\geq 0} u(1,m)t^m v^1 = u(1,1) \cdot t^1 v^1 = tv$$

Using the calculation tricks in the above remarks and avoid repeatable term in the subtraction, we obtain that,

$$\sum u(n+1,m)t^m v^{n+1} = \mathcal{U}(t,v) - \text{ terms where the exponent of 'v' is '0'}$$
$$= \mathcal{U}(t,v) - \sum_{m\geq 0} u(0,m)t^m v^0$$
$$= \mathcal{U}(t,v) - u(0,0)t^0 v^0$$
$$= \mathcal{U}(t,v) - 1$$

$$\sum u(n+1,m+1)t^{m+1} v^{n+1} = \mathcal{U}(t,v) - \text{ terms where the exponent of 't' is '0'}$$
$$\text{or the exponent of 'v' is '0'}$$
$$= \mathcal{U}(t,v) - \sum_{n\geq 1} u(n,0)t^0 v^n$$
$$- \sum_{m\geq 0} u(0,m)t^m v^0$$
$$= \mathcal{U}(t,v) - 0 - u(0,0)t^0 v^0$$
$$= \mathcal{U}(t,v) - 1$$

Therefore, the above equation becomes

$$\frac{\mathcal{U}(t,v) - 1 - tv}{tv^2} - \frac{\mathcal{U}(t,v) - 1}{v} - \mathcal{U}(t,v) - \frac{\mathcal{U}(t,v) - 1}{tv} = 0$$

After arrangement, the equation becomes

$$(1 - tv - v - tv^2) \cdot \mathcal{U}(t,v) = 1 + tv - tv - v$$

Then, we get

$$\mathcal{U}(t,v) = \frac{1 - v}{1 - tv - v - tv^2}$$

8. Now, we have obtained $\mathcal{U}(t,v)$, then use 2-variable GF and the equation

$$s(n,m) \;=\; u(n+1, m+1) - u(n,m)$$

to get the generating function $\mathcal{S}(t,v)$. Multiplying by $t^m v^n$ and summing from $n \geq 0$, $m \geq 0$, same as above in step 7, to to simplify the notation, in the next calculation, use '$\sum$' represent $\sum\limits_{m \geq 0, n \geq 0}^{\infty}$, then we get,

$$\sum s(n,m) t^m v^n = \frac{\sum u(n+1, m+1) t^{m+1} v^{n+1}}{tv} - \sum u(n,m) t^m v^n$$

Then, we get

$$\mathcal{S}(t,v) + \mathcal{U}(t,v) = \frac{\mathcal{U}(t,v) - \sum_{n \geq 1} u(n,0) t^0 v^n - \sum_{m \geq 0} u(0,m) t^m v^0}{tv}$$

From the remark in the step 7, we have known that,

$$\sum_{n \geq 1} u(n,0) t^0 v^n = 0 \cdot t^0 v^n = 0 \quad and \quad \sum_{m \geq 0} u(0,m) t^m v^0 = u(0,0) \cdot t^0 v^0 = 1$$

therefore,

$$\mathcal{S}(t,v) + \mathcal{U}(t,v) = \frac{\mathcal{U}(t,v) - 1}{tv}$$

After arrangement, it can be obtained that

$$
\begin{aligned}
\mathcal{S}(t,v) &= \left(\frac{1}{tv} - 1 \right) \cdot \mathcal{U}(t,v) - \frac{1}{tv} \\
&= \left(\frac{1}{tv} - 1 \right) \cdot \frac{1-v}{1 - v - tv - tv^2} - \frac{1}{tv} \\
&= \frac{1}{tv} \cdot \frac{(1-tv)(1-v) - (1 - v - tv - tv^2)}{1 - v - tv - tv^2} \\
&= \frac{1}{tv} \cdot \frac{2tv^2}{1 - v - tv - tv^2} \\
&= \frac{2v}{1 - v - tv - tv^2}
\end{aligned}
$$

Thus, we get our aim generating function for count sequence $\{g(n,m)\}_{n \geq 0, m \geq 0}$ for the number of $N-$ step Self-avoiding Walks of fixed height $'m'$, and starting from the original point $(0,0)$, which has directions {up, right, left}. That is,

$$
\begin{aligned}
\mathcal{G}(t,v) &= \mathcal{U}(t,v) + \mathcal{S}(t,v) \\
&= \frac{1+v}{1 - v - tv - tv^2}
\end{aligned}
$$

9. Finally, we would like to expand the coefficient of the term with fixed height 'm' and of length 'n' in the expansion of 2-variable generating function $\mathcal{G}(t, v)$. The trick here is in the expansion first fix the exponent of 't' as 'm', thus, reduce a summation, then set the exponent of 'v' as 'n'. That is,

$$
\begin{aligned}
\frac{1+v}{1-v-tv-tv^2} &= (1+v) \cdot \frac{1}{1-v-tv-tv^2} \\
&= (1+v) \cdot \sum_{i \geq 0}^{\infty} (v+tv+tv^2)^i \\
&= (1+v) \cdot \sum_{i \geq 0}^{\infty} v^i \cdot (1+t(1+v))^i \\
&= (1+v) \cdot \sum_{i \geq 0}^{\infty} v^i \cdot \sum_{j \geq 0}^{\infty} \binom{i}{j} t^j (1+v)^j
\end{aligned}
$$

We would like the path ending with a fixed height 'm', so, we only interested the term 't^m', when the internal summation is reduced as one term when 'j=m', then the expansion becomes,

$$
(1+v) \cdot \sum_{i \geq 0}^{\infty} v^i \cdot \binom{i}{m} t^m (1+v)^m = t^m (1+v)^{m+1} \sum_{i \geq 0}^{\infty} \binom{i}{m} v^i
$$

Then, we are interested terms where the exponent of 'v' is 'n', we need to expand $(1+v)^{m+1}$ on the RHS, we get expansion,

$$
t^m \cdot \sum_{j \geq 0}^{m+1} \binom{m+1}{j} v^j \cdot \sum_{i \geq 0}^{\infty} \binom{i}{m} v^i
$$

when 'j' is fixed, to make the exponent of 'v' is equal to 'n', we must set 'i=n-j', thus, we reduce the internal summation over 'i' by fixing 'i=n-j' , the expansion becomes,

$$
t^m v^n \cdot \sum_{j \geq 0}^{m+1} \binom{m+1}{j} \cdot \binom{n-j}{m}
$$

therefore, we get the coefficient of term $t^m v^n$, that is $g(n, m) = \sum_{j \geq 0}^{m+1} \binom{m+1}{j} \cdot \binom{n-j}{m}$

Remark: from the generating function $\mathcal{G}(t, v)$, when we take $t = 1$, we will get the generating function for the count sequence that he number of $N-$ step Self-avoiding Walks with free(non-fixed) height, and starting from the original point $(0, 0)$, which has directions {up, right, left} on square lattice, that is $\mathcal{G}(1, v)$.

4.4   Count unrestricted lattice paths

In this section, a question of counting the number of unrestricted lattice paths advised by supervisor is solved. The question can be related to the counting binary sequences of fixed length and fixed number of bits 1, which was solved using symbolic method by Sedgewick, Robert, and Philippe Flajolet in the book [14, Chapter 3.8]. Counting number of bitstrings of fixed length using symbolic method is also presented on page 243 in the book [14, Chapter 5.4].

In the section, we solve the counting of unrestricted lattice paths by firstly constructing a recurrence equation, then applying the method of solving recurrence equations using generating function method which is taken from [22, Lec-31].

Problem : Calculate the number of paths from (0,0) to (n,m). Denote the number of paths from (0,0) to (n,m) by $a_{n,m}$, Assuming each step can go right or up direction. Solution: First, try to obtain the recurrence relation by reducing the problem, that is,

$$a_{n,m} = (\text{move to right}) \cdot a_{n-1,m} + (\text{move to up}) \cdot a_{n,m-1}, \ \forall n \geq 1, m \geq 1 \quad (4.2)$$

Secondly, define the generating function $f_n(x) = \sum_{m=0}^{\infty} a_{n,m} \cdot x^m$, as $\forall n, m > 0, a_{0,m} = a_{n,0} = 1$, we also define $a_{0,0} = 1$, then we have

$$f_0(x) = \sum_{m=0}^{\infty} a_{0,m} \cdot x^m = \sum_{m=0}^{\infty} x^m = 1 + x + x^2 + \ldots = \frac{1}{1-x}$$

(Here, we does not care the convergence of the sequence because our interest is the coefficients of the sequence.)
by the recurrence equation (1), we calculate

$$
\begin{aligned}
a_{n,1} &= a_{n-1,1} + a_{n,0} \\
&= a_{n-2,1} + a_{n-1,0} + a_{n,0} \\
&= a_{n-3,1} + a_{n-2,0} + a_{n-1,0} + a_{n,0} \\
&= a_{n-n,1} + a_{n-(n-1),0} + a_{n-(n-2),0} + \cdots + a_{n-0,0} \\
&= a_{0,1} + n \cdot 1 = 1 + n
\end{aligned}
$$

similarly using the recurrence equation (1), we can calculate

$$a_{1,m} = 1 + m$$

Thirdly, using the generating function defined as above to calculate the sequence $\{a_{n,m}\}$:

$$
\begin{aligned}
a_{n,m} &= a_{n-1,m} + a_{n,m-1} \\
\implies \sum_{m=1}^{\infty} a_{n,m} \cdot x^m &= \sum_{m=1}^{\infty} a_{n-1,m} \cdot x^m + \sum_{m=1}^{\infty} a_{n,m-1} \cdot x^m \\
\implies f_n(x) - a_{n,0} &= f_{n-1}(x) - a_{n-1,0} + f_n(x) \cdot x \\
f_n(x) &= f_{n-1}(x) + f_n(x) \cdot x \\
f_n(x) &= (1-x)^{-1} \cdot f_{n-1}(x)
\end{aligned}
$$

By induction,

$$
f_n(x) = (1-x)^{-n} \cdot f_0(x) = (1-x)^{-(n+1)} \ \left(by \ \ f_0(x) = \frac{1}{1-x}\right)
$$

Lastly, we got $f_n(x) =^{def} \sum_{m=0}^{\infty} a_{n,m} \cdot x^m = (1-x)^{-(n+1)}$, and then looking for the coefficient of x^m :

in the process of finding the coefficient, we need a general fact about binomial coefficient formula

$$
\begin{aligned}
\binom{-k}{n} &= \frac{(-k)(-k-1)(-k-2)\dots(-k-n+1)}{n!} \\
&= \frac{(-1)^n \cdot (k+n-1)(k+n-2)\dots k}{n!} \\
&= \frac{(-1)^n \cdot (k+n-1)!}{n! \cdot (k-1)!} \\
&= (-1)^n \cdot \binom{k+n-1}{n}
\end{aligned}
$$

By $(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} \cdot x^k, \quad \forall \, \alpha \in \mathbb{R}$,

$$
\begin{aligned}
f_n(x) \quad &= \quad (1-x)^{-(n+1)} = \sum_{m=0}^{\infty} \binom{-(n+1)}{m}(-x)^m \\
&=^{\text{general fact}} \sum_{m=0}^{\infty} (-1)^m \cdot \binom{n+1+m-1}{m} \cdot (-1)^m \cdot x^m \\
&= \quad \sum_{m=0}^{\infty} \binom{n+m}{m} \cdot x^m
\end{aligned}
$$

It can be verified that for $n = 0, \quad f_0(x) = \sum_{m=0}^{\infty} x^m$.
Therefore, for a fixed $n \geq 0$, it can be found that the sequence

$$
a_{n,m} = \binom{n+m}{m}
$$

4.5   Counting the number of path with forbidden city

In this section, one example of counting number of paths from one point to another point with given forbidden line segment is solved using combinatorial methods.

Method 1 is to use the pascal identity to construct the pascal triangle number and solve the counting problem. The pascal triangle identity is a well-known knowledge and can be referred to [37]. The Method 2, we derived the solution to counting a path not touching the given bold line segment and the technique is illustrated in the example.
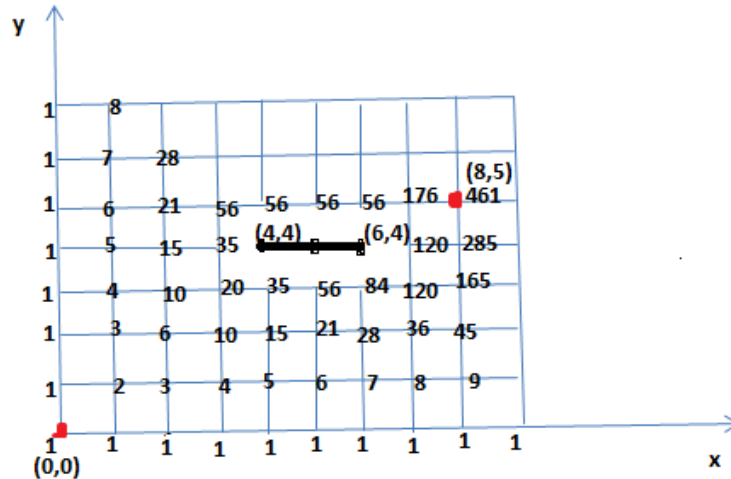

Figure 4.1: Pascal counting

Question: Assuming the step choice is {up, right}. Count the number of paths from one point (0,0) to (8,5) not touching the given segment(bold line segment from (4,4) to (6,4)).

Answer: Method1
in the above graph, use Pascal's triangle to get the number of paths from (0,0) to (8,5) not touching the given bold line segment is 461.
As observed, we count every points except for bold line segment. Starting from the given boundary value 1, there are only one path from (0,0) to each of points on x and y axis. Then, use Pascal's triangle, to get the solution as the above graph. In the hole box of the graph, the top line segment and the right line segment have several same number 56 and 120, because use Pascal's triangle, in these number's down step and left step take number 0 respectively.

Answer: Method2

This method is combinatorial method, we use complement method. Firstly, we would calculate the number of paths touching the line segment. Then, subtract ♯ of paths $(0,0) \to (8,5)$ by ♯ of paths $(0,0) \to (8,5)$ and touching the given bold line segment.

Here, there is a tricky. when do the subtraction, if we consider touching each point in the given bold line segment separately, it is worried that we may subtract more because the set of paths touching each point on the line segment is not disjoint set. Question: How to make the set of paths touching each point on segment is disjoint. The trick used is to consider first touching each point on the given bold line segment. Look at the following graph.
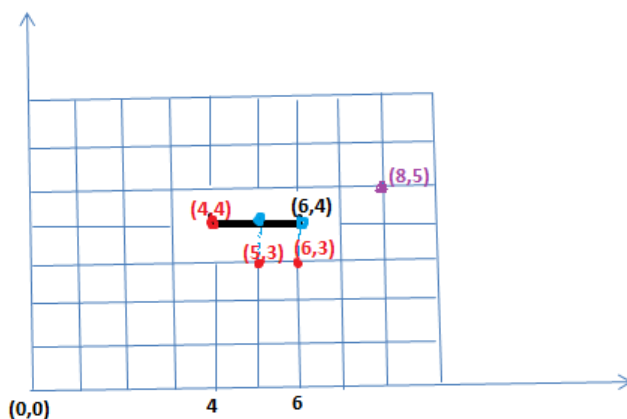
Figure 4.2: avoiding path

We have known from previous section that,

suppose employing two steps $\{up, right\}$, the number of paths from one point (a,b) to another point (x,y) is equal to $\binom{(y-b)+(x-a)}{x-a}$.

If firstly touch the point (4,4), the number of path is [♯ $(0,0) \to (4,4)$ together with $(4,4) \to (8,5)$], that is,

$$\binom{8}{4} \cdot \binom{4+1}{1} = 350$$

If firstly touch the point (5,4), the number of path is [♯ $(0,0) \to (5,3)$ together with $(5,4) \to (8,5)$], that is,

$$\binom{8}{3} \cdot \binom{3+1}{1} = 224$$

If firstly touch the point (6,4), the number of path is [♯ (0,0) → (6,3) together with (6,4) → (8,5)], that is,

$$\binom{9}{3} \cdot \binom{2+1}{1} = 252$$

The total number of paths from (0,0) to (8,5) is

$$\binom{13}{5} = 1287$$

Therefore, the number of paths avoiding the given bold line segment in the graph is equal to

$$1287 - 350 - 224 - 252 = 461$$

It is obvious that the two method get the same counts result.

4.6   Path calculation in gambler's ruin

In this section, we solve the gambler ruin problem using path calculation method. Gambler's ruin is a well-known problem and can be referred to [38] and [39, Chapter 2].

Gambler's ruin; Markov chain, X_n with state space $S = 0, 1, \ldots N$ and transition probability

$$P(0 \to 0) = 1, P(N \to N) = 1 (\text{stay})$$
$$P(k \to k + 1) = p, \ 0 \le k < N (\text{jump up})$$
$$P(k \to k - 1) = 1 - p, \ 0 < k \le N (\text{jump down})$$

Let

$$a_k = a(k) = P(X_\infty = N | X_0 = k)$$

be a probability of winning. It is noted that a_k satifies equation $a_k = a_{k+1}p + a_{k-1}q, \ 0 < k < N$ derived by conditioning on the first step, for $p = q = 1/2$, $a_k = \frac{k}{N}$

We apply the path calculation approach to derive new path calculations.
Path calculation formula is

$$P(\text{winning}) = \sum_{\sigma \in \Omega} P(\text{winning}|\sigma) \cdot P(\sigma)$$

$$\text{where} \sum_{\sigma \in \Omega} P(\sigma) = 1$$

Let Ω be a class of all paths,
Let Ω be a class of all paths stopped at time n, that is

$$X_{n-1} \ne X_n = X_{n+1} = X_{n+2} \quad \text{Stay after time n}$$

n is a stopping time

Path calculation formula is

$$P(\text{winning}) = \sum_{\sigma \in \Omega} P(\text{winning}|\sigma) \cdot P(\sigma)$$

$$\text{where} \sum_{\sigma \in \Omega} P(\sigma) = 1$$

The classical gambler's ruin $P(\sigma \in \Omega \text{ and winning}) = 0$
hence,

$$P(\text{winning}) = \sum_{n} \sum_{\sigma \in \Omega} P(\text{winning}|\sigma_n) P(\sigma_n)$$

Case study , We take for simplicity $N = 3, k = 1$
$N = 3, k = 1$

Notice that for $n = 2m + 2$,

$$P(\text{winning}|\sigma_{2m+2}) = 1$$

provided σ_{2m+2} is finished on 3 (σ_{2m+2} stopped at time $n = 2m+2$) and also

$$\text{Prob(such} \quad \sigma_{2m+2}) = \frac{1}{2^{2m+2}}$$

To win from $k = 1$ we need paths (up,up), (up,down, up,up),(up,down, up,down,up,up),
overall m times (up, down) plus (up,up,...)

Hence,

$$
\begin{aligned}
P(\text{winning}) &= \sum_{n} \sum_{\sigma \in \Omega_n} P(\text{winning}|\sigma_n) P(\sigma_n) \\
&= \sum_{n=2m+2} P(\text{winning}|\sigma_{2m+2}) P(\sigma_{2m+2}) \\
&= \sum_{m=0} \frac{1}{2^{2m+2}} \\
&= \frac{1}{3} = \frac{k}{N}
\end{aligned}
$$

only one path stopped at even time if starting from $k = 1$, also only one path stopped at odd time if starting from $k = 2$.

By introducing different probability $P(\sigma)$, we derive the different gambler's ruin schemes.

Example:

$$P(\sigma_4) = a, \; P(\sigma_3) = 1 - a, P(\text{any other path}) = 0$$

Notice:
σ_4 up, down, up,up, stay, stay, ...at (winning)
σ_3 up, down, down,stay, stay, ...at (losing)

$$
\begin{aligned}
P(\text{winning}) &= \sum_n \sum_{\sigma \in \Omega_n} P(\text{winning}|\sigma_n)P(\sigma_n) \\
&= \sum_{n=2m+2} P(\text{winning}|\sigma_{2m+2})P(\sigma_{2m+2}) \\
&= \sum_{m=1} P(\sigma_4) \\
&= a
\end{aligned}
$$

Chapter 5

Modelling share price paths via Binary trees

The main idea is to construct the path as a realization of a certain binary tree.

The possibility comes from the Bijection between the set of binary trees and the set of Dyck paths motivated by [17] or the bijection between the general rooted ordered trees and Dyck paths motivated by [29, Page 4]. Motivated from the algorithm of constructing share price path from a set of full binary trees we derived in the section 5.1.1 motivated by [17], we derive a construction of share price paths from a set of general binary trees in the section 5.3.1.

Suppose a share price value is x on a path, each jump up is associated with move $x \to xu$ and each jump down with move $x \to xd$.

By constructing a random binary tree, we are able to give another construction of the random share price path.

5.1 A bijection between the set of binary trees and the set of Dyck paths

The definition 54 is motivated from the page 66 in the book [15, Chapter I] by Sedgewick, Robert, and Philippe Flajolet and by Flajolet in the book [14, chapter 3.8].

Definition 54 *(full binary tree)*
A binary tree is a node (called root) and a sequence of 0 or 2 binary trees; each node has 0 or 2 children. If a node has two children, then the node is called internal nodes(vertices), otherwise, it is called external nodes or leaves. The binary tree can also be called a full binary tree.

The definition 55 is motivated from the page 12 in the thesis [23].

Definition 55 *(Dyck path)*
A Dyck path is a sequence of up and down steps, the size of each step has equal

length, it begins and ends on the same level, and never fall below the height it began on.

The theorem 56 is motivated by the method stated by Federico in the enumerative combinatorics homework exercise 3.3 [17].

Theorem 56 *Let $\mathcal{P}$ be the set of Dyck paths which starts from $(0,0)$, going up or down in each step one unit length, never going below the zero level, the path ends at zero horizontal level.*

Let $\mathcal{T}$ be the set of binary trees.

Let f be a function from the set $\mathcal{P}$ to the set $\mathcal{T}$

Then, there is a one-to-one corresponcdence between the set of binary trees with $n+1$ nodes and the set of dyck paths of length n.

5.1.1 Bijection algorithm from binary trees

In this section, the construction algorithm from binary trees to share price paths is introduced, and the method is motivated and developed from the exercise 3.3 by Federico in the enumerative combinatorics homework [17].

Let P be a Dyck path of length $n = 2m$ from the set $\mathcal{P}$, Let T be a binary tree with $n+1$ nodes from the set $\mathcal{T}$, then, the construction between the bijection is as follows,

Firstly, construct T from P.

Step 1: starting from $S_0 = (0,0)$, draw the root of $f(P)$,

Step 2: Going through from $i = 1$ to $i = n$, each time when going up in the path P, that is, $(t_i, S_i) \rightarrow (t_{i+1}, S_{i+1})$, with $S_{i+1} = S_i + 1$,

then, draw left son from the last vertex which was drawn in $f(P)$.

Step3: Going through from $i = 1$ to $i = n$, each time when going down in the path P, that is, $(t_i, S_i) \rightarrow (t_{i+1}, S_{i+1})$, with $S_{i+1} = S_i - 1$,

then, go up one or several vertices in $f(P)$, until you can draw the right son in $f(P)$.

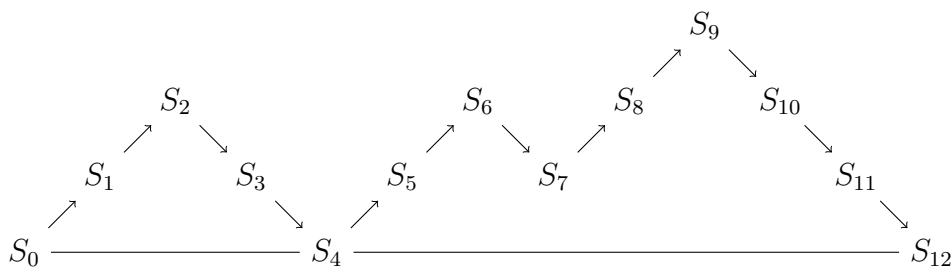Say, if P is a Dyck path of length 12, and given in the Figure 5.1.

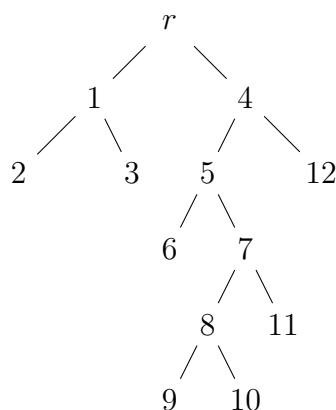Figure 5.1: a dyck path of length   12

It corresponds to a binary tree

Figure 5.2: the corresponding binary tree

Conversly, construct P from T.

Step 1: starting from the root of T, draw S_0 at position $(t_0, S_0) = (0,0)$ in $f^{-1}(T)$, It is noted that the tree T can be labelled using depth-search or has no labels.

Step 2: starting from the current position, if you can go down left to a new vertex in the tree T,

then, draw a path $(t_i, S_i) \to (t_{i+1}, S_{i+1})$, with $S_{i+1} = S_i + 1, t_{i+1} = t_i + 1$ where $0 \leq i \leq n - 1$.

It is noted that the first vertex must be in the left branch of the tree.

Step3: starting from the current position, if you cannot go down left to a new vertex in the tree T,

then, go up one or several vertices in T, until you can go down right to new vertex where you have not visisted before in the tree, and then draw a path $(t_i, S_i) \to (t_{i+1}, S_{i+1})$, with $S_{i+1} = S_i - 1, t_{i+1} = t_i + 1$ where $0 \leq i \leq n - 1$.

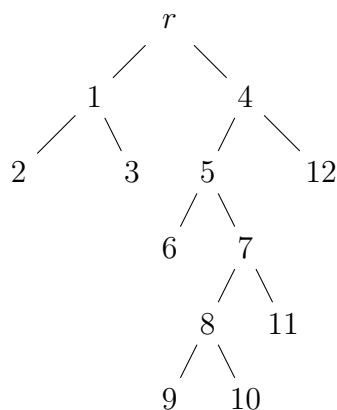For example, if T is a tree of vertices 13, and given in the Figure 5.3.

Figure 5.3: a tree of vertices 13

It corresponds to a share path

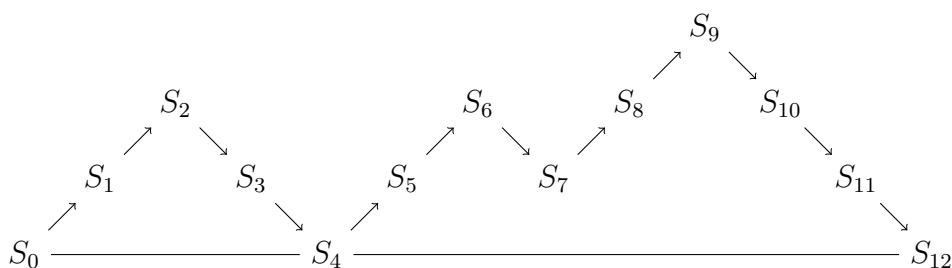Figure 5.4: the corresponding share path

Remarks [17]:

For any Dyck path $P \in \mathcal{P}$, if you go up $a(P)$ steps before going down, you draw $a(P)$ left sons in the tree $f(P)$.

Let $a(P)$ be the number of up steps before you first time going down in the Dyck path P, the next vertex in $f(P)$ is either going down left or going up until you can going down to a new vertex in $f(P)$. It is obvious that you will not go back to the left most branches in $f(P)$, $a(P)$ is the length of the left most branch in $f(P)$.

5.1.2 Share price path interpretation

In the above example of the construction of share path P from the binary tree T, assume that a jump up is a move $S_k \to S_{k+1} = uS_k$, and a jump down is a move $S_k \to S_{k+1} = dS_k$, assume the initial share value S_0 is fixed, then, the share path can be represented by

$$(1, u, u^2, u^2d, (ud)^2, (ud)^2d, (ud)^2d^2, (ud)^2d^3, (ud)^3d^2, (ud)^4d, (ud)^4d^2, (ud)^4d^3, (ud^2)^4)$$

Suppose the defined probability measure is $p(\text{up}) = p_u$ and $p(\text{down}) = p_d$ the corresponding step probability sequence is

$$(1, p_u, p_u, p_d, p_d, p_d, p_d, p_d, p_u, p_u, p_d, p_d, p_d)$$

5.2   A bijection between the set of general rooted ordered trees and the set of Dyck paths

In this section, another bijection between trees and share path is constructed.

The definition 57 is motivated from the page 66 in the book [15, Chapter I] and in the book [14, chapter 6.2] by Flajolet and Sedgewick.

Definition 57 *(rooted ordered trees)*
An rooted ordered tree is a node (called root) and a sequence of ordered rooted trees; each node can have non-negative number of children. In the ordered tree, if a node has more than one child, the order of the children trees is significant.

Consider a set of rooted ordered trees, if the tree is a combinatorial class and we count trees by the number of tree vertices, then, using the symbolic method introduced in later chapter 6 and definition of rooted ordered trees, we construct the tree as

$$\mathcal{G} = \text{a node} \times SEQ(\mathcal{G})$$

Using the transfer theorem of symbolic method, we get the generating function equation

$$G(z) = z(1 + G(z) + G(z)^2 + \ldots) = \frac{z}{1 - G(z)}$$

It is simplified as

$$G(z) = z + G(z)^2$$

Choose the solution

$$G(z) = \frac{1 - \sqrt{1 - 4z}}{2}$$

Extract the coefficient G_N which counts the number of ordered trees of N vertices,

$$G_N = [z^N]\frac{1 - \sqrt{1 - 4z}}{2} = \frac{1}{N}\binom{2(N-1)}{N-1}$$

Recall the Dyck path construction which is introduced in Chapter 7 in detail.

The construction of the set $\mathcal{D}$ of Dyck path can be written as

$$\mathcal{D} = \text{empty set} + \text{Up} \times \mathcal{D} \times \text{Down} \times \mathcal{D}$$

Count the set of Dyck paths by the length of dyck paths, using the transfer theorem, the generating function equation for the dyck paths is

$$D(z) = 1 + z^2 D(z)^2$$

Choose the solution

$$D(z) = \frac{1 - \sqrt{1 - 4z^2}}{2z^2}$$

Extract the coefficient D_{2N} which counts the number of dyck paths of length $2N$,

$$D_{2N} = [z^N]D(z) = [z^{2N+2}]\frac{1 - \sqrt{1 - 4z^2}}{2} = \frac{1}{N+1}\binom{2(N)}{N}$$

It can be seen that

$$G_{N+1} = D_{2N}$$

and it can be convinced that there might exist a bijection between the set of Dyck paths of length $2N$ and the set of ordered trees of $N + 1$ vertices.

The above discussion between the relationship and calculation of dyck paths and general rooted trees can refer to the section 3.4.2 and section 4.2 in the thesis motivated from the book [14]and the thesis [23, Example 2.7].

Next, we construct a bijection between the two sets.

The theorem 58 is motivated by the method stated by Dershowitz and Rinderknecht in the paper [29, Page 4].

Theorem 58 *Let* $\mathcal{P}$ *be the set of Dyck paths which start from* $(0,0)$, *going up or down in each step one unit length, never going below the zero level, the path ends at zero horizontal level.*

Let $\mathcal{G}$ *be the set of rooted ordered trees.*

Let f *be a function from the set* $\mathcal{P}$ *to the set* $\mathcal{G}$

Then, there is a one-to-one correspondence between the set of rooted ordered trees with $n + 1$ *nodes and the set of Dyck paths of length* $2n$.

5.2.1 Bijection algorithm from rooted ordered tree

In the section, using the share price has the same shape as the example motivated by Federico in the enumerative combinatorics homework [17], the construction algorithm between general trees and share price paths is developed and the method is motivated by Dershowitz and Rinderknecht in the paper [29, Page 4].

Let P be a Dyck path of length $2n$ from the set $\mathcal{P}$, Let G be a rooted ordered tree with n edges and $n + 1$ nodes from the set $\mathcal{G}$, then, the construction of the bijection is as follows,

Firstly, construct G from P.

Step 1: starting from $S_0 = (0,0)$, draw the root of $f(P)$,

Step 2: Going through from $i = 1$ to $i = n$, each time when going up in the path P, that is, $(t_i, S_i) \to (t_{i+1}, S_{i+1})$, with $S_{i+1} = S_i + 1$,

then, in the current traversal position in $f(P)$, draw a new left or right child node in the tree $f(P)$; that is, for the current position node in the tree, if there is no left son, then, draw a new left son. Otherwise, draw a new right son in the tree.

Step3: Going through from $i = 1$ to $i = n$, each time when going down in the path P, that is, $(t_i, S_i) \to (t_{i+1}, S_{i+1})$, with $S_{i+1} = S_i - 1$,

then, go up one vertex in $f(P)$, the traversal node is updated to the lower level vertex which is the parent of the prior position node; that is, no new node is drawn in the tree and only the traversal node position is updated.

For example, using the same example in last section, if P is a dyck path of length 12, and given in the Figure 5.5.
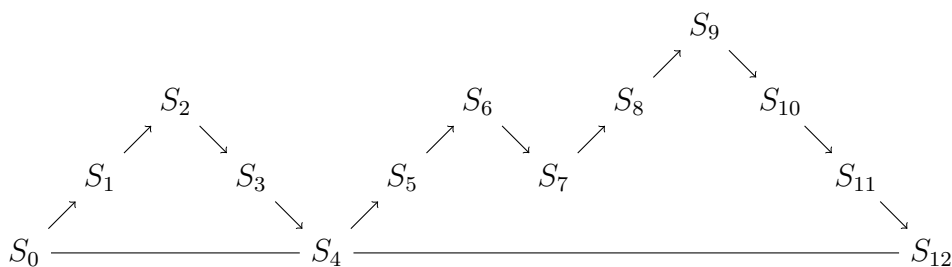

Figure 5.5: a Dyck path of length 12

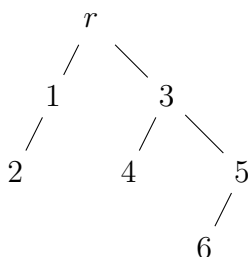It corresponds to an ordered tree of 7 vertices

Figure 5.6: the corresponding ordered tree of 6 edges

Conversly, construct P from T.

Step 1: starting from the root of T, draw S_0 at position $(t_0, S_0) = (0,0)$ in $f^{-1}(T)$, It is noted that the tree T can be labelled using depth-search or has no labels.

Step 2: starting from the current position, check firstly if you can go down left to a new vertex in the tree T, which you have not visited before. If you can do it,

then, draw a path $(t_i, S_i) \rightarrow (t_{i+1}, S_{i+1})$,   with $S_{i+1} = S_i + 1, t_{i+1} = t_i + 1$ where $0 \leq i \leq n - 1$,

and the traversal node position is updated to the child node you just draw in the tree T.

It is noted that the first vertex must be in the left branch of the tree, and the procedure carry on until you traverse a node in the tree which has no left new child.

Step3: starting from the current position, if you cannot go down left to a new vertex in the tree T, then, check if you can go down right to a new vertex in the tree which you have not visited before. If you can do it,

then, draw a path $(t_i, S_i) \rightarrow (t_{i+1}, S_{i+1})$,   with $S_{i+1} = S_i + 1, t_{i+1} = t_i + 1$ where $0 \leq i \leq n - 1$.

If you cannot do it, it means that, from the current position you cannot go down to find new vertex in the tree,

then, go up one vertex in T, update the traversal node position, and draw a path $(t_i, S_i) \rightarrow (t_{i+1}, S_{i+1})$,   with $S_{i+1} = S_i - 1, t_{i+1} = t_i + 1$ where $0 \leq i \leq n - 1$.

Then, based on the updated traversal position, carry on the step 2 and step 3, until you traverse the root node.

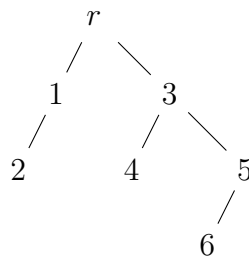For example, if T is a tree of vertices 7, and given in the Figure 5.7.

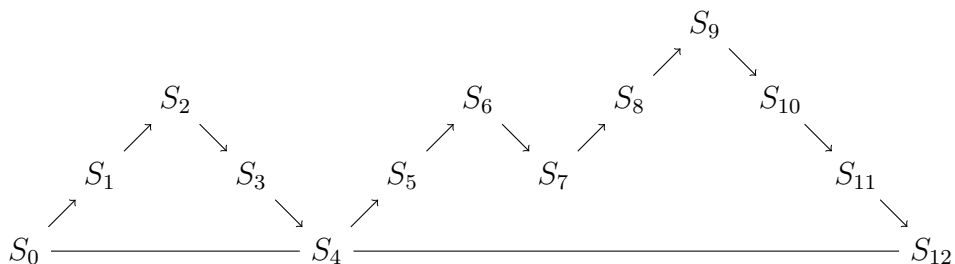Figure 5.7: a tree of vertices 7

It corresponds to a share path

Figure 5.8: the corresponding share path

5.3 An injection construction of share path from general binary rooted ordered trees

In this section, a new construction algorithm is developed and the idea is inspired from the construction algorithm in the Section 5.1.1.

It can be observed that binary trees defined in the Section 5.1 is a special case of ordered trees; each node in a binary tree has non-negative number of children vertices which take values in the set $\{0, 2\}$, the binary tree can also be called full binary tree.

In the bijection of Section 5.1, since the construction of Dyck paths P can be done from a binary tree which is either labelled using depth-search or not labelled, to make the algorithm is valid for the the tree unlabelled, we have to make sure that, for each node, there always has a left branch(child) before having a right branch(child).

Thus, it give us chance to construct a path from a general binary ordered rooted tree, where the number of children vertices takes values in the set $\{0, 1, 2\}$.

The definition 59 is motivated by Flajolet and Sedgewick on the page 66 in the book [15, Chapter I], named unary–binary trees.

Definition 59 *(general binary ordered rooted trees)*
A general ordered binary tree is a node (called root) and a sequence of 0, 1 *or* 2 *ordered binary trees; each node has non-negative number of children which take values in the set* $\{0, 1, 2\}$.

The assumption is that if a node has only one child in the tree, we treat the node's child as its left branch node.

Next, construct an injection from a general binary ordered trees to a share path using the one side of bijection algorithm between the set of binary trees and the set of share paths.

5.3.1 An injection construction algorithm

Let T be a rooted ordered tree in the set $\mathcal{G}$.
Let P be a share path which starts from $(0, 0)$, going up or down in each step one unit length, never going below the zero level.
Then, we construct a path P from a binary ordered tree T as follows,

Step 1: Using the assumption that if a node has only one child in the tree, we treat the node's child as its left branch node, transfer the general binary tree to a binary tree where each has left branch or right branch.

Step 2: starting from the root of T, draw S_0 at position $(t_0, S_0) = (0,0)$ in $f^{-1}(T)$, It is noted that the tree T can be labelled using depth-search or has no labels.

Step 3: starting from the current traversal position, if you can go down left to a new vertex in the tree T,

then, draw a path $(t_i, S_i) \rightarrow (t_{i+1}, S_{i+1})$, with $S_{i+1} = S_i + 1, t_{i+1} = t_i + 1$ where $0 \leq i \leq n - 1$.

It is noted that the first vertex must be in the left branch of the tree.

Step 4: starting from the current position, if you cannot go down left to a new vertex in the tree T,

then, go up one or several vertices in T, until you can go down right to new vertex where you have not visited before in the tree, and then draw a path $(t_i, S_i) \rightarrow (t_{i+1}, S_{i+1})$, with $S_{i+1} = S_i - 1, t_{i+1} = t_i + 1$ where $0 \leq i \leq n - 1$.

Step5: carry on the step 3 and step 4, until you finish the traversal of the given general binary tree.

Based on the construction algorithm, we developed a simple example to illustrate the construction algorithm of share price path from a given binary tree of vertices 4.

Example.

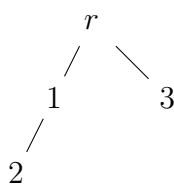Given a general binary tree of vertices 4 as follows,

Figure 5.9: a general ordered binary tree of 3 edges, 4 vertices

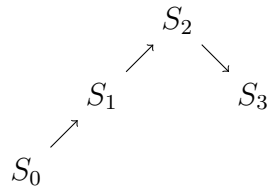Using the injection algorithm, the constructed share path is,

Figure 5.10: the corresponding share path of length 3

using the injection algorithm, it is noted that the constructed path is a kind of share price path which has number of up steps greater than or equal the number of down steps, and the share path never ends at zero horizontal level.

If the general binary tree is a full binary tree, then, the constructed share path has the same number of up steps as the number of down steps, and the constructed share path is a Dyck path.

If we use the bijection algorithm in the Section 3.2, the constructed share path is
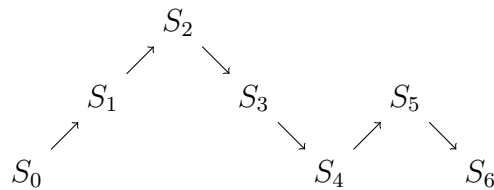

Figure 5.11: the constructed share path from the general binary tree

Remarks: In the above example, from the same given general binary tree, we get two constructed share paths, compare the two constructed share price paths, the injection construction path steps correspond to the up step in the constructed Dyck path. That is,
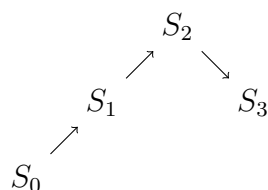
the path

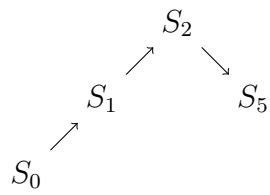Figure 5.12: the corresponding share path of length 3

can be labelled as

$$S_2$$
$$\nearrow \qquad \searrow$$
$$S_1 \qquad\qquad S_5$$
$$\nearrow$$
$$S_0$$

Figure 5.13: the corresponding share path of length 3

Chapter 6

Count all paths not allowing given down steps in Binomial model

In this chapter, motivated by Flajolet symbolic method and generating function approach[14] [15], and motivated by quantum path calculation method in [2] [11], we calculated the option price in the finite restricted binomial model.

6.1   Construction

Motivated by the construction method of words over the binary alphabet letters $\{a, b\}$ which do not have k consecutive a described on the page 51 motivated by by Flajolet and Sedgewick in the book [15] and definitions of sequence symbolic construction in the book [15, Page 25] and [14, Chapter 5.2], we apply the method and develop the construction of the set $\mathcal{A}_k$ of restricted share price path and its path generating function.

Suppose share prices is modelled by a binomial model, and the model use the probability measure defined by $p((up)) = p_u$, $p((down)) = p_d$, with $p_u + p_d = 1$. Then, put a restriction on the share values in the model, the restriction is that not allowing share prices has consecutive downward changes greater and equal than k steps, where k is a given positive integer.

$\mathcal{A}_k$ denotes the set of possible share values of $(S_0 -> S_1 -> S_2 -> \ldots)$ of the restricted binomial model. So, the set $\mathcal{A}_k$ is all the possible share price paths in the restricted binomial model.

Assuming the set $\mathcal{A}_k$ is a combinatorial class, then, construct any share value path in the restricted model using the symbolic method motivated by by Flajolet and Sedgewick in the book [15] and the book [14, Chapter 5.2] as follows,

For any share path σ in the restricted model sample space set $\mathcal{A}_k$, it is a path not allowing k down steps for any length n, and can be defined by a path consisting of consecutive $\leq k$ downward steps either followed by an empty path or followed

by a one Up step connected to a path not allowing k down steps for any length n, that is,

$$
\begin{aligned}
(\mathcal{A}_k) \quad = \quad & (\text{empty path} + \text{path of one down step} \\
& + \text{path of two down steps} + \ldots + \\
& + \text{path of } k \text{ down steps}) \\
& \times (\text{empty path} \\
& \quad + \text{Up} \times (\mathcal{A}_k))
\end{aligned}
$$

Let each valid path in the restricted model $\mathcal{A}_k$ is denoted by σ_k,

count the set of restricted paths by the path lengths,

the generating function for the path σ_k is defined by

$$
\sigma_k(z) = \sum_{i \geq 0}^{\infty} S_i(\sigma) z^{|\sigma_1 \sigma_2 \ldots \sigma_i|}.
$$

Then, the path generating function for the restricted model can be defined by

$$
A_k(z) = \sum_{\sigma \in \mathcal{A}_k} \sigma_k(z) = \sum_{i=0}^{\infty} \sum_{\sigma \in \mathcal{A}_{k,n}} S_i(\sigma) z^i
$$

from the above construction of the restricted paths, using the symbolic method and its transfer theorem introduced in the Chapter 6, the path generating function of the restricted model satisfies the equation

$$
A_k(z) = (1 + z + \ldots + z^{k-1})(1 + z A_k(z))
$$

After extracting the coefficient of the generating function $A(z)$,

$$
[z^N] A_k(z) = \sum_{\sigma \in \mathcal{A}_{k,n}} S_N(\sigma)
$$

it is the sum of the possible state values at time N in the restricted binomial model.

6.2   Calculate option price in the finite restricted binomial model

Motivated by the quantum binomial option pricing in the paper [2, Section 2.3], using the path calculation method which is motivated from the Feynman path integral method or sum-over-history method in [12] [13], we derive the calculation of option

price in the finite restricted binomial model, which can be a generalization of the path calculation lemma in the paper [2] [11].

Consider the restricted n steps binomial model, the path generating function for the finite restricted model is

$$
A(z; k, n) \quad := \quad \sum_{\sigma \in \mathcal{A}_{k,n}} \sum_{i=0}^{n} S_i(\sigma) z^{|\sigma_1 \sigma_2 \dots \sigma_i|}
$$

$$
= \quad \sum_{i=0}^{n} \sum_{\sigma \in \mathcal{A}_{k,n}} S_i(\sigma) z^i
$$

$$
\text{where,} \quad \sigma = \sigma_1 \sigma_2 \dots \sigma_n
$$

$$
S_i(\sigma) = \text{the share value at time } i \text{ of the path}
$$

$$
\sigma \quad \text{in the restricted model } A_{k,n}
$$

In the finite restricted model, let each possible path σ has probability $CP_k(\sigma)$, where $C = \frac{1}{\sum_{\sigma \in \mathcal{A}_{k,n}} P_k(\sigma)}$, which is the probability normalization constant.

The path generating function for the restricted model is

$$
P(w; k, n) \quad := \quad \sum_{\sigma \in \mathcal{A}_k} \sum_{i=0}^{n} CP_k^i(\sigma) w^{|\sigma_1 \sigma_2 \dots \sigma_i|}
$$

$$
\text{where,} \quad CP_k^i(\sigma) = CP_k(\sigma_1 \sigma_2 \dots \sigma_i)
$$

$$
= \quad \text{the state probability at time } i \text{ of the valid path}
$$

$$
\sigma \quad \text{in the restricted model } \mathcal{A}_{k,n}
$$

Thus, in the restricted n steps binomial model $\mathcal{A}_{k,n}$, the expected share price value at time n can be calculated by

$$
E[S_n(\sigma)] = \sum_{\sigma = \sigma_1 \sigma_2 \dots \sigma_n \in \mathcal{A}_{k,n}} S_n(\sigma) CP_k^n(\sigma)
$$

The restricted model $\mathcal{A}_{k,n}$ is a n steps binomial model, suppose the risk-free interest rate is r.

Next is to calculate the price of the exotic option claim in the restricted model.

Suppose the option claim for the share price at time t is $C_t = f(S_t)$, for $1 \le t \le n$.
Let T be the final time, and $1 \le t \le T$.

The exotic option claim in the model is defined by $Claim = \sum_{t=1}^{T} f(S_t)$

Thus, the option value of the exotic claim in the finite restricted model is stated as

Theorem 60 *the option value of the exotic option claim in the finite restricted model* $\mathcal{A}_{k,n}$, *at time* $t = 0$ *is calculated using the given risk-free interest rate* r *and it is*

$$
\begin{aligned}
OP(Claim) \;&=\; OP(\sum_{t=1}^{T} f(S_t)) = \sum_{t=1}^{T} OP(f(S_t)) \\
&=\; \sum_{t=1}^{T} [(1+r)^{-t} \cdot E(f(S_t))] \\
&=\; \sum_{t=1}^{T} [(1+r)^{-t} \cdot \left(\sum_{\sigma \in \mathcal{A}_{k,n}} f(S_t(\sigma)) CP_k^t(\sigma) \right)] \\
&=\; \sum_{\sigma = \sigma_1 \sigma_2 \ldots \sigma_n \in \mathcal{A}_{k,n}} \sum_{t=1}^{T} \frac{f(S_t(\sigma)) CP_k^t(\sigma)}{(1+r)^t}
\end{aligned}
$$

The path calculation method in the theorem generalize the restricted path calculation lemma, which can be compared to the Theorem 7 which is a generalization of path calculation lemma in the paper [2] [11].

In the model $\mathcal{A}_{k,n}$, since the starting state of the share price is fixed, for each path $\sigma \in \mathcal{A}_{k,n}$, set $CP_k^0(\sigma) = 1$. From the path probability generating function $P(w; k, n)$ of the restricted n times binomial model, it can be observed that

$$
\begin{aligned}
P(0; k, n) \;&=\; \sum_{\sigma \in \mathcal{A}_{k,n}} CP_k^0(\sigma) \\
&=\; \sum_{\sigma \in \mathcal{A}_{k,n}} 1 \\
&=\; \sharp \text{ of the possible paths in the restricted model } \mathcal{A}_{k,n}
\end{aligned}
$$

Suppose each path has equal probability in the restricted model $\mathcal{A}_{k,n}$, then, the probability of valid path $\sigma \in \mathcal{A}_{k,n}$ can be calculated by

$$
P(\sigma) = \frac{1}{P(0; k, n)}
$$

Chapter 7

Count paths using Parallelogram polyominoes

7.1 Definitions

The definition 61 is motivated from the handout of Interior, closure, and boundary by Conrad, Brian [33].

Definition 61 *(Interior of a set)*
Let S be a set of points in the Cartesian plane(a two-dimensional Euclidean plane), Then, the interior of S is the set of all points in S which does not belong to the boundary of S, it is denoted by $int(S)$, so, it is the largest open set contained in S and it is the union of all open sets contained in S.

The definition of open ball in the discussion of the next remarks is motivated by Weisstein in [34].

Remarks: A point $x \in S$ if and only if there exists $\epsilon > 0$ such that an open ball centred at x with radius ϵ is completely contained in S. The open ball is denoted by $B(x, \epsilon)$, and defined by $B(x, \epsilon) := \{y \in \mathbb{R}^2 \mid d(x, y) < \epsilon\}$

The definition 62 is motivated by Delest in the introduction of the paper [32].

Definition 62 *(cell)*
A unit square with vertices staying on integer points in $\mathbb{R}^2$ is called a cell.

The definition 63 is motivated by Delest in the introduction of the paper [32].

Definition 63 *(polyomino)*
A polyomino is a finite connected union of cells such that its interior is connected.

Motivated from the discussion of the definition of polyomino by Weisstein [35] and the definition of interior of a set in the definition 61, we develop a small example of not being polyomino as an illustration of understanding the definition.

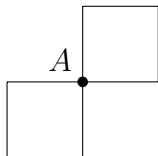**Example 7.1.1** *A finite connected union of cells which is not polyomino is:*

Figure 7.1: an example of not being polyomino

It is observed that the point A is not in the interior of the polyomino. The interior of the polyomino is not connected.

The definition 64 is motivated by Delest and Viennot from page 170 in the paper [30].

Definition 64 *(column or row convex polyomino)*
A polyomino is said to be column- (respectively row-) convex if the intersection of any infinite vertical (respectively horizontal) line with the polyomino is a connected line segment.

To illustrate the definition of convex polyomino, we gave a column-convex polyomino and a row-convex polyomino in the following example.

Example 7.1.2 *a column-convex polyomino is:*
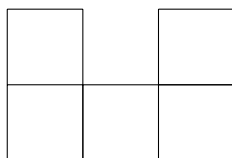

Figure 7.2: an example of column-convex polyomino
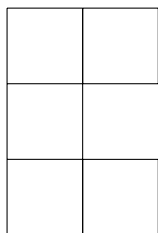
a row-convex polyomino is:

Figure 7.3: an example of row-convex polyomino

The definition 65 is motivated by Delest in the introduction of the paper [32].

Definition 65 *(convex polyomino)*
A polyomino is said to be convex polyomino if it is both column-convex and row-convex polyomino.

We gave a simple example as follows to illustrate the definition of convex polyomino in the definition 65.
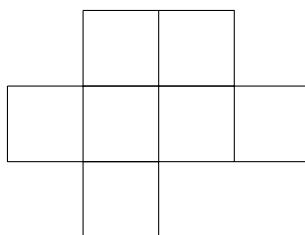
Example 7.1.3 *A convex polyomino is:*

Figure 7.4: an example of convex polyomino

Introducing notation, and give another definition of convex polyomino as follows.

The definition 66 is motivated from the discussion of the definition of polyomino by Weisstein [36] and the page 177 by Viennot from the paper [30].

Definition 66 *(convex polyomino)*

A convex polyomino, denoted by P, is a polyomino whose perimeter is equal to that of its minimal bounding box (Bousquet-Mélou et al. 1999). The perimier of the polyomino is the length of its border. The minimal bounding box of a polyomino is the smallest rectangle that contains P, denoted by $Rect(P)$. The rectangle is also a convex polyomino by the definition.

A small example is given as follows, in which the discussions and notations are motivated by Delest and Viennot in the paper [30, Page 177-178].

Example 7.1.4 *A convex polyomino P with its smallest rectangle $Rect(P)$ is:*
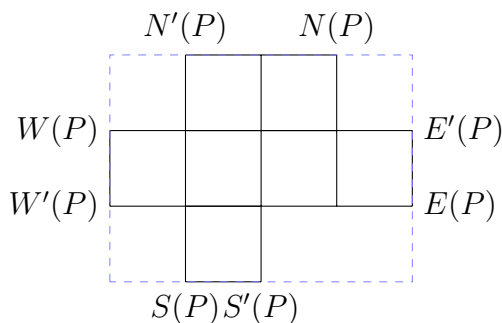
Figure 7.5: an example of convex polyomino with smallest bounding box

It is observed that going through four directions anti-clock-wisely (South, East, North, West), the rectangle $Rect(P)$ touches the convex polyomino P at four connected line segments. Anti-clock-wisely, the extreme points of the four segments are denoted by $[S(P), S'(P)]$, $[E(P), E'(P)]$, $[N(P), N'(P)]$, $[W(P), W'(P)]$.

The definition 67 is motivated by Delest and Viennot in the paper [30, Page 178].

Definition 67 *(Parallelogram polyomino)*

A convex polyomino P is called a parallelogram polyomino if its smallest bounding rectangle $Rect(P)$ touches the polyomino at four connected line segments with the condition that (the South left extreme point = the West bottom extreme point) and (the East top extreme point = the North right extreme point), if label the extreme points of the four segments by $[S(P), S'(P)]$, $[E(P), E'(P)]$, $[N(P), N'(P)]$, $[W(P), W'(P)]$,   the condition is $S(P) = W'(P)$ and $N(P) = E'(P)$. It is observed that the extreme point $N(P)$ is the rightmost top point and $S(P)$ is the leftmost bottom point.

A small example is given by us to illustrate the definition of parallelogram polyominoe.

Example 7.1.5 *A parallelogram polyomino P with its smallest rectangle $Rect(P)$ is:*
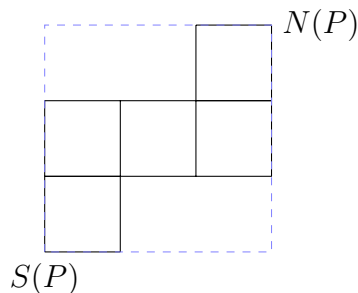
Figure 7.6: an example of parallelogram polyomino

To relate a parallelogram polyomino to a connected lattice path matroid, introduce the next definition of parallelogram polyomino. The definition of lattice path matroid can be referred to the Definition 3.1. in the paper [60]. The definition 68 is motivated from the discussion by Delest and Viennot on the page 178 in the paper [30, Page 178].

Definition 68 *(Parallelogram polyomino)*

A convex polyomino P is called a parallelogram polyomino if its smallest bounding rectangle $Rect(P)$ is characterized by two lattice paths which have the same initial point $S(P)$ and same ending point $N(P)$, each lattice path on the border is formed of East and North unit steps. Denote the top bounding lattice path by w and denote the bottom bounding lattice path by η. The top path w is above the bottom path η, and the two paths do not intersect each other except the initial point $S(P)$ and the ending point $N(P)$, otherwise, it contradicts the definition of being a polyomino.

The same small example as above given by us using the bounding pair paths is illustrated to understand the definition 68.

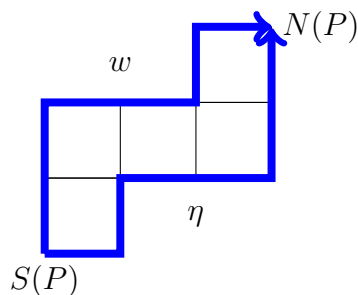**Example 7.1.6** *A parallelogram polyomino P with its bounding pair paths (w, η) is:*

Figure 7.7: the parallelogram polyomino with its bounding paths (w, η)

It is observed that the parallelogram polynomino is characterized by two paths with same inital point $S(P)$, and ending point $N(P)$. The pair (w, η) of paths is as follows,

$$
\begin{aligned}
w &= N_1 \to N_2 \to E_3 \to E_4 \to N_5 \to E_6 \\
\eta &= E_1 \to N_2 \to E_3 \to E_4 \to N_5 \to N_6
\end{aligned}
$$

7.2   Construction of dyck paths from lattice path matroids

7.2.1   Construction

In this section, the construction algorithm from lattice path matroid to a special typle of share price paths, say Dyck paths is introduced and the method is modified by us and motivated from the paper [30, Page 179], in the paper, Delest and Viennot developed a bijection 2-colored Motzkin words of length $n-1$ and Dyck words of length $2n$.

Next, construct a bijection between the set of lattice path matroid and the set of Dyck paths. Using the above example of parallelogram polynomino, the construction is as follows,

In the example, Let $S(P) = (0,0)$, each step has length 1, the polyomino is defined by two paths (w, η), each path from $S(P) = (0,0)$ to $N(P) = (3,3)$ of length $n+1 = 6$, and has perimeter $2(n+1) = 12$. The number of cells in the polyomino is $n = 5$.

Firstly, construct a Dyck path from a pair of paths $[w, \eta]$.

Step 1: Set up delimiter intervals $[\Delta_i, \Delta_{i+1}]$, where Δ_i denotes the line $y = -x+i$, $1 \leq i \leq n-1 = 4$. It is noted that the delimilter interval contains all pairs steps of paths (w, η) except the first pair steps (N_1, E_1) and the last pair steps (E_6, N_6).

Step 2: Going through the delimiliter intervals $[\Delta_i, \Delta_{i+1}]$, for $1 \leq i \leq n-1 = 4$, use the pair steps (w_j, η_j) with $2 \leq j \leq n$ to construct the 2-colored Motzkin path $\gamma = \gamma_1 \gamma_2 \ldots \gamma_{n-1} = \gamma_1 \gamma_2 \ldots \gamma_4$ according to the following scheme,

$$
\begin{aligned}
(N_j, E_j) &\longleftrightarrow \gamma_{j-1} = \text{up diagonal step } (1,1), \\
(N_j, N_j) &\longleftrightarrow \gamma_{j-1} = \text{ blue colored horizontal step } (1,0), \\
(E_j, E_j) &\longleftrightarrow \gamma_{j-1} = \text{red colored horizontal step } (1,0), \\
(E_j, N_j) &\longleftrightarrow \gamma_{j-1} = \text{down diagonal step } (1,-1),
\end{aligned}
$$

then, the pairs steps in the pair path (w, η) is

$$(N_2, N_2) \rightarrow (E_3, E_3) \rightarrow (E_4, E_4) \rightarrow (N_5, N_5)$$

It corresponds to the 2-colored Motzkin path

$$(0,0) \longrightarrow (1,0) \longrightarrow (2,0) \longrightarrow (3,0) \longrightarrow (4,0)$$

Figure 7.8: the corresponding 2-colored Motzkin path from the example polyomino

Step 3: Based on the given path $\gamma = \gamma_1 \gamma_2 \ldots \gamma_4$, using a bijection h, which is defined by the scheme,

$$h(\gamma_i) = \begin{cases} \text{(up diagonal, up diagonal)} & \text{if } \gamma_i = \text{up diagonal step}(1,1) \\ \text{(up diagonal, down diagonal)} & \text{if } \gamma_i = \text{blue colored horizontal step}(1,0) \\ \text{(down diagonal, up diagonal)} & \text{if } \gamma_i = \text{red colored horizontal step}(1,0) \\ \text{(down diagonal, down diagonal)} & \text{if } \gamma_i = \text{down diagonal step}(1,-1) \end{cases}$$

that is, h is a bijection between dyck path steps(except initial and final steps) and the 2-colored Motzkin path following the scheme

$$h(\gamma_i) = \begin{cases} & \\ & \end{cases}$$

Thus, in the example, the path $h(\gamma)$ is constructed as the following,

$$(1,1) \qquad\qquad (7,1)$$

$$(0,0) \quad (2,0) \quad (4,0) \quad (6,0) \quad (8,0)$$

$$(3,-1) \quad (5,-1)$$
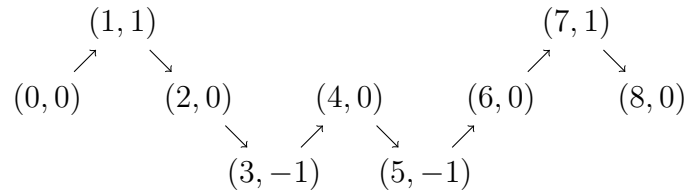
Figure 7.9: the corresponding path $h(\gamma)$ from the 2-colored Motzkin path

Step4: construct the Dyck path $(\text{up diagonal}, h(\gamma), \text{down diagonal})$ as follows,

$$(2,2) \qquad\qquad (8,2)$$

$$(1,1) \quad (3,1) \quad (5,1) \quad (7,1) \quad (9,1)$$

$$(0,0) \text{————} (4,0) \text{——} (6,0) \text{————} (10,0)$$
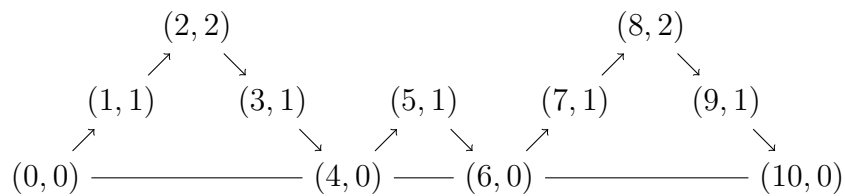
Figure 7.10: the final constructed Dyck path from the example polyomino

It can be observed that the construction from the parallelogram polyomino to the Dyck path is bijection, therefore, the converse construction from the Dyck path to the paralleogram can also be done similarly.

7.2.2 Share price Path interpretation

In this section, we gave the explicit interpretations of the share price path constructed from the section 7.2.1, which is important for financial modelling.

Share price Path interpretation

Assume that a jump up is a move $S_k \to S_{k+1} = uS_k$.
Then, a jump down is a move $S_k \to S_{k+1} = dS_k$.
Finally, a stay jump is a move $S_k \to S_{k+1} = S_k$.

The Figure 7.9 is then interpreted as

$$
\begin{aligned}
S_1 &= S_0 u^0 = S_0 \quad \text{(stay)} \\
S_2 &= S_1 u^1 = S_1 u \quad \text{(jump up)} \\
S_3 &= S_2 u^0 = S_2 \quad \text{(stay)} \\
S_4 &= S_3 u^{-1} \quad \text{(jump down)} \\
S_5 &= S_4 u^0 = S_4 \quad \text{(stay)} \\
S_6 &= S_5 u^{-1} \quad \text{(down)} \\
S_7 &= S_6 u^0 = S_6 \quad \text{(stay)} \\
S_8 &= S_7 u^1 = S_7 u \quad \text{(up)} \\
S_8 &= S_8 u^0 = S_8 \quad \text{(stay)}
\end{aligned}
$$

Then, this corresponds to compound Possion behaviour, before going up and down, stay at random time, compared to binomial model.

Chapter 8

Probability Methods

In the first part of the chapter, we apply a well-known probability results and a simple argument in probability to solve combinatorial problem, the probability result can be referred to [40, Section 1.1]. In the next part of the chapter, we go further and modify certain probability methods to solve another restricted path counting combinatorial problem.

8.1 Path counts using a probability method

8.1.1 Probability of simple random walks from $(0, a)$ to (n, b)

Considering a nearest neighbour random walk on $\mathbb{Z}$. Imagine a walker starts at $a \in \mathbb{Z}$ and at every integer time $n \in \mathbb{N}$ the walker has two options; moves one step to up with $P(X_n = 1) = p$ or moves one step to down with $P(X_n = -1) = q = 1 - p$. Denote the position of the walker at time n by S_n, the nearest neighbour random walk is defined by the following formula,

$$S_0 = a, \quad S_n = S_0 + \sum_{i=1}^{n} X_i$$

the set of realization of the walker is the set of sequences $S = (s_0, s_1, \ldots)$ with $s_0 = a$ and $s_{i+1} - s_i = \{+1, -1\} \ \forall i \in \mathbb{N}$, the sequence in the set is called a sample path of the random walk. The probability that the first n steps of the walk follow a given path $S = (s_0, s_1, \ldots, s_n)$ is $p^{up} q^{down}$, where

$$
\begin{aligned}
up &= \text{number of up steps in a path S} \\
&= \sharp\{i : s_{i+1} - s_i = 1\} \\
down &= \text{number of down steps in a path S} \\
&= \sharp\{i : s_{i+1} - s_i = -1\}
\end{aligned}
$$

Assuming the position of the random walk at time n is $s_n = b$, initial position is $s_0 = a$.

Denote $N_n^{up}(a, b)$ is the number of paths $(s_0, s_1, \ldots)$ with $s_0 = a$, $s_n = b$. An equation relating to number 'up' and 'down' is derived as follows,

$$
\begin{aligned}
up + down &= n \\
up - down &= (s_m - s_{m-1})^+ + (s_m - s_{m-1})^- \\
&= s_n - s_0 = b - a
\end{aligned}
$$

$\implies \quad \{up = \frac{1}{2}(n + b - a), down = \frac{1}{2}(n - b + a)\}$

Then, any event for the random walk can be expressed as follows,

$$
\begin{aligned}
P(S_n = b) &= N_n^{up}(a, b) p^{up} q^{down} \text{(here 'up' is a fixed number)} \\
&= \binom{n}{\frac{1}{2}(n + b - a)} p^{\frac{1}{2}(n+b-a)} q^{\frac{1}{2}(n-b+a)}
\end{aligned}
$$

It is obvious to see that computing certain random walk events are after counting the size of corresponding set of paths.

Therefore, by recalling the interesting question in percolation theory, a problem involving the number of ways of going from one point to another point in a square lattice would be an important question in the Problem 1. Before doing problem 1, the following definition is given.

Definition 69 *[41] Lattice path is a path consisting of connected horizontal and vertical line segments, which are restricted to north and east steps. assuming the line segment has nearest integer distance 1.*

8.1.2   Path counts using a simple probability argument

A simple argument in probability:

Assuming the sample space is Ω, if all the elements in Ω are equally likely, then the probability that an event $\{A\}$ happens in the sample space is

$$
P(A) = \frac{\text{number of elements in A}}{\text{number of elements in the sample space}}
$$

therefore, based on the discussion in section 8.1.1, assuming the lattice path employing two steps {up, right}, then we get,

$$
\begin{aligned}
P(S(n) = m) &= \frac{\text{number of paths of length } N \text{ and } k \text{ 'up' steps}}{\text{number of paths of length N}} \\
&= \frac{a_{n,m}}{a_n}
\end{aligned}
$$

Next, we provide an example of counting paths using the simple probability argument and the well-known probability results as follows.

Firstly, assuming that each step exists a random variable,

$$\varepsilon_i = \begin{cases} 1, \text{if up with } \frac{1}{2}0 \\ 0, \text{if right with } \frac{1}{2} \end{cases}$$

The number of up steps is:

$$\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = \sum_{i=1}^{3} \varepsilon_i = \sum_{i=1}^{m+n} \varepsilon_i = m$$

For each path, the total steps is (n+m) so, the number of right steps is

$$(n+m) - \sum_{i=1}^{n+m} \varepsilon_i = \sum_{i=1}^{n+m} (1 - \varepsilon_i)$$

Then, suppose each possible path happens with a equal possibility $\frac{1}{2^u}$, with $u = m + n$, by assumption,

{the number of paths from (0,0) to (n,m)}={the number of up steps=m}, the event is Bernoulli trials $(n+m, \frac{1}{2})$.

As the number of total possible paths is equal to 2^u, the number of paths from (0,0) to (n,m) is the product of 2^u and the possibility of {the number of up steps=m}.

Say m=1,n=2; the number of total possible paths is equal to 2^3,

$$P(\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 1) = P(Bin(3, \frac{1}{2}) = 1) = \binom{3}{1}\frac{1}{2^3}$$

so, the number of paths from(0,0) to (2,1) is

$$2^3 \cdot P(\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 1) = 3$$

8.2   Count paths via unusual stochatic modelling

Consider a finite N time steps discrete time model in which share prices is modelled by a restricted N time step binomial model with restriction not allowing k consecutive down steps. We aim to calculate the number of restricted share price paths not allowing k consecutive down steps at time $1 \leq T \leq N$ using Markov-type technique stochastic modelling.

We will illustrate the technique using two small examples, one is to calculate the number of restricted share paths not allowing 2 down steps in a row at $T = 3$ time step. The other is to calculate the number of restricted share paths not allowing 3 consecutive down steps at time $T = 4$.

We will firstly recall the definition of Markov chain, and its two important theorems, then, we would like to use the Markov-type technique to do the path calculation.

8.2.1 Markov chain

Definition 70 *[27]*

Consider a set of states $S = \{s_0, s_1, \ldots, s_n, \ldots\}$, *a discrete Markov chain is a family of random variables* $\{X_n\}_{n \in \mathbb{N}_0}$ *which take values in the set* S *and satisfies the Markov property*

$$
\begin{aligned}
& P(X_n = s_n | X_{n-1} = s_{n-1}, \ldots, X_0 = s_0) \\
= \; & P(X_n = s_n | X_{n-1} = s_{n-1}) \\
= \; & P_{s_{n-1}, s_n}(n-1)
\end{aligned}
$$

Consider the time-homogeneous Markov chains, the transition probability of a time homogeneous discrete Markov chain is defined as

$$
\begin{aligned}
P_{i,j}(n-1) \; = \; & P_{i,j} \\
= \; & P(X_{n-1} = j \mid X_n = i), \\
& where \sum_{j \in S} P_{i,j} = 1, \; for \; i \in S
\end{aligned}
$$

The initial state X_0 *of the Markov chain has a distribution defined as*

$$
\phi(i) = P(X_0 = i), \quad for \quad i \in S
$$

Using the Markov property, the distribution of a time-homogeneous discrete Markov chain is defined as the point probabilities

$$
\begin{aligned}
& P(X_n = s_n, X_{n-1} = s_{n-1}, \ldots, X_0 = s_0) \\
= \; & P(X_n = s_n | X_{n-1} = s_{n-1}) \times \ldots \times P(X_1 = s_1 | X_0 = s_0) \\
= \; & P_{s_{n-1}, s_n}(n-1) \times P_{s_{n-2}, s_{n-1}}(n-2) \times \ldots \times P_{s_0, s_1}(0) \\
= \; & P_{s_{n-1}, s_n} \times P_{s_{n-2}, s_{n-1}} \times \ldots \times P_{s_0, s_1}
\end{aligned}
$$

for $s_n, s_{n-1}, \ldots, s_0 \in S$ *and* $n \in \mathbb{N}_0$.

Theorem 71 *[27] [25, chapter 11]*

Consider a time-homogeneous discrete time Markov chain $\{X_n\}_{n \in S}$ *defined on a set of finite states,* $S = \{s_0, s_1, \ldots, s_n\}$,

Let $\mathbf{P}$ *be the transition matrix set up for the Markov chain, an (i,j)-th entry of the matrix is*

$P_{i,j} = P_{s_i, s_j} = P(X(n+1) = s_j \mid X(n) = s_i) = P(X(1) = s_j \mid X(0) = s_i),$

Let $R_{i,j}^{(n)} = R_{s_i, s_j}^{(n)} = P(X(n) = s_j \mid X(0) = s_i)$ *denotes the probability that the Markov chain will be in the state* s_j *after* n *time steps, if it initially starts from the state* s_i,

Then, the probability is calculated as the (i,j)-th entry of the matrix $\mathbf{P}^n$.

Here, a simple proof using conditional probabilities is given and the idea is from [26].

Proof:

$$
\begin{aligned}
R_{i,j}^{(n)} = R_{s_i,s_j}^{(n)} &= P(X(n) = s_j \mid X(0) = s_i), \quad \text{for } n \in \mathbb{N},\ s_i, s_j \in S \\
&= \sum_{k=1}^{n} P(X(n) = s_j, X(1) = s_k \mid X(0) = s_i)\ [\text{Conditioning on } X_1] \\
&= \sum_{k=1}^{n} P(X_n = s_j \mid X_1 = s_k) P(X_1 = s_k \mid X_0 = s_i)\ [\text{Markov property}] \\
&= \sum_{k=1}^{n} P(X_{n-1} = s_j \mid X_0 = s_k) P(X_1 = s_k \mid X_0 = s_i)\ [\text{Time-homogeneous}] \\
&= \sum_{k=1}^{n} P_{s_i,s_k} R_{s_k,s_j}^{(n-1)} = \sum_{k=1}^{n} P_{i,k} R_{k,j}^{(n-1)}
\end{aligned}
$$

Using matrix notation,

$$ R^{(n)} = \mathbf{P} \times R^{(n-1)} $$

then, it can be easily found that starting from any initial state s_i, for $1 \leq i \leq n$, the probability matrix after n time steps is calculated as

$$ R^{(n)} = \mathbf{P}^n $$

Theorem 72 *[27] [25, chapter 11]*

Consider a time-homogeneous discrete time Markov chain $\{X_n\}_{n \in S}$ defined on a set of finite states, $S = \{s_0, s_1, \ldots, s_n\}$,

Let $\mathbf{P}$ be the transition matrix set up for the Markov chain, an (i,j)-th entry of the matrix is

$P_{i,j} = P_{s_i,s_j} = P(X(n+1) = s_j \mid X(n) = s_i) = P(X(1) = s_j \mid X(0) = s_i),$

Let $\mathbf{P}^n$ be the n time step transition probability matrix that describes the probabilities that the Markov chain starts from any state $s_i \in S$ and arrives at a state $s_j \in S$ after n time steps. An (i,j)-th entry of the matrix is

$P_{i,j}^n = P_{s_i,s_j}^n = P(X(n) = s_j \mid X(0) = s_i),$

Let the initial state distribution of the Markov chain is

$\phi(i) = P(X_0 = i), \quad for \quad i \in S$

Then, the n time step state $X(n)$ of the Markov chain has a distribution

$$ P(X(n) = s_j) = \sum_{s_i \in S} \phi(s_i) P^n(s_i, s_j), \quad for \quad s_j \in S $$

Here, a simple proof using total probabilities is given by us and the notation representation is motivated from [27] [25, chapter 11].

Proof:

$$
\begin{aligned}
P(X(n) = s_j) &= \sum_{s_i \in S} P(X(n) = s_j \,,\, X(0) = s_i), \quad \text{for } s_j \in S \quad \text{[total probability]} \\
&= \sum_{s_i \in S} P(X_0 = s_i) P(X(n) = s_j \mid X(0) = s_i) \ \text{[Conditioning on } X_0] \\
&= \sum_{s_i \in S} \phi(s_i) \, \mathbf{P}^n(s_i, s_j)
\end{aligned}
$$

Using matrix notation,

if denote the initial distribution by a vector

$$
\phi = \{\phi(1), \ldots, \phi(n)\} = \{\phi(s_1), \ldots, \phi(s_n)\}
$$

denote the state distribution of the Markov chain at time n by a vector

$$
\begin{aligned}
\phi^{(n)} &= \{\, \phi^{(n)}(1), \ldots, \phi^{(n)}(n) \,\} = \{\, \phi^{(n)}(s_1), \ldots, \phi^{(n)}(s_n) \,\} \\
&= \{\, P(X(n) = s_1), \ldots, P(X(n) = s_n \,\}
\end{aligned}
$$

Then,

$$
\phi^{(n)} = \phi \times \mathbf{P}^n
$$

8.2.2   Method Inspiration

For a discrete time markov chain, the important is how to define the set of states to work on and how to set up the transition probabilities with which the states move in the markov chain.

We found in the restricted share price model, each step the share price can either go up or go down, as time goes progress, the number of possible valid states at each time is increasing. To avoid using the set of increasing possible states, we observe that at each time step, the share price can only go up or go down.

Consider all possible share price changes at each time step be the set of possible states we are interested in. If no restriction is applied, the share price model can be modelled as a regular markov chain, because the possible states at each time step is independent of each other time step.

In the restricted model, if the restriction is not allowing k consecutive down steps, then the restriction make sense starting after the $k-1$ time step.

Consider the restricted share price model is modelled by a markov-type chain. In the chain, the set of possible states in the model is the set of all possible paths in the first $k-1$ time steps.

When the restriction is applied in the share price model, construct a transform probability matrix to be a square matrix with the dimension 2^{k-1}, which describes the probabilities of the possible share price change at time k given the first $k-1$ time steps possible paths.

In the matrix, each row except the last row represents the probabilities of the possible changes at the time step k, which is independent of the first $k-1$ time steps share price paths. That is, $P(\text{up}) = p$ and $P(\text{down}) = 1-p$. Here, we will use a Markov-type model to do the share price paths calculation with the given restriction. So, we consider the case $p = \frac{1}{2}$

By analogy with the n step transition matrix of markov chains, we would like to introduce a square matrix that has dimension 2^{k-1} to denote a tranform probability matrix, which can help to predict probabilities after n steps. However, it is noted that each row only have two possible share price changes with $P(\text{up}) = \frac{1}{2}$ and $P(\text{down}) = \frac{1}{2}$.

Therefore we introduce a proper notation to solve the conflict between the number of possible states 2 and the number of the matrix dimension 2^{k-1}. We will illustrate the technique using two example in the implementation section.

Concerning the last row in the probability matrix, it is natural to have two ways to set up the transform probability matrix, the first way is to set up

$$
\begin{cases}
P(X(k) & = & \text{up}|X(k-1) = \text{down}, \ldots, X(0) = \text{down}) = 1 \\
P(X(k) & = & \text{down}|X(k-1) = \text{down}, \ldots, X(0) = \text{down}) = 0
\end{cases}
$$

the other way is to set up

$$
\begin{cases}
P(X(k) & = & \text{up}|X(k-1) = \text{down}, \ldots, X(0) = \text{down}) = \frac{1}{2} \\
P(X(k) & = & \text{down}|X(k-1) = \text{down}, \ldots, X(0) = \text{down}) = 0
\end{cases}
$$

8.2.3 Transform probability matrix setting up

We claim that the second way of setting up the last row probabilities in the transform matrix would be a proper way. It can be tested and illustrated using an example. After that, we will also explain the reasonable idea behind the setting up of the last row probabilities in the transform matrix using an example.

Example1: Setting up the last row probabilities in a probability matrix

Consider a three time steps binomial share price model with the restriction not allowing two consecutive down steps, the probability measure is $P(\text{up}) = \frac{1}{2}$ and $P(\text{down}) = \frac{1}{2}$.

Consider the restricted share price model is modelled by a Markov-type chain, the restriction in the model starts work after the first one time step. In the first one

time step, share price has $2 = 2^1$ possible changes, the set of states is denoted by $S = \{s0, s1\} = \{\text{up}, \text{down}\}$. Denote the up step by 1 and the down step by 0, then, $S = \{s0, s1\} = \{1, 0\}$.

There are two possible ways to set up the second row transform probabilities in the transform probability matrix.

1. One possible way of setting up

$$
\begin{cases}
P(X(1) & = & \text{up}|\ X(0) = \text{down}) = P(X(k) = 1|X(0) = 0) = 1 \\
P(X(1) & = & \text{down}|\ X(0) = \text{down}) = P(X(k) = 0|X(0) = 0) = 0
\end{cases}
$$

The model is modelled by a markov chain, the transform matrix actually is a transition matrix. The second up step given the first down step can be seen as an absorption state. That is,

$$
F = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}
$$

In the following, using notation $P(X(2), X(1) \mid X(0))$ to denote the corresponding conditional path, the up step and down step are denoted by 1 and 0 repectively.

Use the transition matrix to predict probabilities of
(3rd step state | 1st step state), it is a two step transition matrix,

$$
\begin{aligned}
F^2 & = \begin{bmatrix} \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times 1 & \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times 0 \\ 1 \times \frac{1}{2} + 0 \times 1 & 1 \times \frac{1}{2} + 0 \times 0 \end{bmatrix} \\
& = \begin{bmatrix} P(11 \mid 1) + P(10 \mid 1) & P(01 \mid 1) + P(00 \mid 1) \\ P(11 \mid 0) + P(10 \mid 0) & P(01 \mid 0) + P(00 \mid 0) \end{bmatrix} \\
& = \begin{bmatrix} P(1 \mid 1)^{(2)} & P(0 \mid 1)^{(2)} \\ P(1 \mid 0)^{(2)} & P(0 \mid 0)^{(2)} \end{bmatrix}
\end{aligned}
$$

From the entry $F^2(2, 1)$,

the probability of (3rd step $=$ up | 1st step $=$ down) $= 1 \times \frac{1}{2} = \frac{1}{2}$, it represents one path in the three time-steps share price model.

However, given the first step state, the path should be considered having the probability $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$, if considering each path is equal and without restrictions in the three time-steps model.

The reason is that when doing path counting, without restriction in the three-time-step model, the number of total possible paths in the last two time-steps is $4,$ and if put the restriction on the model, the number of total possible path is 3, consider each path is still equal, actually, the probability $\frac{1}{2}$ means it count one path (down,up,up) as two paths.

In general, setting up the probability of (3rd step = up | 1st step = down) to be an absorbing state, For any one path containing the patterns (down, up), if the number of the patterns (down, up) in the path is a number $k,$ this path will be counted as $2k$ paths, although it only represents the probability of one path in the n time-steps share price model without restriction applied in the model. As time T increases, if not allowing the consecutive two down steps and calculate the number of valid paths at time T will result in more improper path counts because it might contains more number of patterns (down, up).

Consider a general restriction, not allowing k consecutive down steps, we have to set up a transform matrix be a probability matrix that describes all possible share price change probabilities given the first 2^{k-1} time steps possible paths, and if using the first way of last row probability setting up, when doing path calculation, it counts any k step path containing the pattern (k consecutive downs, up) as two paths, as time goes progress, it will lead to more inaccurate restricted path counting.

Therefore, we claim that the probability of (3rd step = up | 1st step = down) $= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}.$

2. Another proper way of setting up

   The probabilities of the other rows in the transform probability matrix describes all possible share price change probabilities at time step k given the first 2^{k-1} time steps.

   The share price change has two possible changes in the set {up, down}, when the restriction not allowing k consecutive down steps is applied in the n time steps model, the two possible share price change probability at time $k,$ is independent of the first $k-1$ time steps information except the last row in the matrix, so, we set the probabilities of the rows from 1 to $2^{k-1}-1$ to be $P(\text{up} = \frac{1}{2}),\quad P(\text{down} = \frac{1}{2}).$

   To construct the transform probability matrix to predict the probabilities of states after n time steps, the matrix should be a square matrix, if the restriction is not allowing $k \geq 3$ consecutive down steps, there is a conflict between the number of possible share price change and the number of the first $k-1$ share price paths.

   To solve the conflict, notations are introduced as follows,

   Given the first $k-1$ time steps share price change, the valid state is to go to

a state which has the first $k-2$ share price change same as the last $k-2$ share price change in the given state, considering equal probabilities at each time step.

Let one possible path in the first $k-1$ time steps is denoted by $(s_0, s_1, \ldots, s_{k-2})$, where $s_i \in S = \{\text{ up, down }\} = \{1, 0\}$,

then, the probabilities of rows from 1 to $2^{k-1} - 1$ in the square transform matrix is

$$P((s_0, s_1, \ldots, s_{k-2}) \to ((s_1, s_2, \ldots, s_{k-2}, \text{up})$$
$$= P(s_0, s_1, \ldots, s_{k-2}, \text{up}) = \frac{1}{2}$$

$$P((s_0, s_1, \ldots, s_{k-2}) \to ((s_1, s_2, \ldots, s_{k-2}, \text{down})$$
$$= P(s_0, s_1, \ldots, s_{k-2}, \text{down}) = \frac{1}{2}$$

The last row $(2^{k-1}th)$ row of the probability matrix has probabilities

$$P((s_0 = \text{down}, \ldots, s_{k-2} = \text{down}) \to P((s_0 = \text{down}, \ldots, s_{k-2} = \text{down}, s_{k-1} = \text{up})$$
$$= P(s_0 = \text{down}, s_1 = \text{down}, \ldots, s_{k-2} = \text{down}, s_{k-1} = \text{up}) = \frac{1}{2}$$

$$P((s_0 = \text{down}, \ldots, s_{k-2} = \text{down}) \to P((s_0 = \text{down}, \ldots, s_{k-2} = \text{down}, s_{k-1} = \text{up})$$
$$= P(s_0 = \text{down}, s_1 = \text{down}, \ldots, s_{k-2} = \text{down}, s_{k-1} = \text{down}) = 0$$

Next, a reasonable idea behind the setting up of the last row probabilities as above in the transform matrix is explained using an example.

Consider a two-steps share price model, suppose the probability measure set up for the model is $P(\text{up}) = \frac{1}{2}$ and $P(\text{down}) = \frac{1}{2}$, each possible share price path has equal probability $\frac{1}{4}$.

Consider a restriction is not allowing 2 consecutive down steps and put the restriction on the two-steps share price model, the possible share price paths with corresponding probabilities and corresponding normalized probabilities are as follows,

$$
\begin{cases}
\qquad \nearrow & \text{with probability} \quad \tfrac{1}{4}, \\
\nearrow & \text{with normalized probability} \quad \tfrac{1}{3} \\[2em]
\nearrow \searrow & \text{with probability} \quad \tfrac{1}{4}, \\
& \text{with normalized probability} \quad \tfrac{1}{3} \\[2em]
\searrow \nearrow & \text{with probability} \quad \tfrac{1}{4}, \\
& \text{with normalized probability} \quad \tfrac{1}{3}
\end{cases}
$$
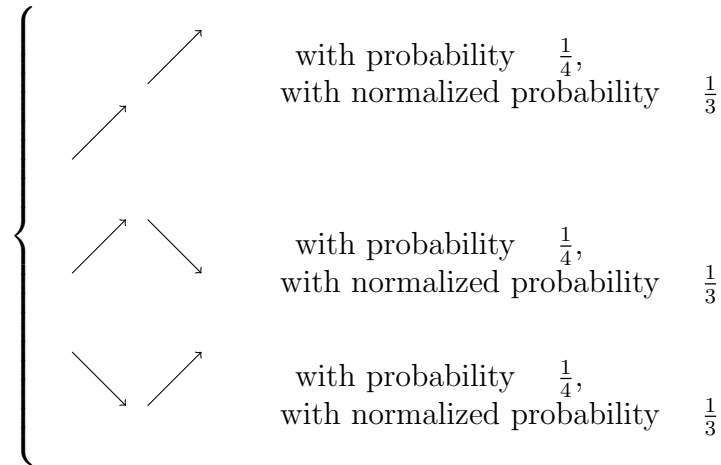
It can be observed that each path still has equal probability after put restriction on the model, if setting up the transform matrix as

$$
F \;=\; \begin{bmatrix}
P(1\mid 1)=\tfrac{1}{2} & P(0\mid 1)=\tfrac{1}{2} \\
P(1\mid 0)=\tfrac{1}{2} & P(0\mid 0)=0
\end{bmatrix}
$$

8.2.4   Result and Discussion

Lemma 73 *Consider a restricted finite N time steps binomial model with two possible share price change at each time step, either going up or going down, and the restriction not allowing k consecutive down steps.*

Consider the model is modelled as a Markov-type chain, with the initial state distribution is denoted by a vector

$$
\begin{aligned}
\phi = [\,\phi(s_1),\ldots,\phi(s_{2^{k-1}})\,] \;&=\; [\,P(X(0)=s_1),\ldots,P(X(0)=s_{2^{k-1}})\,] \\
&=\; [\,p_1,\ldots,p_{2^{k-1}}\,] = [\,\frac{1}{2},\ldots,\frac{1}{2}\,]
\end{aligned}
$$

Let the transform probability matrix is denoted by $\mathcal{A}$, the possible states in the matrix is ordered as the decreasing order of the $(k-1)$ digits binary expression of $\{0,\,1\}$ starting from $(1,1,\ldots,1)$ to $(0,0,\ldots,0)$, the matrix $\mathcal{A}$ is defined as a square matrix with dimension 2^{k-1},

$$
\mathcal{A} \;=\; \begin{bmatrix}
\tfrac{1}{2} & \tfrac{1}{2} & 0 & 0 & \ldots & 0 & \tfrac{1}{2} \\
0 & 0 & \tfrac{1}{2} & \tfrac{1}{2} & \ldots & 0 & \tfrac{1}{2} \\
\vdots & & & \ddots & & & \vdots \\
0 & 0 & \ldots & \tfrac{1}{2} & \tfrac{1}{2} & 0 & 0 \\
0 & 0 & 0 & 0 & \ldots & \tfrac{1}{2} & 0
\end{bmatrix}
$$

Then, the number of restricted paths not allowing k consecutive down steps at time $1 \leq T \leq N$ is defined as

$$\psi_{T,k} = 2^T \times \sum_{i \geq 1}^{2^{k-1}} \phi \cdot \mathcal{A}^{T-(k-1)}(i)$$

Example 8.2.1 *Consider a finite time discrete time binomial share price model, what is the number of restricted paths not allowing 2 down steps in a row at time 3.*

Solution: $k = 2$, $T = 3$, the restriction starts work after $(k-1) = 1$ steps,

so, by decreasing order of the 1 digits binary expression of $\{0, 1\}$ the set of states is $S = \{s_0, s_1\} = \{\text{up, down}\} = \{1, 0\}$,

set up the 2×2 matrix

$$\mathcal{A} \;=\; \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$$

the initial state distribution is

$$\begin{aligned}
\phi \;&=\; [\, \phi(s_1), \; \phi(s_2) \,] = [\, p_1, \; p_2 \,] \\
&=\; [\, P(X(0) = 1), \; P(X(0) = 0) \,] = [\, \frac{1}{2}, \; \frac{1}{2} \,]
\end{aligned}$$

the restricted state distribution at time $T = 3$ is a vector

$$\begin{aligned}
\phi \cdot \mathcal{A}^2 \;&=\; [\, \phi(1), \; \phi(0) \,] \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} \\
&=\; \frac{1}{8} \cdot [\, 1, \; 1 \,] \cdot \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = [\, \frac{3}{8}, \; \frac{2}{8} \,]
\end{aligned}$$

Then, the number of restricted paths is

$$\begin{aligned}
\psi_{3,2} \;&=\; 2^3 \times \sum_{i \geq 1}^{2} \phi \cdot \mathcal{A}^2(i) \\
&=\; 2^3 \times \frac{5}{8} = 5
\end{aligned}$$

Example 8.2.2 *Consider a finite time discrete time binomial share price model, what is the number of restricted paths not allowing 3 down steps in a row at time 5.*

Solution: $k = 3$, $T = 5$, the restriction starts work after $(k - 1) = 2$ steps,
so, by decreasing order of the $(k - 1) = 2$ digits binary expression of $\{0, 1\}$ the
set of states is

$$
\begin{aligned}
S &= \{s_0, s_1, s_2, s_3\} \\
&= \{(\text{up, up}), (\text{up, down}), (\text{down, up}), (\text{down, down}) \} \\
&= \{11, 10, 01, 00\}
\end{aligned}
$$

set up the 4×4 matrix

$$
\mathcal{A} =
\begin{bmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 \\
0 & 0 & \frac{1}{2} & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 \\
0 & 0 & \frac{1}{2} & 0
\end{bmatrix}
$$

the initial state distribution is

$$
\begin{aligned}
\phi &= [\, \phi(s_1), \; \phi(s_2), \; \phi(s_3), \; \phi(s_4) \,] \\
&= [\, p_1, \; p_2, \; p_3, \; p_4 \,] \\
&= [\, P(X(0) = 11), \; P(X(0) = 10), \; P(X(0) = 01), \; P(X(0) = 00) \,] \\
&= [\, \frac{1}{4}, \; \frac{1}{4}, \; \frac{1}{4}, \; \frac{1}{4} \,]
\end{aligned}
$$

the restricted state distribution at time $T = 4$ is a vector

$$
\begin{aligned}
\phi \cdot \mathcal{A}^2 &= [\, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \,] \cdot
\begin{bmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 \\
0 & 0 & \frac{1}{2} & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 \\
0 & 0 & \frac{1}{2} & 0
\end{bmatrix}^2
= \frac{1}{16} \cdot [\, 1, 1, 1, 1 \,] \cdot
\begin{bmatrix}
1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}^2 \\[2mm]
&= \frac{1}{16} \cdot [\, 1, 1, 1, 1 \,] \cdot
\begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 \\
1 & 1 & 0 & 0
\end{bmatrix}
= \frac{1}{16} \cdot [\, 4, 4, 3, 2 \,]
\end{aligned}
$$

Then, the number of restricted paths is

$$
\begin{aligned}
\psi_{4,3} &= 2^4 \times \sum_{i \geq 1}^{4} \phi \cdot \mathcal{A}^2(i) \\
&= 2^4 \times \frac{13}{16} = 13
\end{aligned}
$$

In general case, following the same argument and proof discussion in the Section 8.2.1, 8.2.2, 8.2.3, we provided the Lemma 74 as follows.

Lemma 74 *Consider a restricted finite N time steps binomial model with two possible share price change at each time step, either going up or going down. The restriction is not allowing any given fixed steps pattern $y_1 y_2 \ldots y_k$ with $y_i \in \{1, 0\}$. The length of the steps pattern is k.*

Consider the model is modelled as a Markov-type chain, with the initial state distribution is denoted by a vector

$$
\begin{aligned}
\phi &= [\, \phi(s_1), \ldots, \phi(s_{2^{k-1}}) \,] \\
&= [\, P(X(0) = s_1), \ldots, P(X(0) = s_{2^{k-1}}) \,] \\
&= [\, p_1, \ldots, p_{2^{k-1}} \,] = [\, \frac{1}{2}, \ldots, \frac{1}{2} \,],
\end{aligned}
$$

Where, s_j $= x_1 x_2 \ldots x_{k-1}$ *with* $x_i \in \{1, 0\}$
and with $decimal(s_1) > decimal(s_2) > \ldots > decimal(s_{2^{k-1}})$
   and $decimal(s_j) = decimal(x_1 x_2 \ldots x_{k-1}) = x_1 \cdot 2^{k-2} + x_2 \cdot 2^{k-3} + \ldots + x_{k-1} \cdot 2^0$

Let the transform probability matrix is denoted by B,   the possible states in the matrix is ordered as the decreasing order of the $(k-1)$ digits binary expression of $\{0,\ 1\}$ starting from $(1, 1, \ldots, 1)$ to $(0, 0, \ldots, 0)$,   the matrix B is defined as a square matrix with dimension 2^{k-1}.   The matrix B is set up as follows,

Let $s_i = x_1 x_2 \ldots x_{k-1}, \quad s_j = \tilde{x}_1 \tilde{x}_2 \ldots \tilde{x}_{k-1}$,

If $x_2 x_3 \ldots x_{k-1} = \tilde{x}_1 \tilde{x}_2 \ldots \tilde{x}_{k-2}$,

then,

$$
\begin{aligned}
(s_i \to s_j) &= (x_1 x_2 \ldots x_{k-1} \to \tilde{x}_{k-1}) \\
&= (\text{the k-th step is } \tilde{x}_{k-1} \mid \text{the first k-1 steps equals } x_1 x_2 \ldots x_{k-1})
\end{aligned}
$$

The matrix elements in the matrix B is set up as

$$
B_{i,j} = P(s_i \to s_j) = \begin{cases} \frac{1}{2} & \text{if } (x_1 x_2 \ldots x_{k-1} \to \tilde{x}_{k-1}) \neq (y_1 y_2 \ldots y_{k-1} \to y_k) \\ 0 & \text{if } (x_1 x_2 \ldots x_{k-1} \to \tilde{x}_{k-1}) = (y_1 y_2 \ldots y_{k-1} \to y_k) \end{cases}
$$

If $x_2 x_3 \ldots x_{k-1} \neq \tilde{x}_1 \tilde{x}_2 \ldots \tilde{x}_{k-2}$, then,

$$
(s_i \to s_j) \quad \text{not forms a possible path}
$$
$$
\text{and} \quad P(s_i \to s_j) = 0
$$

Then, the number of restricted paths not allowing the given pattern $y_1 y_2 \ldots y_k$ at

time $1 \leq T \leq N$ is defined as

$$\psi_{T,k} = \psi_T = 2^T \times \sum_{i \geq 1}^{2^{k-1}} \phi \cdot B^{T-(k-1)}(i)$$

Where, $\psi_{T,k} = \psi_T$, because the length of
given avoiding pattern k is a fixed number

Example 8.2.3 *Consider a finite time discrete time binomial steps share price model, what is the number of restricted paths not allowing (down, up) steps pattern in a row at time 3.*

Solution: The length of the avoiding pattern (down, up) is $k = 2$, $T = 3$, the restriction starts work after $(k - 1) = 1$ step.

The possible states are $s_j = x_1 x_2 \ldots x_{k-1}$ with $1 \leq j \leq 2^{k-1}$, $x_i \in \{0, 1\}$

so, the possible state in the model are $s_j = x_1$ with $1 \leq j \leq 2$, $x_1 \in \{0, 1\}$

Let $decimal(s1) > decimal(s_2)$, then, the possible initial states are $s_1 = 1, s_2 = 0$, the possible states are obtained by decreasing order of the 1 digits binary expression of $\{0, 1\}$, the set of possible states in the model is $S = \{s_1, s_2\} = \{\text{up, down}\} = \{1, 0\}$.

Next, calculate the initial state distribution as follows,

$$
\begin{aligned}
\phi &= [\, \phi(s_1),\ \phi(s_2) \,] \\
&= [\, P(X(0) = s_1),\ P(X(0) = s_2) \,] \\
&= [\, P(X(0) = 1),\ P(X(0) = 0) \,] = [\, \frac{1}{2}, \frac{1}{2} \,]
\end{aligned}
$$

Next, set up the transform probability matrix for the model. Let the transform probability matrix is denoted by B.

The matrix B is a square matrix with dimension $2^{k-1} \overset{k=2}{=} 2$

the 2×2 matrix B is set up as follows,

Let $s_i = x_1 x_2 \ldots x_{k-1} \overset{k=2}{=} x_1$, $s_j = \tilde{x}_1 \tilde{x}_2 \ldots \tilde{x}_{k-1} \overset{k=2}{=} \tilde{x}_1$ It can be observed that $x_2 x_3 \ldots x_{k-1} = \tilde{x}_1 \tilde{x}_2 \ldots \tilde{x}_{k-2} = \text{empty}$, then,

$$
\begin{aligned}
(s_i \to s_j) &= (x_1 x_2 \ldots x_{k-1} \to \tilde{x}_{k-1}) \\
&= (x_1 \to \tilde{x}_1), \\
&= (\text{the second step is } \tilde{x}_1 \mid \text{the first step equals } x_1) \\
&\quad \text{with } x_1 \tilde{x}_1 \in \{01\}
\end{aligned}
$$

The matrix elements in the matrix B is set up as

$$B_{i,j} = P(s_i \to s_j) = \begin{cases} \frac{1}{2} & \text{if } (x_1 \to \tilde{x}_1) \neq (0 \to 1) \\ 0 & \text{if } (x_1 \to \tilde{x}_1) = (0 \to 1) \end{cases}$$

So, the matrix elements B is

$$
\begin{aligned}
B_{1,1} &= P(1 \to 1) = \frac{1}{2} \\
B_{1,2} &= P(1 \to 0) = \frac{1}{2} \\
B_{2,1} &= P(0 \to 1) = 0 \\
B_{2,2} &= P(0 \to 0) = \frac{1}{2}
\end{aligned}
$$

and

$$B = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}$$

Then, the restricted state distribution at time $T = 3$ is a vector

$$
\begin{aligned}
\phi \cdot B^{T-(k-1)} \overset{k=2}{\underset{T=3}{=}} \phi \cdot B^2 \\
&= [\frac{1}{2}, \frac{1}{2}] \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} \\
&= [\frac{1}{8}, \frac{3}{8}]
\end{aligned}
$$

Then, the number of restricted paths not allowing the pattern $y_1 y_2 = \text{down, up} = 01$ at time $T = 3$ is definied as

$$
\begin{aligned}
\psi_{3,2} = \psi_3 &= 2^3 \times \sum_{i \geq 1}^{2} \phi \cdot B^2(i) \\
&= 2^3 \times \frac{4}{8} = 4
\end{aligned}
$$

It is easy to check that in the binomial steps share price model and set up restriction of avoiding pattern $y_1 y_2 = \text{down} \quad \text{up}$, there only have 4 valid path at time $T = 3$.

In order to construct the transform matrix B in the Lemma 74 more easily using the computer programming language, we provide another Lemma 75 for the restricted path counting in a the same model.

Lemma 75 *Consider a restricted finite N time steps binomial model with two possible share price change at each time step, either going up or going down. The restriction is not allowing any given fixed steps pattern $y_1 y_2 \ldots y_k$ with $y_i \in \{1, 0\}$. The length of the steps pattern is k.*

Consider the model is modelled as a Markov-type chain, with the initial state distribution is denoted by a vector

$$
\begin{aligned}
\phi & = [\, \phi(s_1), \ldots, \phi(s_{2^{k-1}}) \,] \\
& = [\, P(X(0) = s_1), \ldots, P(X(0) = s_{2^{k-1}}) \,] \\
& = [\, p_1, \ldots, p_{2^{k-1}} \,] = [\, \frac{1}{2}, \ldots, \frac{1}{2} \,],
\end{aligned}
$$

Where, $\quad s_j \quad = x_1 x_2 \ldots x_{k-1}$ with $x_i \in \{1, 0\}$
and with $\quad decimal(s_1) > decimal(s_2) > \ldots > decimal(s_{2^{k-1}})$
and $\quad decimal(s_j) = decimal(x_1 x_2 \ldots x_{k-1}) = x_1 \cdot 2^{k-2} + x_2 \cdot 2^{k-3} + \ldots + x_{k-1} \cdot 2^0$

Let the transform probability matrix is denoted by B, the possible states in the matrix is ordered as the decreasing order of the $(k-1)$ digits binary expression of $\{0, 1\}$ starting from $(1, 1, \ldots, 1)$ to $(0, 0, \ldots, 0)$, the matrix B is defined as a square matrix with dimension 2^{k-1}. The matrix B is set up as follows,

Let $s_i = x_1 x_2 \ldots x_{k-1}, \quad s_j = \tilde{x}_1 \tilde{x}_2 \ldots \tilde{x}_{k-1},$

The matrix index of $(s_i \to s_j)$ with $x_2 x_3 \ldots x_{k-1} = \tilde{x}_1 \tilde{x}_2 \ldots \tilde{x}_{k-2}$

and $(x_1 x_2 \ldots x_{k-1} \to \tilde{x}_{k-1}) = (y_1 y_2 \ldots y_{k-1} \to y_k)$ is set up as

$(2^{k-1} - decimal(y_1 y_2 \ldots y_{k-1}), \quad decimal(y_1 y_2 \ldots y_{k-1}) + 1),$

therefore, the matrix elements $B_{i,j}$ can be directly set up as

$B_{i,j} = 0$ with $i = 2^{k-1} - decimal(y_1 y_2 \ldots y_{k-1})$, and $j = decimal(y_1 y_2 \ldots y_{k-1}) + 1$

and

$B_{i,j+1} = \dfrac{1}{2}$ with $i = 2^{k-1} - decimal(y_1 y_2 \ldots y_{k-1})$, and $j = decimal(y_1 y_2 \ldots y_{k-1}) + 1$

and

$B_{i,j} = 0$ with $i = 2^{k-1} - decimal(y_1 y_2 \ldots y_{k-1})$, and $\quad j \neq decimal(y_1 y_2 \ldots y_{k-1}) + 1$
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ and $j \neq decimal(y_1 y_2 \ldots y_{k-1}) + 2$

Then, the next is to set up the other matrix elements $B_{i,j}$ with $i \neq 2^{k-1} - decimal(y_1 y_2 \ldots y_{k-1})$, and the set up is as follows,

$$
B_{i,j} = P(s_i \to s_j) = \begin{cases} \frac{1}{2} & \text{if } x_2 x_3 \ldots x_{k-1} = \tilde{x}_1 \tilde{x}_2 \ldots \tilde{x}_{k-2} \\[2mm] 0 & \text{if } x_2 x_3 \ldots x_{k-1} \neq \tilde{x}_1 \tilde{x}_2 \ldots \tilde{x}_{k-2} \end{cases}
$$

Then, the number of restricted paths not allowing the given pattern $y_1 y_2 \ldots y_k$ at time $1 \leq T \leq N$ is defined as

$$\psi_{T,k} = \psi_T = 2^T \times \sum_{i \geq 1}^{2^{k-1}} \phi \cdot B^{T-(k-1)}(i)$$

Where, $\psi_{T,k} = \psi_T$, because the length of
given avoiding pattern k is a fixed number

8.2.5   Implementation

The Lemma in the section 4 can be implemented using computer language MAT-LAB. When doing the implementation using computer, it can be noted that the size of the dimension of the transform matrix of the discrete time stochastic process increase exponentially as the restriction k increases.

To save the computer memory, it can be observed that the transform matrix can be constructed using sparse matrix, therefore, the counts can be easily done using the MATLAB.
The implementation code using MATLAB is put in the appendix.

Chapter 9

Matroid

9.1 Definitions of Matroids

Definition 76 *[46, Page 2] [47, Definition 1.1]*
A matroid M is a pair $(E, \mathcal{I})$, E is a finite set(are called ground set), I is a collection of subsets of E (the subsets are called independent sets such that

(i) $\emptyset \in \mathcal{I}$.

(ii) If $I \subset J, J \in \mathcal{I}$, then $I \in \mathcal{I}$.

(iii) If $I, J \in \mathcal{I}$ and $|I| < |J|$, then, there exists $j \in J - I$, such that $(I \cup J) \in \mathcal{I}$

Note: 1. $\mathcal{I}$ is the collection of independent sets.

2. A matorid is often denoted by listing its ground set and its family of independent sets, that is, denoted by $M = (E, \mathcal{I})$ or to emphasize the matroid, denoted by $M = (E(M), \mathcal{I}(M))$.

Definition 77 *[46, Page 5]*
In a matroid $M = (E, \mathcal{I})$, a subset B of the ground set is called maximal independent set iff the subset is independent(i.e. $B \in \mathcal{I}$) and no subsets of the ground set that containing it is independent(i.e. $\nexists C \subseteq E$, such that $C \supset B$ and $C \in \mathcal{I}$).

Remark:

1. Axiom (iii) of the definition of a matroid implies that all maximal independent sets have the same cardinality. The proof following [46, Page 5] is as follows,

   Proof: if B_1, B_2 are two maximal independent sets, suppose they do not have the same cardinality, Let $|B_1| < |B_2|$, B_1, B_2 are independent sets, by the definition of a matroid, its independent sets satisfies the third axiom (iii), that is,

   $$B_1 \in \mathcal{I}, B_2 \in \mathcal{I}, |B_1| < |B_2| \implies \exists b_2 \in (B_2 - B_1) \text{ such that } (B_1 \cup b_2) \in \mathcal{I}$$

   That is, B_1 is contained in the set $(B_1 \cup B_2)$ which is a bigger independent set. So, B_1 is not a maximal independent set. It contradicts the given condition that B_1 is a maximal independent set. Therefore, all maximal independent sets of a matorid have the same cardinality.

Definition 78 *[46, page 5]*
A maximal independent set is called a basis(or base) of the matorid, often denoted by B .

Remark [46, page 12, 14]:

1. Notice that a matroid can also be defined as $M = (E(M), \mathcal{B}(M))$. As given a set $\mathcal{B}(M)$ of all bases of a matroid, the family of independent sets $\mathcal{I}(M)$ can be recovered by taking all subsets of some bases B in the set $\mathcal{B}(M)$ .That is, a collection of independent sets is defined as
$\mathcal{I}(M) = \{I \subseteq E : I \subseteq B \text{ for some } B \in \mathcal{B}(M)\}$.

2. By using the definition $M = (E(M), \mathcal{B}(M))$, it is cost saving definition of a matroid, as no need to list all the independent sets of the matroid, only list the maximal independent sets of the matroid.

Three examples are given as follows to illustrate the definition of the matroid.

Example 9.1.1 *uniform matroid [48]*

The unifrom definition in the example is motivated from [48], then, we gave the explicit proof to show the proof of uniform matroid.

Let $M = (E, \mathcal{I})$ with the collection of independent sets is defined as $\mathcal{I} = \{X \subseteq E : |X| \leq k\}$, for a given $k \in \mathbb{N}$. Then, $M = (E, \mathcal{I})$ is a matroid. Following the definition of the matroid, we give the proof as follows,

Proof: Obviously, axiom (i),(ii) of definition of a matorid is satisfied; as the size of the empty set $|\emptyset| = 0$, if $Y \in \mathcal{I}$, then, it has size less than the given length k, so, all its subsets $X \subset Y$, has size no more than the given length k,that is, $X \in \mathcal{I}$.

It is left to check axiom (iii). That is, Let $X, Y \in \mathcal{I}$ and $|X| < |Y|$, then $|X| \leq (k-1)$ and $|Y| \leq k$, so, there must exists an element $e \in (Y - X)$ such that $|X \cup e| \leq k$.

Example 9.1.2 *partition matroid [49]*

The definition and proof of the partition matroid is motivated from [49], we added clear explanation.

Let $M = (E, \mathcal{I})$, Let $E_1, E_2, \ldots, E_l$ be disjoint sets that partition the ground set E. Let the collection of independent sets is defined as $\mathcal{I} = \{X \subseteq E : |X \cap E_i| \leq k_i \quad \text{for all} \quad i = 1, 2, \ldots, l\}$, for some given length parameters $k_1, k_2, \ldots, k_l$. Then, $M = (E, \mathcal{I})$ is a matroid.

Proof: Analogous to the example 1, it is obvious that axiom (i),(ii) of definition of a matorid is satisfied.

It is left to check axiom (iii). That is, Let $X, Y \in \mathcal{I}$ and $|X| < |Y|$, as the sets E_i are disjoint sets, then, there must exist i such that $|Y \cap E_i| > |X \cap E_i|$, so, it must

$\exists\ e \in (E_i \cap |Y - X|)$, such that, if the element e is added into X, there will be $|(X \cup e) \cap E_i| \le k_i$, this is because $|Y \cap E_i| \le k_i$.

Remark:

The example and discussion in the remarks are motivated from the discussion on page 1 in the [49].

1. *Notice that M will not be a matroid iff the sets E_i are not disjoint. For example, let $M = (E, \mathcal{I})$, Let E_1, E_2 be disjoint sets that partition the ground set E, with $E_1 = \{1, 2\}$, $E_2 = \{2, 3\}$, and $\mathcal{I}$ is defined as $\mathcal{I} = \{X \subseteq E : |X \cap E_i| \le k_i \quad \text{for all} \quad i = 1, 2 \}$, for some given length $k_1 = 1, k_2 = 1$.*

   Then, $Y = \{1, 3\}$, $X = \{2\}$. Both X and Y obviously satisfies $\{X \subseteq E : |X \cap E_i| \le k_i \quad \forall i = 1, 2 \}$ with $k_i = 1$, that is, $X, Y \in \mathcal{I}$ and $|X| < |Y|$, but if add any element $e \in |Y - X|$, i.e. elements 1 or 3, it will lead $|(X \cup e) \cap E_i| = 2$, which is greater than the given length $k_i = 1$, $i = 1, 2$. therefore, $(X \cup e)$ will not be in the set $\mathcal{I}$, thus not satisfying the axiom (iii) of the definition of the matroid.

Motivated from the free matroid mentioned on page 1 in the [49], we give another example of matroid which is the special case of example 1, in which the independent set is the power set of ground set.

Example 9.1.3 *a matroid where the independent set is the power set*

Let $M = (E, \mathcal{I})$, with the collection of independent sets defined to be the power set of the ground set E. Actually, if let the size of the ground set E, $|E| = n$, the power set of the ground set can be defined as $\mathcal{I} = \{X \subseteq E : |X| \le n\}$, which is a special case $(k = |E| = n)$ of Uniform matroid in example 1. Then, $M = (E, \mathcal{I})$ is a matroid.

Proof: By definition of the matroid, it is obvious that the power set 2^E satisfies the three axioms of independent sets of a matroid, because it contain all subsets of the ground set E. Also, it is observed that this matroid is a uniform matroid with the given length $(k = |E| = n)$.

There are three common classes of matroid which are linear matroid, graphic matroid and matching matroid(also called transversal matroid).[46, Page 4]

1. Linear Matroid

Proposition 79 *[49, section 4.1] [50, section 3.2, 4.2]*
Let a ground set E be the index set of columns of a matrix A, Let A_X be the submatrix of A consisting only of those columns indexed by a subset $X \subseteq E$, Let the collection of independent sets are defined as $\mathcal{I} = \{X \subseteq E : rank(A_X) = |X|\}$ Then, $M = (E, \mathcal{I})$ is a matroid.

Remark:

In the remarks, the knowledge of rank of a matrix can be referred to [52]

1. Recall the definition of rank of a matrix;

$$rank(A_X) = dim\,(span(\text{the set of columns of submatrix }\ A_X))$$
$$= \text{maximum number of independen column vectors of } A_X$$
$$= dim\,(span(\text{the set of rows of submatrix }\ A_X))$$
$$= \text{maximum number of independen row vectors of } A_X$$

2. By the definition of rank of the set of column vectors, if $rank(A_X) = |X| = $ number of columns of submatrix A_X, it is known that a index set $X \subseteq E$ are independent set iff its corresponding columns are linearly independent.

Following the idea in [49, section 4.1], we give more clear proof.

Proof: Firstly, it is obvious that axiom (i),(ii) of definition of a matorid are satisfied; as the subsets of linearly independent columns of matrix are also linearly independent columns.

It is left to check axiom (iii). That is, Let $X, Y \in \mathcal{I}$ and $|X| < |Y|$, then, $rank(A_X) = |X|$, $rank(A_Y) = |Y|$, it means, A_X has a full column rank, with the dimension of the span space of the set of columns of A_X equal to the size of the index set $|X|$, i.e.

$$dim\,(span(\text{the set of columns of submatrix } A_X)) = |X|$$

Similarly,

$$rank(A_Y) = |Y| \implies dim\,(span(\text{the set of columns of submatrix} A_Y)) = |Y|$$

If $|X| < |Y|$, there must exists a column e of A_Y, which is not in the $span(\text{the set of columns of} A_X)$ that means, the column is linearly independent with the set of columns of A_X, when adding this column to the submatrix A_X, the rank of new matrix $(A_X \cup e)$ is increased by 1, it is observed that the size of the index set is also increased by 1, therefore,

$$rank(A_X \cup e) = |X| + 1 = |A_X \cup e|$$

it satisfies the axiom (iii) of the definition of the matroid.

Remark:

1. In the above proposition, the linear matroid is defined as a full matrix column index set(ground set) and subsets of the ground set that are independent sets, where independent sets are defined as index sets where their corresponding columns are linearly independent columns, i.e. $\mathcal{I} = \{X \subseteq E : rank(A_X) = |X|\}$. However, there is another way to define a linear matroid which are given as the following proposition.

Proposition 80 *[46, page 3] [51]*
Let a ground set E be a finite set of vectors in a vector space V , let $\mathcal{I}$ be a collection of linear independent subsets of E .
Then, $M = (E, \mathcal{I})$ is a matroid.

Proof:

(I1) check axiom (i) of definition of a matorid; for empty set $\emptyset$, as it has no vectors in the set, therefore, it has no linear relations, that is, it is a linear independent subset of E , i.e $\emptyset \in \mathcal{I}$

(I2) check axiom (ii) If $I \subset J$ and $J \in \mathcal{I}$, then I is also an independent set, i.e., $I \in \mathcal{I}$

(I3) it is left to check axiom (iii). the proof for axiom (iii) is followed from [51]. That is, Let $I, J \in \mathcal{I}$ and $|I| < |J|$,   Let $I = \{i_1, i_2, \ldots, i_n\}$,

suppose $I \cup j$ is not independent set for any $j \in J - I$ then,

$$c_1 i_1 + c_2 i_2 + \ldots + c_n i_n + c \cdot j = 0, \quad \text{where } c_i, c \text{ are scalars, and } c \neq 0$$
$$\implies j = \frac{-c_1 i_1 - c_2 i_2 - \ldots - c_n i_n}{c}$$
$$\implies j \in span(i_1, i_2, \ldots, i_n), \quad \text{i.e.} \quad j \in span(I)$$
$$\implies (J - I) \subseteq span(I), \quad \text{as } j \text{ is for any } j \in (J - I)$$
$$\overset{I \subseteq span(I)}{\implies} ((J - I) \cup I) \subseteq span(I), \quad \text{i.e.} \quad J \subseteq span(I)$$
$$\implies span(J) \subseteq span(I)$$

As J is an independent set, that is, it contain $|J|$ linear independent vectors, so, $dim(span(J)) = |J|$.   Similarly, $I \in \mathcal{I} \implies dim(span(I)) = |I|$. Therefore,

$$span(J) \subseteq span(I) \implies dim(span(J)) \leq dim(span(I))$$
$$\implies |J| \leq |I|$$

It contradicts the given condition $|J| > |I|$ in the axiom (iii) of the definition of a matroid.

Definition 81 *[46, Page 4]*
Two matroids (E, I) and (E', I') are isomorphic iff there is a bijection from E to E' which induces a bijection from I to I' . In other words, the two matroids are the same matroid with different names for the elements of the ground set.

Remark:

In the remarks, we gave the discussion that the two definition of linear matroid actually gives two matroids which are isomorphic.

1. In the above two propositions, the linear matroid is defined as a full matrix column index set(ground set) which is has one-to-one correspondence to a finite set of column vectors in a column vector space. Besides, the collection of index sets for linearly independent columns obviously has one-to-one correspondence to a collection of linear independent subsets of the finite set of vectors of a columns vector space.

2. Graphic Matroid

The proposition 82 and its proof idea is stated according to the video lecture 2 Matroids given by Federico [51].

Proposition 82 *Let $G = (V, E)$ be a graph, Let $\mathcal{I}$ be a collection of independent sets of edges, the independent sets are defined as*
$\mathcal{I} = \{X \subseteq E : X$ does not contain any cycle$\}$,
Then, $M = (E, \mathcal{I})$ is a matroid.

Proof:

(I1) check axiom (i) of definition of a matorid; for empty set $\emptyset$, as it has no edges in the set, therefore, it has no cycles, that is, $\emptyset \in \mathcal{I}$.

(I2) check axiom (ii) If $I \subset J$ and $J \in \mathcal{I}$, because $J \in \mathcal{I}$ and then J has no cycles, all its subgraph will not contain cycles, so, I has no cycles and $I \in \mathcal{I}$.

(I3) it is left to check axiom (iii). That is, Let $I, J \in \mathcal{I}$ and $|I| < |J|$,   to prove there exists an edge element $j \in (J - I)$, if it is added into the graph (V, I), then, it forms an augmented independent set $(I \cup j) \in \mathcal{I}$.

Note: If an edge subset I of the graph G's edge set E is an independent set, that is, $I \in \mathcal{I}$, then, the graph (V, I) will have $|V| - |I|$ connected components. Actually, assuming each vertex is itself a connected component, given vertices V, when an edge is added between two vertices, then one connected component is lost. That is, every time when an edge is drawn to connect two connected components, then, the new graph lose one connected component compared to the original graph. The process ends until a graph (V, I) is formed.

Also noticed that when adding an edge inside a connected component, the new graph will have a cycle, which is not allowed because the edge set I of the graph (V, I) is an independent set, and then no cycle is allowed when adding an edge starting from single vertices.

Therefore, for any independent set $I \in \mathcal{I}$, graph (V, I) has $|V| - |I|$ connected components. then, start check axiom (iii) in the following proof:

Let $I, J \in \mathcal{I}$, $|I| < |J|$.

Suppose that for every edge $j \in J - I$, $(I \cup j) \notin \mathcal{I}$, that means, for every edge $j \in J - I$, when it is added into the graph (V, I) , it forms a cycle in (V, I).

When adding an edge, it forms a cycle, then, no two connected components are connected, instead, only two vertices that were already connected is connected.

So, graph $(V, I \cup J)$ also has $|V| - |I|$ connected components, and the graph $(V, I \cup J)$ has no new connected component compared with the original graph (V, I).

From the graph, remove the edges that are only in I (some edges can be both in I and J), then, the new graph is a graph (V, J), as when disconnecting edges of a graph, the number of connected components will be increased, so,

the number of connected components in the graph $(V, J) \geq |V| - |I|$

But, because J is an independent set, the number of connected components for the graph (V, J) equal to $|V| - |I|$, so,

$$|V| - |J| \geq |V| - |I|$$

then,
$$|J| \leq |I|$$

It contradicts the given condition of the axiom (iii) $|J| \geq |I|$.

3. Matching Matroid

Definition 83 *[55, Definition 1.6.] A graph is called bipartite graph iff it satisfies the following conditions,*

(i) Vertices are divided into two disjoint sets L and R.

(ii) Every edge connects a vertex in L and another vertex in R.

Proposition 84 *[54, Page 4]*
Let G be a bipartite graph with bipartition (L, R).

Let $\mathcal{I}$ be the collection of subsets of R(right side) which can be matched to L. Then, $(R, \mathcal{I})$ is a matroid.

Before proceeding the proof of the above proposition, several related definitions are stated as follows.

Definition 85 *[56]*
A subgraph $G' = (V', E')$, of a graph $G = (V, E)$ is a graph where its vertex set V' is the subset of V, $V' \subseteq V$ and its edge set E' is the subset of E, $E' \subseteq E$, which satisfies if an edge $(v_1, v_2) \in E'$, then, $v_1 \in V'$ and $v_2 \in V'$.

Definition 86 *[56]*
Given a graph $G = (V, E)$, an induced (generated) subgraph is a subset S of the vertices of the graph together with all the edges of the graph between the vertices of this subset. Denoted by $G[S] = (S, E')$, where $E' = \{(v_1, v_2) : (v_1, v_2) \in E \text{ and } \{v_1 \in S \wedge v_2 \in S\}\}$.

Definition 87 *[55, Definition 1.1.]*
A matching of graph G is a set of edges such that every edge shares no vertex with any other edge. That is, each vertex in matching M has degree one.

Note: a matching is one-to-one correspondence, i.e. a matching is a bijection.

Definition 88 *[55, Definition 1.5.]*
A matching of a given graph G is complete iff it contains all of G's vertices. Sometimes this is also called a perfect matching.

Definition 89 *[55, Definition 1.2.]*
The size of a matching is the number of edges in that matching.

Definition 90 *[55, Definition 1.3.]*
A matching is maximum when it has the largest possible size.

Note: for a given graph G, there may be several maximum matchings.

Example 9.1.4 *Example that is not a matroid: [49, Section 4.1]*

Let a graph $G = (V, E)$, with $V = \{1, 2, 3, 4\}$ and $E = \{(1, 2), (2, 3), (3, 4)\}$. Let the ground set be E and let the collection of independent set is defined as $\mathcal{I} = \{F \subseteq E : F \text{ is a matching}\}$,
Then, $(E, \mathcal{I})$ is not a matroid.

Proof:

Motivated from the discussion on the page 2 in [49, Section 4.1], we gave more detailed proof.

Consider an example that not satisfying the axioms of the definition of a matroid. Let $X = \{(2, 3)\}$, $Y = \{(1, 2), (3, 4)\}$, Obviously, X, Y both are matching, because they are a set of pairwise disjoint edges. That is, X, $Y \in \mathcal{I}$. and $|X| < |Y|$. However, X cannot be extended to a matching by adding an element $e \in (Y - X)$. so, the axiom(iii) of the definition of a matroid is not satisfied.

Based on the above example that is not a matroid, using the same graph in the example, if change the definition of the ground set from an edge set to a vertex set, and change definition of a collection of independent sets based on the subsets of the edge set to based on the subsets of the vertex set, a matching matroid for a given general graph can be found from the graph.

Definition 91 *[57] For a given graph, Let a matching formed from the graph is denoted by M. A vertex v is covered by a matching M iff there exists an edge $e \in M$, such that v is one endpoint of the edge e.*

Proposition 92 *[47, Example 6]*
Let $G = (V, E)$ be a graph with a vertex set V and an edge set E, Let the ground set be V and let the collection of independent set is defined as $\mathcal{I} = \{U \subseteq V : \exists$ a matching that covers all vertices of $U\}$.
Then, $(V, \mathcal{I})$ is a matroid.

Note: for a given graph $G = (V, E)$, if there is an independent set U, it must exist a matching that covers all vertices in U, but the matching does not precisely cover U, it might also covers other vertices, because one vertex is covered by an edge, only need it is one endpoints of that edge.

The proof is motivated from [47, Example 6], in which we gave more clear explanation of the check of axiom (iii) in (I3).

Proof:

(I1) check axiom (i) of definition of a matorid; for empty set of vertices $\emptyset$, as it can be covered by the empty matching, that is, $\emptyset \in \mathcal{I}$.

(I2) check axiom (ii) If $I \subset J$ and $J \in \mathcal{I}$, because $J \in \mathcal{I}$ and then there must exist a matching M that covers all vertices in J. Since the vertices set I is the subset of vertex set J, so, I must also be covered by the matching M. That is, $I \in \mathcal{I}$.

(I3) it is left to check axiom (iii). That is,

Let $J \in \mathcal{I}$, $I \in \mathcal{I}$, with $|J| > |I|$,

Then, let M_J is a matching that covers the vertex set J, M_I is a matching that covers the vertex set I,

1. If there exists an element $a \in (J - I)$ that is already covered by an edge in the matching M_I, then, $I \cup \{a\} \in \mathcal{I}$, since it can be covered by a matching M_I.

2. If there exists an element $a \in (J - I)$ that is covered by an edge $e \notin M_I$, but $e \in M_J$; since any element in $(J - I)$ must be either covered by M_I or covered by M_J.

   If for the edge $e \in M_J$, its other endpoint b is not in the vertex set I, denote the edge e by (a, b), that is, $a \notin I$, $b \notin I$.

   Then, since $I \cup \{a\}$ can be covered by a matching $M_I \cup \{(a, b)\} = M_I \cup \{e\}$, so, $I \cup \{a\} \in \mathcal{I}$.

3. If $\exists\, a \in (J - I)$ that is covered by an edge $e \in M_J$, but $e \notin M_I$.

   If for the edge $e \in M_J$, its other endpoint b is in the vertex set I, denote the edge e by (a, b).

   Since $b \in I$, but $e = (a, b) \notin M_I$, b must be covered by an edge in the matching M_I, denoted by b is covered by an edge $f \in M_I$.

   (31.) If for the edge $f \in M_I$, its other endpoint c is not in the vertex set I, denote the edge f by (b, c).
   That is, $f = (b, c) \in M_I$, and $c \notin I$.

   Then, $I \cup \{a\}$ can be covered by a matching $(M_I - \{f\}) \cup \{e\}$, so, $I \cup \{a\} \in \mathcal{I}$.

   (32.) If for the edge $f \in M_I$, its other endpoint c is in the vertex set I, denote the edge f by (b, c).
   That is, $f = (b, c) \in M_I$, and $c \in I$.

   It is known that $b, c \in I$, $f = (b, c) \in M_I$, $a \in (J - I)$.

Since $|J| > |I|$, there must exists two more vertices that belong to the vertex set J.

Without loss of generality, consider the simplest graph case, Let J contains two more vertices apart from the vertex a. Let $J = \{a, b, c\}$,

since $c \in J$, $e = (a, b) \in M_J$, c must be covered by an edge in the matching M_J that shares no vertex with other edges in M_J, that is, (b, c) must not in the matching M_J.

So, c must be covered by another edge $g \in M_J$. Denote the edge's other endpoint by d, that is, $g = (c, d) \in M_J$.

It is known that $f = (b, c) \in M_I$, by the definition of a matching, we know $e = (a, b) \notin M_I$, $g = (c, d) \notin M_I$,

Then, $I \cup \{a\}$ can be covered by a matching $(M_I - \{f\}) \cup \{e, g\}$, so, $I \cup \{a\} \in \mathcal{I}$.

In the above definitions and proposition related to matching matroid formed from a given general graph is discussed, the following will give the theory regarding the matroid that is constructed from a given bipartite graph.

Definition 93 *[58]*
For a given bipartite graph $G = (L, R, E)$, a complete matching is a matching which has the same number of edges as vertices down one side.

Definition 94 *[59]*
For a given bipartite graph $G = (L, R, E)$, a maximum matching is a matching which has the maximal numbers of edges based on the given bipartite graph.

Motivated from the definition 93 of complete matching and the definition 94 of maximum matching, we gave a simple example to understand the difference.

Example 9.1.5 *(Maximal matching VS Complete matching): Let a bipartite graph $G = (L, R, E)$, with $L = \{1, 2, 3\}$, $R = \{A, B, C\}$,*

$E = \{(1, A), (1, C), (2, B), (3, B)\}$, find a maximal matching and a complete matching.

Answer: from the given bipartite graph, a maximum matching can be formed as $\{(1, A), (2, B)\}$, or be formed as $\{(1, C), (3, B)\}$. But there is no complete matching, since the right hand side elements A, C only have relation with one element 1 in the left hand side.

Definition 95 *[59]*
For a given bipartite graph $G = (L, R, E)$, let a matching be formed from the given graph, denoted M, an alternating path with respect to M is defined as a path that alternates between edges in M, and $E - M$, which starts with an unmatched edge and ends with an unmatched edge.

Note: Let a given graph be $G = (L, R, E)$, let a matching be formed from the given graph, denote the matching by M, an edge $e \in E$, is called an unmatched edge with respect to M iff $e \in E$, but $e \notin M$.

An edge $e \in E$, is called an matched edge with respect to M iff $e \in E$, and $e \in M$.

In the remarks, the algorithm of finding maximum matching is taken from [59], it follows by an example which is modified by us from the example of maximal matching VS complet matching.

Remarks:

1. For a given bipartite graph G, there may be several maximum matchings, but from possible edges in the given graph G, a complete matching cannot always be found.

2 A maximal matching can always be found using the maximum matching algorithm. The algorithm can be separated into four step: firstly, give an initial matching M, that can be formed from the possible edges in the given graph $G = (L, R, E)$, Secondly, find an alternating path with respect to the given initial matching M, that is, find an alternating path that starts from an edge $e_1 \in E$ but $e_1 \notin M$, ends with an edge $e_2 \in E$ but $e_2 \notin M$, which alternates between edges in M, and $E - M$. Thirdly, from the beginning edge e_1, change unmatched edge to matched edge, and matched edge to unmatched edge until the end edge e_2. then, a new matching is formed by grouping the new matched edges. Lastly, repeat the second and third steps until a maximal matching is formed.

Note: Each time the algorithm is implemented, the number of matches are increased by one, the procedure ends until a maximal matching is formed from the possible edges in the given graph.

Example 9.1.6 *Let a bipartite graph $G = (L, R, E)$, with $L = \{1, 2, 3\}$, $R = \{A, B, C\}$, $E = \{(1, A), (1, B), (1, C), (2, B), (3, B)\}$, find a maximal matching.*

Answer:

1st, *Find an initial matching M, which contains one edge that have two endpoints which connects another two unmatched edges in the given graph G, Let the initial matching is $M = \{(1, B)\}$.*

2nd, *Let $1 = B$ denotes $(1, B)$ is matched edge with respect to M, let $1 - A$, $B - 2$ denotes $(1, A)$, $(B, 2)$ are unmatched edges with respect to M. An alternating path with respect to the initial matching is given as $A - 1 = B - 2$.*

3rd, *After changing the status of matched and unmatched edges, a new matching $M = (1, A), (2, B)$ can be get from a path $A = 1 - B = 2$.*

4th, *The matching $M = (1, A), (2, B)$ is the maximal matching that can be formed from the given graph G.*

Back to the Proposition 84, recall definition of a matching matroid formed from a given bipartite graph. The collection of independent sets is defined as $\mathcal{I} = \{T \subseteq R :$ $\exists$ a matching that covers all vertices of T $\}$. It is known that a matching is a one-to-one correspondence, and it is a bijection, rewrite the Proposition 84 as follows.

Motivated from the definition of transversal matroid [54], the discussion of relationship between bipartitie graph and set system and definition of transversal matroid [61, Section 2.1], we gave the proposition 96. The proof of the proposition 96 is given by us following the definition of the matroid.

Proposition 96 *Let G be a bipartite graph with bipartition (L, R). Let the ground set be R and let the collection of independent sets is defined as $\mathcal{I} = \{T \subseteq R :$ for some $K \subseteq L$, $\exists$ a bijection from the set K to the set T $\}$.*
Then, $(R, \mathcal{I})$ is a matroid.

Proof:

(I1) check axiom (i) of definition of matorid. For empty set $\emptyset \subseteq R$, $\exists \emptyset \subseteq L$, such that $\emptyset \to \emptyset$ is a bijection. So, $\emptyset \in \mathcal{I}$.

(I2) check axiom (ii) If $A \subset B$ and $B \in \mathcal{I}$, because $B \in \mathcal{I}$ and then $B \subseteq R$, $\exists K \subseteq L$, such that $K \to B$ is a bijection.

Since $A \subset B$, then $A \subseteq R$, there must exist a bijection $K' \subset K \subseteq L$, such that $K' \to A$ is a bijection.

(I3) it is left to check axiom (iii). That is,

Let $A \in \mathcal{I}$, $B \in \mathcal{I}$, with $|A| < |B|$.

It is known that $A \in \mathcal{I}$, $B \in \mathcal{I}$, then,

$$A \in \mathcal{I} \implies A \in R, \exists K_1 \subseteq L, \text{ such that } K_1 \to A \text{ is a bijection.}$$
$$B \in \mathcal{I} \implies B \in R, \exists K_2 \subseteq L, \text{ such that } K_2 \to B \text{ is a bijection.}$$

Since $|A| < |B|$, there must exists an element $b \in (B - A)$, which has a unique pairing element k in $(K_2 - K_1)$, denoted by an edge (k, b).

Therefore, there must exists an element $b \in (B - A)$, such that $A \cup \{b\} \in \mathcal{I}$, since $A \cup \{b\} \in R$, $\exists (K_1 \cup \{k\}) \subseteq L$, such that $(K_1 \cup \{k\}) \to A \cup \{b\}$ is a bijection.

Following the proposition 96, we develop a construction algorithm of bipartite graph from any given matroid and explain the algorithm using a small example as follows.

Example 9.1.7 *Result:*

1. *For any given matroid $M = (V, \mathcal{I})$, a bipartite matroid can be constructed in this way: take $L = R = V$, connect all pairs from the left side vertex to the right side vertex.*

2. *An example is given as follows regarding relationship between a given matroid and a bipartite matroid constructed from the given matroid.*

   Example: Let a matroid is given as $M = (V, \mathcal{I})$, with the ground set $V = \{1, 2, 3\}$, the independent set is given as

   $\mathcal{I} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}\}$.   *Then, the matroid can be constructed as bipartite matroid using the following way:*

   1st, Take $L = R = V$ as the vertices of a bipartite graph.

   2nd, In the following, denote (a, b) be an edge in the bipartite graph where vertex a is from the left side vertex set L , the vertex b is from the right side vertex set R.

   Take the maximum independent sets from the given matroid, that is, $\{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$,

   since $\{1, 2, 3\}$ is not an independent set, so, when connecting the pairs, make one element in the left side vertex set L not be connected to the right side set R. For instance, choose an left side element $2 \in L$ not be connected to an edge,

   then, the pairs between the left side and right side can be connected in the way that

   $\{1, 2\} \iff edge{:}(\{(1, 2), (2, 1)\} - \{(2, 1)\}),\ that\ is,\ connect\ edge\ (1, 2).$

   $\{1, 3\} \iff edge{:}(\{(1, 3), (3, 1)\}\}),\ that\ is,\ connect\ edge\ (3, 1), (1, 3).$

   $\{2, 3\} \iff edge{:}(\{(3, 2), (2, 3)\} - \{(2, 3)\}),\ that\ is,\ connect\ edge\ (3, 2).$

   Then, check the correspondence edges for the maximum independent set in the bipartite matching matroid $M = (R, \mathcal{I})$, with $\mathcal{I} = \{T \subseteq R : \exists$ a matching that covers all vertices of $T\}$.

   3rd, Check: A matching in the constructed bipartite graph can be found for any maximum independent set in the given matroid. that is,

   $$\{1, 2\} \iff edge{:}\quad \{(3, 1), (1, 2)\}$$
   $$\{1, 3\} \iff edge{:}\quad \{(3, 1), (1, 3)\}$$
   $$\{2, 3\} \iff edge{:}\quad \{(3, 2), (1, 3)\}$$

3. *in the above example, to generalize the method to construct a bipartite graph from any given matroid, the method is that, given a matroid $M = (E, I)$, firstly, let the left vertices and right vertices equal to the ground set E . Secondly, starting from each maximum independent set(basis) of the given matroid by finding the basis from the set I , construct the edges like the small example. Lastly, one bipartite graph can be constructed and checked using the definition of the matching matroid like checking the small example.*

9.2 Lattice path matroid

In this section, the definition of transversal, full transversal and partial transversal will be introduced, which will be useful to talk the transversal matroid and its application to count number of lattice paths.

Definition 97 *[61, Section 2.1]*
A multiset is a set for which repeated elements are considered. That is, the same element may appear multiple times in the multset and are treated as different elements in the multset.

Note: a set also allows repeated elements, but the repeated elements are treated as one element in the set.

For example, $\{a, a, b\}$ and $\{a, b\}$ are different multisets, but they are the same set.

Definition 98 *[61, Section 2.1]*
A set system $(S, \mathcal{A})$ is a finite set S along with a finite multiset $\mathcal{A}$ of subsets of S. Let J be an index set corresponding to all elements in the multiset $\mathcal{A}$, assume $|J| = r$, then, let $J = [r] = \{1, 2, \ldots, r\}$, $\mathcal{A}$ can also be denoted by a sequence $(A_j : j \in J) = (A_1, A_2, \ldots, A_r)$, where the same set in $\mathcal{A}$ are indexed by different elements in J.

Remark: [61, Section 2.1]

1. The different possible orders in which the sets are listed in the multiset $\mathcal{A}$ is not distinguished. That is, for any permutation $\sigma = (\sigma(1), \sigma(2), \ldots, \sigma(r))$ of the index set $[r] = \{1, 2, \ldots, r\}$,
   $(A_1, A_2, \ldots, A_r) = (A_{\sigma(1)}, A_{\sigma(2)}, \ldots, A_{\sigma(r)})$

2. A set system $(S, \mathcal{A})$ can be represented by a bipartite graph (L, R, E), where the vertex sets $L = \mathcal{A}$, $R = S$. In the edge set E, an edge connects an element(set) $A_j \in \mathcal{A}$ and an element $x \in S$ precisely when $x \in A_j$. Similarly, each bipartite graph (L, R, E) can be seen as representing a set system $(S, \mathcal{A})$, where $S = R$, $\mathcal{A} = L$ is a multiset of subsets of R.

Definition 99 *[61, Section 2.1]*
Let $(S, \mathcal{A})$ be a set system, a multiset $\mathcal{A}$ of subsets of S is denoted by $\mathcal{A} = (A_j : j \in J)$ having the size $|J| = r$, denoted by $\mathcal{A} = (A_1, A_2, \ldots, A_r)$.
A transversal of the set system is defined as a subset T of S, such that there exists a bijection from the index set $J = [r] = \{1, 2, \ldots, r\}$ to the set T, $\phi : J \to T$, with $\phi(j) \in A_j$, for all $j \in J$. That is, $T \subseteq S$ can be denoted by $T = \{a_1, a_2, \ldots, a_r\}$, where $a_i \in A_i$, and the set T has r distinct elements.

Motivated from the discussion on page 2-3 in [60, Section 2.1], we gave the remarks as follows.

Remark:

1. If represent the set system by a bipartite graph (L, R, E) with $L = \mathcal{A}$, $R = S$, an edge in E connects an element in R and an element(set) in L, then, transversals of the bipartite graph can be defined as

   $\{T \subseteq R : \exists$ a bijection from the set L to the set T $\}$,

2. A transversal is also called a full transversal, since it is a subset T of right-hand side elements R, that could be matched by all the left-hand side elements, that is, transversals is
   $\{T \subseteq R : \exists$ a matching that covers all vertices of L and T $\}$,

3. Not all set systems have transversals.

   For example: Let a set system $(S, \mathcal{A})$ is given by $S = \{1, 2, 3, 4\}, \mathcal{A} = \{\{1\}, \{2\}, \{1, 2\}, \{3, 4\}\}$, this set system has no transversals, since the first three of the sets contain only two elements and three different elements cannot be taken from three sets.

   A related theorem without proof is given as follows.

Theorem 100 *[61, Theorem 2.1.]*

A finite set system $(S, (A_j : j \in J))$ has a transversal if and only if $\left| \bigcup_{i \in K} A_i \geq K \right|$, for all $K \subseteq J$.

Remark: It is obvious that to make a set system have a transversal, a set containing $|K|$ different elements must be taken from $(A_i : i \in K)$, for all $K \subseteq J$, which must have at least as many elements as number of sets $|K|$, the proof of sufficient condition for the theorem did not given.

Based on the definition of transversal, the definition of partial transversal is given as follows.

Definition 101 *[61, Page 3]*
Let $(S, \mathcal{A})$ be a set system, a multiset $\mathcal{A}$ of subsets of S is denoted by $\mathcal{A} = (A_j : j \in J)$ having the size $|J| = r$, denoted by $\mathcal{A} = (A_1, A_2, \ldots, A_r)$.
A partial transversal of the set system is defined as a subset T of S, such that, for some $K \subseteq J$, there exists a bijection from the index subset set K to the set T. Denoted by $\phi : K \to T$, with $\phi(i) \in A_i$, for all $i \in K$. That is, for some $K \subseteq J$, $T \subseteq S$ can be denoted by $T = \{a_1, a_2, \ldots, a_{|K|}\}$, where $a_i \in A_i$, and the set T has $|K|$ distinct elements.

Remark:

1. If represent the set system by a bipartite graph (L, R, E) with $L = \mathcal{A}$, $R = S$, an edge in E connects an element in R and an element(set) in L, then, partial transversals of the bipartite graph can be defined as
   $\{T \subseteq R : $ for some $K \subseteq L$, $\exists$ a bijection from the set K to the set T $\}$.

2. A partial transversal of a set system $(S, \mathcal{A})$ with $\mathcal{A} = (A_j : j \in J)$ is a transversal of some subsystem $(S, \mathcal{A}')$, where $\mathcal{A}' = (A_k : k \in K)$, with $K \subseteq J$.

   3. A partial transversal is a subset T of right-hand side elements R, that could be matched by a subsets of left-hand side elements, that is, partial transversals is
   $\{T \subseteq R : \exists$ a matching that covers all vertices of T $\}$.

From the definition of partial transversal in above remarks(1) and recall the Proposition 96 which gives a construction of a matching matroid, the partial transversals of a given bipartite graph is defined same as the collection of independent sets in the Proposition 96, that is,
$\mathcal{I} = \{T \subseteq R :$ for some $K \subseteq L$, $\exists$ a bijection from the set K to the set T $\}$.

Recall another definition of a matching matroid in the Proposition 84, where the collection of independent sets are given by
$\mathcal{I} = \{T \subseteq R : \exists$ a matching that covers all vertices of T $\}$, which is also the same as definition of partial transversal in above remarks(3).

Then, the definition 102 of transversal matroid is given as follows, which is motivated from the discussion on page 2-4 in [61, Section 2.1].

Definition 102 *Let $J = [\|J\|] = \{1, 2, \ldots, |J|\}$ be an index set of size $|J|$,*
Let $(S, \mathcal{A})$ be a set system, where S is a finite set, $\mathcal{A} = (A_j : j \in J)$ is a multiset of subsets of the finite set S. Let the ground set be S, let the collection of independent sets is defined be
$\mathcal{I} = \{partial\ transversals\ of\ the\ set\ system\ (S, \mathcal{A})\}$, that is, $\mathcal{I} = \{T \subseteq S :$ for some $K \subseteq J$, $\exists$ a bijection $\phi : K \to T$, with $\phi(i) \in A_i$ $\}$,
Then, by the proposition 96, $(S, \mathcal{I})$ is a matroid. Such matroid is called a transversal matroid. The multiset $\mathcal{A}$ is called a presentation of the transversal matroid.

Remarks:

   1. From above definition, partial transversals of a set system $(S, \mathcal{A})$ are the independent sets of a transversal matroid with the ground set S. From the definition of transversal, it is a subset of S that can be matched by all the left-hand side elements(sets) of $\mathcal{A}$, so, it is a maximal independent set. that is, a transversal is a basis of a transversal matroid. A collection of partial transversals can be obtained by taking all the subsets of the transversals.

Then, definition of a lattice path matroid is given as follows, which is a transversal matroid.

Definition 103 *[61, Section 4.1.]*
A lattice path is a sequence of east(right) and north(up) steps of unit length, which starts from $(0, 0)$. Denote north step by N, denote east step by E, a lattice path can be written as a string of E's and N's.

Instead of an individual lattice path, a collection of lattice path that stay in a given region in a plane is interested.

Lemma 104 *[61, Page 17-18]*
Let P, Q be two lattice paths from $(0,0)$ to (m,r) with the path P never going above Q (they may meet at points except the ends).
Fix the two lattice paths P and Q, then let $\mathcal{P}$ be the set of lattice paths from $(0,0)$ to (m,r) that stay in the region that P and Q bound.
Each lattice path in $\mathcal{P}$ is a sequence of E's and N's which has $m+r$ steps containing m East steps and r North steps. The positions for the $m+r$ steps are denoted by $1, 2, \ldots, m+r$, For each lattice path in $\mathcal{P}$, the length of the path is fixed and is $m+r$, if the positions of north steps is confirmed, then the other positions are east steps, the lattice path is identified. Conversely, given a lattice path, the set of positions of north steps in the lattice path can be taken.
Then, each lattice path in $\mathcal{P}$ is determined by its north steps. each lattice path in $\mathcal{P}$ can be represented by the set of positions of north steps in the lattice path.

Next, the example of lattice path representation and the following remarks are motivated from the paper by Bonin [60, Definition 3.1.] and [61, Page 17-18].

Example(Lattice path representation):

Let P be a lattice path from $(0,0)$ to (m,r) which is a lower bound for a given region in a plane, the lattice path is a sequence of E's and N's which has $m+r$ steps. Denote the lattice path P by $p_1 p_2 \ldots p_{m+r}$, **that is,** $P = p_1 p_2 \ldots p_{m+r}$. The lattice path in P can be determined by the set of its north steps, the number of north steps in the lattice path P is r, that is, $P = \{p_{u_1}, p_{u_2}, \ldots, p_{u_r}\}$, where $\{u_1, u_2, \ldots, u_r\}$ is the set of positions of north steps in the lattice path, and u_i, is the position for the *ith* north step in the lattice path P, for $1 \le i \le r$.

For the lattice path P from $(0,0)$ to (m,r), since it has $m+r$ steps and the positions for the $m+r$ steps are denoted in natural order from left to right by $1, 2, \ldots, m+r$, if take the set of positions of north steps in the lattice path, $\{u_1, u_2, \ldots, u_r\}$, then $u_1 < u_2 < \ldots < u_r$ and P can also represented by $\{u_1, u_2, \ldots, u_r\}$, that is, $P = \{u_1, u_2, \ldots, u_r\}$.

Remarks:

1. In the above example, the symbol equal $'='$ means 'represented by'.

2. In the above lemma, P, Q be two lattice paths from $(0,0)$ to (m,r) with the path P never going above Q (they may meet at points except the ends), fix the two lattice paths P and Q, then, the region that P and Q bound is given in the plane.

   In the given region, let $\mathcal{P}$ be the set of lattice paths from $(0,0)$ to (m,r) that stay in the region. Different lattice paths in the region are different sequences of E's and N's. The given region is formed by many north steps and many east steps of many different lattice paths.

   For each lattice path, in the given formed region(formed by E's and N's), label the path's north steps by the positions of north steps in the path.

It is obvious that the labels in the given region are the positions of north steps of all possible lattice paths, which is chosen from $1, 2, \ldots, m + r$. For each lattice path in $\mathcal{P}$, the labels for its steps are denoted by $[m + r] = \{1, 2, \ldots, m + r\}$,

In the above lemma, let $P = p_1 p_2 \ldots p_{m+r}$, $Q = q_1 q_2 \ldots q_{m+r}$ is two bounding lattice path for the given region, where P is the lower bound path, it can be written as

$$P \overset{denoted\ by}{=} \{p_{b_1}, p_{b_2}, \ldots, p_{b_r}\} \overset{denoted\ by}{=} \{b_1, b_2, \ldots, b_r\},$$

where $b_1 < b_2 < \ldots < b_r$ is the set of positions of north steps in the lower bound lattice path, and b_i is the position for the ith north step in the lower bound path P, for $1 \leq i \leq r$. Similarly, Q is the upper bound path, it can be written as

$$Q \overset{denoted\ by}{=} \{q_{a_1}, q_{a_2}, \ldots, p_{a_r}\} \overset{denoted\ by}{=} \{a_1, a_2, \ldots, a_r\},$$

where $a_1 < a_2 < \ldots < a_r$ is the set of positions of north steps in the upper bound lattice path. and a_i is the position for the ith north step in the upper bound path P, for $1 \leq i \leq r$.

The positions of north steps of all possible lattice paths from $(0, 0)$ to (m, r) are labelled in the given bound region from bottom to up, from left to right, therefore, the position for the ith north step in the upper bound path must less than or equal to the label for the the ith north step in the lower bound path, that is, $a_i \leq b_i$, for $1 \leq i \leq r$. $a_i = b_i$ happens when the two bounding paths P and Q meet at points except the ends $(0, 0)$ and (m, r).

Let the position of ith north steps of a possible lattice paths that stay in the region that P and Q bound, denoted by s_i, for $1 \leq i \leq r$.

Let

$$\begin{aligned}
N_i = \ & \{ \text{ labels of steps that can be } ith \text{ north step} \\
& \text{ in a lattice path that stay in a given region } \} \\
= \ & \{ s \in [m + r] : s \text{ is the } ith \text{ north step in a lattice path in } \mathcal{P} \} \\
= \ & \text{The integer inverval } [a_i, b_i], \text{ for } 1 \leq i \leq r \\
& \text{where, } r = \text{ no of north steps for any lattice path from } (0, 0) \text{ } to \text{ } (m, r)
\end{aligned}$$

Motivated from the definition of lattice path matroid by Bonin in the paper [60, Definition 3.1.] and definition of transversal matroid on page 2 in [61, Section 2.1], we gave a more constructive lattice path matroid definition 105 as follows.

Definition 105 *Consider a lattice region from* $(0, 0)$ *to* (m, r)*, a lattice path matroid is characterized by two bounding lattice paths from* $(0, 0)$ *to* (m, r)*, say,* (P, Q) *with* P *never going above* Q *(they may meet at points except ends).* P, Q *can be represented by two sets of north steps respectively, and can be represented by two sets of positions(index) of north steps of the two lattice paths. Let*

$$P \overset{denoted\ by}{=} \{p_{b_1}, p_{b_2}, \ldots, p_{b_r}\} \overset{denoted\ by}{=} \{b_1, b_2, \ldots, b_r\} \text{ with } b_1 \leq b_2 \leq \ldots \leq b_r,$$

$$Q \overset{denoted\ by}{=} \{q_{a_1}, q_{a_2}, \ldots, p_{a_r}\} \overset{denoted\ by}{=} \{a_1, a_2, \ldots, a_r\} \text{ with } a_1 \leq a_2 \leq \ldots \leq a_r.$$

Let

$$N_i = \text{The integer interval } [a_i, b_i], \text{ for } 1 \leq i \leq r$$

The lattice path matroid $M = (E(M), \mathcal{I}(M)) = (E, \mathcal{I})$ or $M = (E(M), \mathcal{B}(M)) = (E, \mathcal{B})$ is defined to be isormorphic to transversal matroid and is constructed from a set system which is defined as

$$(S, \mathcal{A}) = (E(M), (\ N_j : 1 \leq j \leq r\ ))$$

with $E(M)$, $\mathcal{I}(M)$, $\mathcal{B}(M)$ defined as

$$E(M) \;=\; \text{a linearly ordered set } = [1, m+r] = \{1, 2, 3, \ldots, m+r\}$$

$$\mathcal{I}(M) \;=\; \{X \subseteq E : X \text{ is a partial transversal of the set system } (E, (\ N_j : 1 \leq j \leq r\ ))\}$$

$$\mathcal{B}(M) \;=\; \{B \subseteq E : B \text{ is a transversal of the set system } (E, (\ N_j : 1 \leq j \leq r\ ))\}$$

9.3   Tutte polynomials

Definition 106 *(Rank in matroid)[62, Page 20] [63, Page 10]*
Let $M = (E, \mathcal{I})$ be a matroid; $E = E(M)$ is the ground set of the matroid, $\mathcal{I} = \mathcal{I}(\mathcal{M})$ is the collection of independent subsets of E.

Let a subset $A \subseteq E$,

Then, the rank of the set A is defined to be the size of a maximal independent subset of A,

the rank of the matroid M is defined to be the rank of the ground set E the size of maximal independent subset of E is the size of a basis in the matroid. That is,

$$r(M) = r(E) = a \text{ size of a basis set}$$

Let $\mathcal{B} = \mathcal{B}(M) =$ the set of maximal independt subsets of E(M) $=$ the set of bases of E

By the definition of a matroid, the subset of each independent set is an independent set, each independent subset must be a subset of one basis of the matroid; the independent subset of the set A is $(A \cap B)$, for some $B \in \mathcal{B}(M)$.

Thus, the rank of the set A can be defined as

$$r(A) = \max_{B \in \mathcal{B}(M)} |A \cap B|$$

The idea in the following example is motivated from the discussion of uniform matroid in the paper [64, Section 2.1].

Example 9.3.1 *Consider a uniform matroid $M = (E, \mathcal{I})$ with $E =$ the ground set.*
For a given $k \in \mathbb{N}$, the collection of independent sets is

$$I = \{X \subseteq E : |X| \leq k\}.$$

By the definition of the rank in the matroid, for any $X \subseteq E$,

$$r(X) = rank(X) = \text{the size of a maximal independent subset of } X$$

for the uniform matroid,
if $|X| \leq k$, by definition of the uniform matroid, X is an independent subset, then, $r(X) = |X|$.
If $|X| > k$,
then, X is not an independent set, the maximal independent subset of X is a subset of size k, so, $r(X) = k$.
Thus, suppose the size of the ground set E is a set of size n; $|E| = n$, given an integer $k \in \mathbb{N}$ to define the uniform matroid, then, the matroid has rank k,
Thus, the uniform matroid can also be denoted by $U_{k,n}$.

Remarks: Consider a subset $A \in E(M)$, it is noted that $0 \leq rank(A) \leq |A|$

If $A \subseteq B \subseteq E(M)$, then, $rank(A) \leq rank(B)$, since the maximal independent subset of B can be expanded by the maximal independent subset of A.

Definition 107 *(Whitney) [65, Section 1.1]*
Consider a subset $A \in E(M)$, the nullity of the set A is defined to be

$$n(A) = nullity(A) = |A| - rank(A) = |A| - r(A)$$

Remarks: Consider a subset $A \in E(M)$, it is noted that $0 \leq n(A) \leq |A|$.

If $A \subseteq B \subseteq E(M)$, then, $n(A) \geq n(B)$.

Motivated from the discussion of rank-nullity in [65, Section 1.1], the lemma 108 and lemma 109 are given as follows following the definition of the independent set.

Lemma 108 *(Rank of independent sets)*
A subset $A \subseteq E$ is an independent set if and only if $rank(A) = |A|$, since the maximal independent subset of an independent set is itself.

Lemma 109 *(nullity of independent sets)*
A subset $A \subseteq E$ is an independent set if and only if $n(A) = nullity(A) = 0$

Definition 110 *(Tutte polynomial) [64, Definition 2.1.] [65, Definition 1.2.]*
Consider a matroid $M = (E(M), \mathcal{I}(M))$, $E(M)$ is the ground set of the matroid.
$\mathcal{I}(M)$ is the collection of independent sets of the matroid. $B(M) =$ the set of all
bases of the matroid $=$ the set of maximal independent subset of $E(M)$.

The Tutte polynomial of the matroid is defined as

$$T_M(x, y) = \sum_{A \subseteq E(M)} (x-1)^{r(E)-r(A)}(y-1)^{|A|-r(A)}$$

$$\text{where, } r(E) = \text{the rank of the ground set } E$$
$$= \text{the size of the maximal independent subset of } E$$
$$= \text{the size of a basis in the matroid}$$

9.4   Calculation of the Tutte polynomials for lattice path matroids

In this section, motivated from the definition 110 and the constructive definition 105 of lattice path matroid, we gave the methodology of the calculation of tutte polynomial from a given lattice path matroid.

Consider a lattice path matroid $M = (E(M), \mathcal{I}(M))$, it can be characterized by two bounding paths (ω, η). The pair paths have a same initial point $S(P) = (0, 0)$ and a same ending point $N(P) = (m + r)$.

Consider lattice paths which stay in the matroid and are formed of north steps and east steps, starting from the initial point $S(P) = (0, 0)$ to the ending point $N(P) = (m + r)$.

Methodology.

The methodology consists in

(i) Given a matroid $M = (E(M), \mathcal{I}(M))$, find $E(M)$, $\mathcal{I}(M)$ $\mathcal{B}(M)$;
(ii) From the definition, calculate the Tutte polynomial by separating subsets $A \subseteq E(M)$ into two parts, one part is subsets $A \in \mathcal{I}(M)$, another part is subsets $A \in E(M) - \mathcal{I}(M)$.
(iii) Before doing the calculation, consider the contribution to terms in the Tutte polynomial of independent subsets, minimal dependent subsets, the ground set $E(M)$ to simplify the later calculation of the Tutte polynomial of the given matroid.

9.4.1 Construction of bijection between lattice paths and set of their north steps

Applying the lemma 104 of relationship between lattice paths and their north steps which is motived from [61, Page 17-18], we gave a construction algorithm between the set of lattice paths and the set of their north steps as follows.

Given a lattice path matroid with initial point $S(P) = (0,0)$ and ending point $N(P) = (m, r)$

Consider a lattice path $\sigma = \sigma_1\sigma_2\ldots\sigma_{m+r}$ from $S(P)$ to $N(P)$, which stays in the region of the matroid.

Construct a north-step set of the path by the following bijection.

Firstly, going through the path $\sigma = \sigma_1\sigma_2\ldots\sigma_{m+r}$ from $i = 1$ to $i = m + r$,

if $\sigma_i =$ North step $= N$, then, put the index i into the set $N(\sigma)$.

Conversely, going through a north step set $N(\sigma)$,

construct a path $\sigma = \sigma_1\sigma_2\ldots\sigma_{m+r}$ by setting the i-th step $\sigma_i = N_i$, for $i \in N(\sigma)$.

Construct the remaining steps in the path by putting them to be East steps; that is, set $\sigma_j = E_i$, for $j \in [m + r] - N(\sigma)$.

9.4.2 The independent sets of a lattice path matroid

In this section, a construction of $E(M)$, $\mathcal{I}(M)$ $\mathcal{B}(M)$ is done according to the definition 105 of lattice path matroid we gave and developed from the paper [60, Definition 3.1.].

Consider a lattice path matroid $M = (E(M), \mathcal{I}(M))$,

Let (ω, η) be the bounding pair paths of the lattice path matroid.

It is known that the lattice path steps in the matroid are labelled from the bottom left to the top right. Each lattice path from $(0,0)$ to (m, r) consists in $r =$ number of north steps.

Since the bounding path $\eta =$ never goes above ω and η is the bottom right of ω, Let

$$\omega = \omega_1\omega_2\ldots\omega_{m+r} = \{\omega_{l_1}, \omega_{l_2}, \ldots, \omega_{l_r}\}$$

with $l_j =$ the j-th north step in the path ω, for $1 \le j \le r$.

$$\eta = \eta_1\eta_2\ldots\eta_{m+r} = \{\eta_{u_1}, \omega_{u_2}, \ldots, \omega_{u_r}\}$$

with $u_j = $ the j-th north step in the path η, for $1 \le j \le r$. It can be observed that the index integer $l_j \le u_j$, for $1 \le j \le r$.

Let

$$N_j = \text{the integer interval } [l_j, u_j] = [l_j, l_j + 1, \ldots u_j]$$

Recall the definition of transversals and partial transversals of the set system which was introduced in the Section 9.2.

Then, a transversal of the set system $\left(N_j : j \in [r] = \{1, 2, 3, \ldots r\} \right)$ is a set which is constructed by taking an element from each set N_j, for $j \in [r] = \{1, 2, 3, \ldots r\}$ and $r = $ the number of lattice path north steps in the matroid.

A partial transversal of the set system $\left(N_j : j \in [r] = \{1, 2, 3, \ldots r\} \right)$ is a set which is constructed by taking an element from some sets N_j, for $j \in [r] = \{1, 2, 3, \ldots r\}$ and $r = $ the number of lattice path north steps in the matroid.

Next is to define the set $E(M)$, $\mathcal{I}(M)$, $\mathcal{B}(M)$, for the given lattice path matroid which is characterized by the pair path (ω, η).

Firstly, define $E(M)$ be a linearly ordered set

$$E(M) = [1, m + r] = \{1, 2, 3, \ldots, m + r\}$$

Secondly, define $\mathcal{I}(M)$ as

$$\mathcal{I}(M) = \{X \subseteq E : X \text{ is a partial transversal of } \left(N_j : j \in [r] = \{1, 2, 3, \ldots r\} \right)\}$$

Lastly, define $\mathcal{B}(M)$ as

$$\begin{aligned} \mathcal{B}(M) &= \text{ the set of maximal independent subset of the ground set E} \\ &= \{B \subseteq E : B \text{ is a transversal of } \left(N_j : j \in [r] = \{1, 2, 3, \ldots r\} \right)\} \end{aligned}$$

9.4.3 The calculation of Tutte polynomials

In this section, motivated from the examples for calculating tutte polynomial which are stated on page 21 in [61, Section 5.1] and on page 1 in [66, Introduction], and inspired by the formula of calculating uniform matroid in [65, Section 3.2], we provide the analysis of the calculation of tutte polynomial according to the definition 110 of tutte polynomial.

Calculate Tutte polynomial using one definition of Tutte polynomial which is based on the rank of subset of the ground set $E(M)$.

Before doing the calculate, it can be observed that the subsets of the ground set can be divided into two parts, one is the independent subsets $A \in \mathcal{I}(M)$, the other is the non-independent subsets $A \in E(M) \setminus \mathcal{I}(M)$.

If a subset $A \in \mathcal{I}(M)$, it contributes a term without y, since

$$
\begin{aligned}
(x-1)^{r(E)-r(A)}(y-1)^{|A|-r(A)} &= (x-1)^{r(E)-|A|}(y-1)^{|A|-|A|} \\
&= (x-1)^{r(E)-|A|}
\end{aligned}
$$

If a subset $A \in E(M) \setminus \mathcal{I}(M)$ and it is a subset which has size greater than the basis size of the matroid,

then, it contributes a term without x, since by definition, $r(A) =$ the maximum independent subsets of A, if it has size greater than the basis size(maximum number of independent subsets of the matroid), then,

$$r(A) = \text{ the size of a basis.}$$

We also know that the maximum number of independent subsets of the ground set $E(M)$ is the basis size, that is,

$$r(A) = \text{ the size of a basis.}$$

Then, it contributes a term without x, since

$$
\begin{aligned}
(x-1)^{r(E)-r(A)}(y-1)^{|A|-r(A)} &= (x-1)^{0}(y-1)^{|A|-r(A)} \\
&= (y-1)^{r(E)-|A|}
\end{aligned}
$$

It can also be observed that if a subset $A \in E(M) \setminus \mathcal{I}(M)$ and it is a subset which has size greater than the basis size of the matroid,

then, $r(A) < |A|$.

To apply the calculation method discussed in the section, we give a smiple example to illustrate the calculation of tutte polynomial from a given lattice path matroid according to the definition 110.

Example 9.4.1 *Consider a lattice matroid given as follows,*
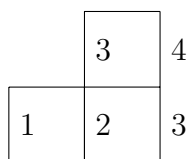

Figure 9.1: an example of lattice path matroid with labelled north steps indexes

Step 1,   Find the value of $E(M)$, $\mathcal{I}(M)$, $\mathcal{B}(M)$,

$$E(M) = 1, 2, 3, 4$$

$$
\begin{aligned}
\mathcal{B}(M) \;=\; &\{\{1,3\}, \{1,4\}, \{2,3\}, \\
&\{2,4\}, \{3,4\}\}
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{I}(M) \;=\; &\{X \subseteq E(M) : X \subseteq B \text{ for some } B \in \mathcal{B}(M)\}, \\
=\; &\{\emptyset, \{1\}, \{3\}, \{1,3\}, \\
&\{4\}, \{1,4\}, \\
&\{2\}, \{2,3\}, \\
&\{2,4\}, \\
&\{3,4\}\}
\end{aligned}
$$

Step 2,   Use the definition of the Tutte polynomial and separate the subsets of ground set $E(M)$ into independent subsets and non-independent subsets.

$$
\begin{aligned}
T_M(x,y) \;=\; & \sum_{A \subseteq E(M)} (x-1)^{r(E)-r(A)}(y-1)^{|A|-r(A)} \\
=\; & \sum_{A \subseteq E(M)} (x-1)^{r(E)-r(A)} + \sum_{A \subseteq E(M)\backslash \mathcal{I}(M)} (x-1)^{r(E)-r(A)}(y-1)^{|A|-r(A)}
\end{aligned}
$$

It can be observed that in this example, not all non-independent sets have a size greater than the basis size. If the non-independent subsets have a size greater than the basis size, it contributes terms without x.

The collection of non-independent sets is

$$
\begin{aligned}
E(M) \setminus \mathcal{I}(M) \;=\; &\{\{1,4\}, \{1,2\}, \{1,2,3\}, \\
&\{1,3,4\}, \{1,2,4\}, \{2,3,4\}, \\
&\{1,2,3,4\}\}
\end{aligned}
$$

Since a basis has size 2,   each three element subsets of $E(M)$ is a non-independent subset greater than the basis size, the three elements subsets in the set $E(M)\backslash \mathcal{I}(M)$ has $|A| = 3, r(A) = 2, r(E) = 2,$ in the second term of Tutte polynomial, they contribute to

$$\binom{4}{3} \cdot (y-1)^1$$

The ground set itself has $|E| = 4$, $r(E) = 2$ it contributes one term

$$(y-1)^2$$

The non-independent set $1, 2$ *is equal to the basis size. It has* $|A| = 1$, $r(E) = 2$, $|A| = 2$, *it contributes one term*

$$(x - 1)^1 (y - 1)^1$$

Then, the second term of the Tutte polynomial of the matroid is

$$\sum_{A \subseteq E(M) \backslash \mathcal{I}(M)} (x - 1)^{r(E) - r(A)} (y - 1)^{|A| - r(A)} = \binom{4}{3} \cdot (y - 1)^1 + (y - 1)^2 + (X - 1)^1 (y - 1)^1$$

$$= (y - 1)(x + y + 2)$$

Consider the independent subsets. In the first term of Tutte polynomial.

Since the basis set in $\mathcal{B}(M)$ *covers the ground set* $E(M) = \{1, 2, 3, 4\}$, *if the independent sets has size* $|A| = 1$, *they contribute to*

$$\binom{4}{1} \cdot (x - 1)^{2-1} = 4(x - 1)$$

Since the basis set in $\mathcal{B}(M)$ *covers all pair subsets of the ground set* $E(M) = \{1, 2, 3, 4\}$, *except* $\{1, 2\}$. *If the independent sets has size* $|A| = 2$, *they contribute to*

$$\left(\binom{4}{2} - 1 \right) \cdot (x - 1)^{2-2} = 5$$

The empty set is an independent set with size $|A| = 0$, *it constributes one term*

$$(x - 1)^2$$

Then, the first term of the Tutte polynomial of the matroid is

$$\sum_{A \subseteq E(M)} (x - 1)^{r(E) - r(A)} = (x - 1)^2 + 4(x - 1) + 5$$

Finally, the Tutte polynomial of the given lattice path matroid is

$$T_M(x, y) = \sum_{A \subseteq E(M)} (x - 1)^{r(E) - r(A)} (y - 1)^{|A| - r(A)}$$

$$= \sum_{A \subseteq E(M)} (x - 1)^{r(E) - r(A)} + \sum_{A \subseteq E(M) \backslash \mathcal{I}(M)} (x - 1)^{r(E) - r(A)} (y - 1)^{|A| - r(A)}$$

$$= (x - 1)^2 + 4(x - 1) + 5 + (y - 1)(x + y + 2)$$

9.5   Modelling of paths via matroids

The idea starts from a given connected lattice path matroid, and then find its tutte polynomial. The tutte polynomial can be seen as a generating function in which the coefficients represent the weights.

From the generating function, if the maximum powers of x and y is m and r respectively, then consider a lattice starting from the bottom left corner $(0,0)$ to top right corner (m,r), and identify each term in the generating function as a node in the lattice graph.

The next is to construct a general binary tree from the identified nodes in the lattice, and then use the general binary tree to construct a share price path.

9.5.1   Tutte polynomial

Given a matroid, one definition of tutte polynomial is via bases of matroid. Following the notations for tutte polynomial as found in [60], the tutte polynomial is the following

$$
\begin{aligned}
T(M;x,y) &= \sum_{B \in \mathcal{B}(M)} x^{i(B)} y^{e(B)} \\
&= \sum_{\substack{\text{Any path } P \text{ that stays} \\ \text{in a bounded lattice region} \\ (0,0) \to (m,r)}} x^{i(P)} y^{e(P)}
\end{aligned}
$$

If the matroid is a lattice path matroid characterized by two bounding lattice paths starting from the bottom left corner $(0,0)$ to top right corner (m,r), the second equation is followed from [60, theorem 3.3], the set of bases of transversal matroid is the bijection of the set of lattice paths staying in the transversal matroid region.

The method in [60, section 6] is to associate each lattice integer point (i,j) in the lattice path matroid to a tutte polynomial. The tutte polynomial is the generating function for two lattice path statistics as stated in [60].

Each integer point (i,j) in the lattice path matroid can represent a set of lattice paths bounded in the matroid which starts from $(0,0)$ to (i,j), where $0 \leq i \leq m$ and $0 \leq j \leq r$.

The lattice paths from $(0,0)$ to (i,j) staying in the matroid region has two basic statistics. Each such lattice path consists of a sequence of north steps and steps.

One statistic, denoted by $i(P)$, is to count the number of the north steps of the lattice path P which lies on the upper bounding lattice path of the matroid. Another statistic, denoted by $e(P)$, is to count the number of the east steps of the lattice path P which lies on the lower bounding lattice path of the matroid.

Suppose the lattice path matroid is characterized by the upper bounding lattice path ω and the lower bounding lattice path η. The two statistics of any lattice

path staying in the matroid is denoted by

$$i(P) \quad = \quad \text{the number of north steps of path } P \ \cap \omega$$
$$e(P) \quad = \quad \text{the number of east steps of path } P \ \cap \eta$$

It is noted that in the given lattice path matroid with (ω, η), where ω, η are two bounding lattice paths staying in the matroid from $(0,0)$ to (m,r). The two statistics satisfy $0 \le i(P) \le r$ and $0 \le e(P) \le m$.

As stated in [60], the Tutte polynomial is the generating function for two basic lattice path statistics. It means that the generating function counts the set of lattice paths staying in the given lattice matroid from $(0,0)$ to (i,j) according to their two statistics $i(P)$ and $e(P)$. This is one relationship between lattice integer points and tutte polynomials introduced in [60, section 6].

9.5.2   Result

We provide another relationship between lattice integer points and tutte polynomials of transversal matroid. The advantage of this connection is to construct a share price path from a given connected lattice path matroid or connected transversal matroid.

We tested that a valid general binary tree can only be constructed from the connected lattice path matroid, which is a paralleogram polyomino. We use an example to illustrate the testing.

Starting from a basic unit cell, which is a connected lattice path matroid. According to the association in [60, section 6], the matroid with labelled tutte polynomial is as follows,
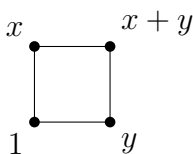

Figure 9.2: connected lattice path matroid containing one unit cell

The top right node's tutte polynomial is $x + y$, its maximum power of x and y is $1, 1$ respectively. Consider a regular integer lattice with the region from $(0,0)$ to $(1,1)$. From the tutte polynomial $x + y$, identify the nodes in the lattice according to the correspondence that each term $x^i y^j$ in the polynomial corresponds to a node (i,j) in the lattice. That is, $x \leftrightarrow (1,0)$, $y \leftrightarrow (0,1)$. In the lattice, the bottom left node $(0,0)$ is identified the initial node. The lattice with identified nodes is the following,
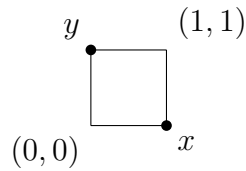
Figure 9.3: lattice with identified nodes one unit cell connected matroid

Set the lattice integer point $(0,0)$ to be the root of the general binary tree, according to the correspondence

$$
\begin{cases}
x & \leftrightarrow \quad \text{left branch in the tree} \\
y & \leftrightarrow \quad \text{right branch in the tree}
\end{cases}
$$

The coefficients of the nodes in the lattice denote the weight of the node in the general binary trees. Then, the corresponded general binary trees of vertices 3 for the two unit cells connected lattice path matroids' lattices are
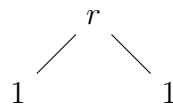
Figure 9.4: One unit cell matroid's tree

According to the definition of connected lattice path matroid in [60, theorem 3.6], connected lattice path matroids containing two unit cells with labelled tutte polynomial are as follows,
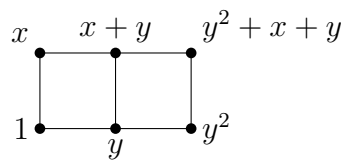
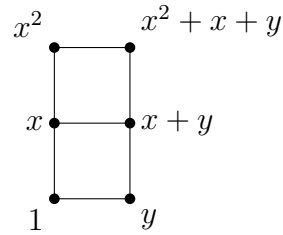Figure 9.5: One two unit cells connected lattice path matroid

Figure 9.6: Another two unit cells connected lattice path matroid

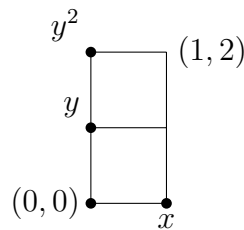The corresponding lattices for the matroids with identified nodes are respectively as follows,

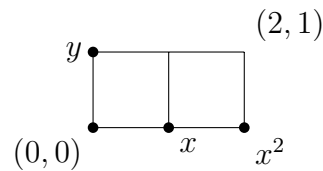Figure 9.7: One two unit cells matroid's lattice

Figure 9.8: One two unit cells matroid's lattice

Set the lattice integer point $(0,0)$ to be the root of the general binary tree, according to the correspondence

$$
\begin{cases}
x & \leftrightarrow \quad \text{left branch in the tree} \\
y & \leftrightarrow \quad \text{right branch in the tree}
\end{cases}
$$

The coefficients of the nodes in the lattice denote the weight of the node in the general binary trees. Then, the corresponded general binary trees of vertices 4 for the two unit cells connected lattice path matroids' lattices are
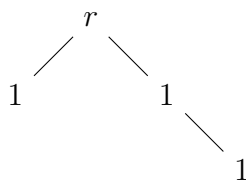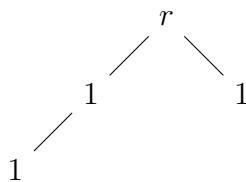
Figure 9.9: One two unit cells matroid's tree

Figure 9.10: Another two unit cells matroid's tree

If the lattice path matroid consiting of two unit cells is not connected, the matroid with labelled tutte polynomial is
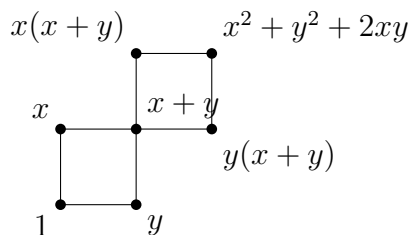

Figure 9.11: One two unit cells connected lattice path matroid

The top right node's tutte polynomial is $x^2 + y^2 + 2xy$, its maximum power of x and y is 2, 2 respectively. Consider a regular integer lattice with the region from $(0,0)$ to $(2,2)$. From the tutte polynomial $x^2 + y^2 + 2xy$, identify the nodes in the lattice, that is,

$$
\begin{cases}
x^2 & \leftrightarrow & (2,0) \\
y^2 & \leftrightarrow & (0,2) \\
2xy & \leftrightarrow & (1,1)
\end{cases}
$$

In the lattice, the bottom left node $(0,0)$ is identified the initial node.
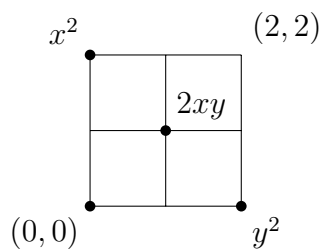
The lattice with identified nodes is the following,

Figure 9.12: lattice with identified nodes one unit cell connected matroid

It can be observed that from the lattice point $(0,0)$, there is no way to the lattice points $(1,1)$, $(2,0)$, $(0,2)$. So, there is no corresponding general binary tree. We conclude that the general binary tree can only be constructed from the connected lattice path matroid.

9.5.3 Representations of share paths via general tree

Motivated by the bijection idea from [17], we provide a construction from a set of general binary trees to the set of share price paths which has the number of down steps greater than or equal to the number of up steps.

The construction is an on-to construction, but it is not a bijection construction. A share price path of n steps can be constructed from many general binary trees of n edges, however, we can construct a specific share price path from one general binary trees of n.

The algorithm from the general binary tree to the share price path is stated as follows,

Step 1: Look to the left. (If there exists left, then there must exist right.)

Step 2: Draw up if left,

draw down if right otherwise. (It always has the left then right.)

Step 3: If nothing, go one level up in the tree.

Consider one unit cell matroid's tree figure 9.4, according to the algorithm, the matroid's share price path of length 2 is
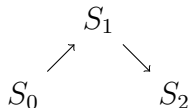

Figure 9.13: the corresponding share path of length 2

Consider the above two unit cells matroids' tree figure 9.9 and figure 9.10 according

to the algorithm, the matroids' share price paths of length 3 are the following
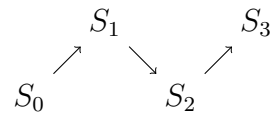
$$S_0 \nearrow S_1 \searrow S_2 \nearrow S_3$$

Figure 9.14: One share path of length 3
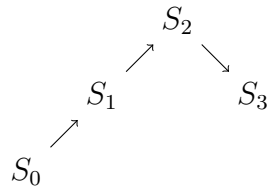
$$S_0 \nearrow S_1 \nearrow S_2 \searrow S_3$$

Figure 9.15: Another share path of length 3

Chapter 10

Conclusions and Future work

10.1 Conclusions

The main topic of the thesis is to develop combinatorial structure and path calculation methods and then apply them to model individual share price path and calculate option prices. The main contributions of this thesis are covered in Chapter 2, 5, 6, 7, 8, 9 and Section 3.1, 3.2, 3.7, 4.5, 4.6. Two working papers including section 5.3, 8.2, 9.5 is now being produced.

In Chapter 2, the main results are the followings. In section 2.1.2, we presented the path calculation method in the formula from the perspective of set theory. Furthermore, originated from [5], we applied the path calculation method to 2 steps, 3 steps binomial models with different setting of probability measures, and present the option pricing formula as lemmas using classical method and path calculation method(Lemma 1, Lemma 2, Lemma 3, Lemma 4, Lemma 5). We also applied the Flajolet generating method to provide two ways of representing 2-steps share price binomial model, we gave new examples of applying the second representation method of share price paths and generating functions(Example 2.3.1, Example 2.3.2). In section 2.3.2, we provided a Lemma 6, which claims a test condition whether or not a generating function can represent a path and illustrated using Example 2.3.3. Moreover, we derived Theorem 7, which is the option value of the exotic claim using path calculation argument and generating function method. In the end of the chapter, applying the generating function knowledge, we provided a argument of relationship between counting paths in a model and the probability generating function of the model and checked that the finite share price binomial model Ω_n is a combinatorial class.

In Chapter 3, the main results are the followings. In section 3.1, 3.2, applying the knowledge of generating function to represent an individual share price path, we gave Example 3.1.1, Example 3.1.2, Example 3.1.3, Example 3.2.1, Example 3.2.2, Example 3.2.3, Example 3.2.4, Example 3.2.5, Example 3.2.6, Example 3.2.7. In section 3.3, based on the Flajolet symbolic method, we gave different presentations of description of sequence construction taken from [15]; multiset construction, cycle construction, powerset construction. In addition, we provide Definition 24(symbolic path transformation) and Corollary 15. In section 3.4.3, we provided the detailed enumeration of general binary strings that contain consecutive 0's less than three using Flajolet symbolic method. In the Example 3.4.2, we added the derivation

of generating function using Flajolet symbolic method. In Example 3.5.1, we gave detailed algorithm how disjoint union of directed cycles is obtained from a permutation with disjoint cycle notation. We also gave more clear and readable proof of proposition 33. In the proof of Proposition 37, we added the derivation of the counts $c(n, k)$ using Flajolet symbolic method. In section 3.5.4, we provided a detailed and applicable algorithm of the natural correspondence from the given integer sequence to a permutation and also give the argument to justify the algorithm is reasonable by analysing the relation between the possible values of a_{ni} and the number of positions in the permutation u_i. Moreover, we derive a Proposition 39 that claims the relationship between integer sequences and permutation of $[n]$. In addition, we added more readable argument of the combinatorial method in the proof of Proposition 48. We also derive a general formula $A(n, k)$ for the counts of the number of set permutation $w \in \mathfrak{S}_n$ with a fixed number of descent $k-1$ (Theorem 48). In the last section of the chapter, we also provided the case study of share price analysis using the knowledge of generating function (Section 3.7); (case I; Example 3.7.1, Example 3.7.2, Example 3.7.3 case II).

In Chapter 4, the main results are the followings. In Section 4.1, we use the simple combinatorial method to count the no of lattice paths of length N. In Section 4.3, we reviewed the enumeration of self-avoiding walks with three directions, in which we added the detailed explanation of derivation of the general recursive before deriving the two variable generating function $\mathcal{G}(t, v)$ and extracting the coefficient $g(n, m)$. In Section 4.6, in the method 2, we derived the solution of counting a path not touching the given bold line segment. In Section 4.6, we applied the path calculation method to solve the winning probability in the gambler ruin problem, in which we got the same answer as using the classical method.

In Chapter 5, the main results are the followings. Motivated from the idea taken from [17] [29], we provided the algorithm of constructing share price path from a set of full binary trees in Section 5.1.1 and gave the share price path interpretation in Section 5.1.2. In Section 5.2.1, we gave the algorithm of constructing share price path from a set of general trees. In Section 5.3.1, we provided an algorithm to represent a share price via a general binary ordered tree.

In Chapter 6, the main results are the followings. We constructed the set of restricted share price paths and its path generating functions in Section 6.1 and then we derived the calculation of option price in the given finite restricted binomial model(Theorem 60).

In Chapter 7, the main results are the followings. In Section 7.1, we gave the Example 7.1.1, Example 7.1.2, Example 7.1.3, Example 7.1.4, Example 7.1.5, Example 7.1.6. In Section 7.2.1, using the idea of bijection between 2-colored Motzkin words and dyck words taken from the paper [30, Page 179], we gave a construction algorithm from parallelogram polyomino to dyck share price paths. It followed with the share price path ineterpreation in Section 7.2.2.

In Chapter 8, the main results are the Section 8.2, in which we provided the method of counting no of restricted paths via unusual stochastic modelling(Lemma 73,

Lemma 74, Lemma 75, Example 8.2.1, Example 8.2.2, Example 8.2.3, Section 8.2.2, Section 8.2.3, Section 8.2.4, Section 8.2.5).

In Chapter 9, the main results are the followings. In general, we gave different presentations of the knowledge of matroid and Tutte polynomial. In Section 9.1, we gave the explicit proof of Example 9.1.1, we added clear explanation in the proof of Example 9.1.2. We also gave Example 9.1.3, Example 9.1.5, more detailed proof of Proposition 79, more detailed proof of Example 9.1.4, more detailed explanation of checking axiom(iii) in (I3) in the proof of Proposition 92. In addition, applying the algorithm of finding maximum matching, we gave Example 9.16. We also gave the proof of Proposition 96 following the definition of the matroid, and then provided a construction algorithm of bipartite graph from any given matroid and explained it using a small example(Example 9.1.7). Furthermore, we provided a more constructive definition of lattice path matroid(Definition 105), and Example 9.3.1 in Section 9.3. In Section 9.4, we gave the methodology of the calculation of Tutte polynomial from a gvien lattice path matroid. Specifically, in Section 9.4.1, applying the Lemma 104, we gave a construction algorithm between the set of lattice paths and the set of their north steps. According to the Definition 105. a construction of E(M), I(M), B(M) is provided for a lattice path matroid M, and we provided an analysis of the calculation of Tutte polynomial $T_M(x, y)$ for the matroid M in Section 9.4 and provided the Example 9.4.1 to illustrate the calculation analysis in Section 9.4.3. Moreover, in the end of the chapter, we provided a new modelling of paths via matroid(Section 9.5).

10.2   Future work

Connection with continuous models is the most interesting direction which was not covered.

In particular, it would be interesting to introduce continuous time binary tree and its application to finance.

Asymptotic results in another interesting direction, in particular connection between number of path and number of path on a lattice, and measures the Hausdorff manifold, preliminary reading and some initial calculation were done.

Appendices

Appendix A

Implementation matlab code

```
// 1. transform_matrix.m
function [W] = transform_matrix(n)
% the forward matrix of dimension n-by-n

 W=zeros(n);
 for r=1:n-1
     cc=r
     if 2*r<=n
         indx=[2*r-1,2*r]
     else
         indx=mod([2*r-1,2*r],n)
     end
     %V=W(r,:);
     W(r,indx)=[1/2,1/2];
     W;
 end
W(n,[end-1,end])=[1/2,0]
end
```

```
// 2. sparse_matrix.m
function [m] = sparse_matrix(n)
%or we can use the sparse forward matrix to allow more restrictions
format rat;
i=[]; j=[]; v=[];
for r=1:n-1
  i=[i,r,r];
  if(2*r<=n)
    j=[j,2*r-1,2*r];
  else
    j=[j,mod([2*r-1,2*r],n)];
  end
  v=[v,1/2,1/2];
end
i=[i,n,n];
j=[j,n-1, n];
```

```matlab
v=[v,1/2,0];
m=sparse(i,j,v,n,n);
end
```

```matlab
//3. test_count.m
function [ ] = test_count
% Input: restriction=the number of not allowing consecutive down steps.
%             T= count the number of restricted paths on which steps, T>=1
% Output: count
% the restriction condition will make sense starting from the time step
   2, so, we consider the given
% information on the first (restriction-1) steps.
prompt = 'the number of not allowing consecutive down steps? ';
restriction = input(prompt)
prompt = 'count the number of restricted paths on which time steps(T>=1)?
   ';
T = input(prompt)

m=restriction-1; % consider the given information on the first
   (restriction-1) steps
format rat;
i=[];      % initial probability of the states of the two steps paths
for j=1:2^m    % it is 2 (each step two choices) to the power m (given
   information on the first  m
                          % steps); say, if m=2, the state set is {11, 10,
                              01, 00}
  i=[i, 1/(2^m)];  % equal probability for jump up and jump down of the
     paths on the given time
                            % steps information
end
n=2^(restriction-1); % the number of rows or columns of the transform
   forward matrix


if  T >= restriction
  % if T greater than restriction, then calculate forward matrix
  %fw= transform_matrix(n) % the forward matrix of dimension n-by-n
  %or we can use the sparse forward matrix to allow more restrictions
  fw=sparse_matrix(n);
  Probb=i*fw^(T-(restriction-1)); % the power denotes the number of time
     steps after the given
                          % time steps information
  CProb= Probb/sum(Probb); % normalized probability at time step T with
     restriction not allowing
                          %  restriction   consecutive down steps
  count=2^T * sum(Probb); % the number of restricted path at time step T
     with the given restriction
else
  count=2^T;  % the number of restricted path at time step T with the
     given restriction
```

```
                          % when T< restriction, the restriction not works,
                                the counts just as normal
end

X = sprintf('The number of restricted paths not allowing %d consecutive
    down steps at time %d is %d',restriction, T,count);
disp(X)

end
```

```
// 4. Test:

>> test_count
the number of not allowing consecutive down steps? 2

restriction =

      2

count the number of restricted paths on which time steps(T>=1)? 3

T =

      3

The number of restricted paths not allowing 2 consecutive down steps at
    time 3 is 5




>> test_count
the number of not allowing consecutive down steps? 3

restriction =

      3

count the number of restricted paths on which time steps(T>=1)? 5

T =

      5

The number of restricted paths not allowing 3 consecutive down steps at
    time 5 is 24

%%%%%%%%%%%%%%%%%if restriction number is big%%%%%%%%%%%%
>> test_count
the number of not allowing consecutive down steps? 15
```

```
restriction =

      15

count the number of restricted paths on which time steps(T>=1)? 25

T =

      25

The number of restricted paths not allowing 15 consecutive down steps at
    time 25 is 33548288
```

Bibliography

[1] Claessens, S., Kose, M.M.A., Laeven, M.L. and Valencia, F., 2014. Financial crises: Causes, consequences, and policy responses. International Monetary Fund.

[2] Hao, W., Lefèvre, C., Tamturk, M. and Utev, S., 2019. Quantum option pricing and data analysis. Quantitative Finance and Economics, 3(3), p.490.

[3] Lento, C. and Gradojevic, N., 2013. The Effectiveness of Option Pricing Models During Financial Crises. In Rethinking Valuation and Pricing Models (pp. 1-11). Academic Press.

[4] Black, F. and M. Scholes. (1973). The Pricing of Options and Corporate Liabilities, Journal of Political Economy 81, No. 3 (May-June 1973), pp. 637-654.

[5] Cox, J.C., Ross, S.A. and Rubinstein, M., 1979. Option pricing: A simplified approach. Journal of financial Economics, 7(3), pp.229-263.

[6] Cox, J.C. and S.A. Ross. (1976). The Valuation of Options for Alternative Stochastic Processes, Journal of Financial Economics 3, No. 1 (January-March 1976) pp. 145-166.

[7] Henderson, V., 2014. Black–Scholes Model. Wiley StatsRef: Statistics Reference Online.

[8] Baaquie, B.E., 2009. Interest rates and coupon bonds in quantum finance. Cambridge University Press.

[9] Ma, X. and Utev, S., 2012. Modelling the share prices as a hidden random walk on the lamplighter group. In Mathematical and Statistical Methods for Actuarial Sciences and Finance (pp. 263-270). Springer, Milano.

[10] Karadeniz, R.S. and Utev, S., 2015. Modelling share prices via the random walk on the lamplighter group.

[11] Lefèvre, C., Loisel, S., Tamturk, M. and Utev, S., 2018. A quantum-type approach to non-life insurance risk modelling. Risks, 6(3), p.99.

[12] Taylor, E.F., Vokos, S., O'Meara, J.M. and Thornber, N.S., 1998. Teaching Feynman's sum-over-paths quantum theory. Computers in Physics, 12(2), pp.190-199.

[13] Sinha, S. and Sorkin, R.D., 1991. A Sum-over-histories Account of an EPR (B) Experiment. Foundations of Physics Letters, 4(4), pp.303-335.

[14] Sedgewick, Robert, and Philippe Flajolet. An introduction to the analysis of algorithms. Addison-Wesley, 2013.

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

[16] https://math.stackexchange.com/questions/1633762/using-generating-functions-to-answer-how-many-bit-strings-of-length-n-have-no-00

[17] Federico, Class EC Homework. http://math.sfsu.edu/federico/Clase/EC/Homework/3.3.Jorge.pdf

[18] Brylawski, T. and Oxley, J., 1992. The Tutte polynomial and its applications. Matroid applications, 40, pp.123-225.

[19] Mohanty, G., 2014. Lattice path counting and applications. Academic Press.

[20] Racorean, O.S., 2014. Crossing Stocks and the Positive Grassmannian I: The Geometry behind Stock Market. Available at SSRN 2512437.

[21] Stanley, R.P., 2011. Enumerative Combinatorics Volume 1 second edition. Cambridge studies in advanced mathematics.

[22] Chandran,Combinatorics. Mod-04 Lec-31 Solving recurrence relations using generating functions - Part (2)

[23] Wallner, Michael. Lattice Path Combinatorics. na, 2013.

[24] Williams, Lauren K. "Enumerating up-side self-avoiding walks on integer lattices." The Electronic Journal of Combinatorics 3.R31 (1996): 2.

[25] Grinstead, C.M. and Snell, J.L., 2006. Grinstead and Snell's introduction to probability. Chance Project.

[26] John Tsitsiklis. Class MIT 6.041 Probabilistic Systems Analysis and Applied Probability, Fall 2010 Lecture 16 Markov Chains I.

[27] Tolver, A., 2016. An introduction to Markov chains. Department of Mathematical Sciences, University of Copenhagen.

[28] Utev, Sergey, 2018. Summary of Lectures 1-38. Financial mathematics lecture notes. Department of Mathematics, University of Leicester.

[29] Dershowitz, N. and Rinderknecht, C., 2015. The average height of Catalan trees by counting lattice paths. Mathematics Magazine, 88(3), pp.187-195.

[30] M.P. and Viennot, G., 1984. Algebraic languages and polyominoes enumeration. Theoretical Computer Science, 34(1-2), pp.169-206.

[31] Leroux, P. and Rassart, E., 1999. Enumeration of symmetry classes of parallelogram polyominoes. arXiv preprint math/9901135.

[32] Delest, M.P., 1988. Generating functions for column-convex polyominoes. Journal of Combinatorial Theory, Series A, 48(1), pp.12-31.

[33] Conrad, Brian. Handout. Interior, closure, and boundary. http://virtualmath1.stanford.edu/ conrad/diffgeomPage/handouts/closure.pdf

[34] Weisstein, Eric W. "Open Ball." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/OpenBall.html

[35] Weisstein, Eric W. "Polyomino." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Polyomino.html

[36] Weisstein, Eric W. "Convex Polyomino." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/ConvexPolyomino.html

[37] Alexander Bogomolny. PascalTriangleProperties. https://www.cut-the-knot.org/arithmetic/combinatorics/PascalTriangleProperties.shtml

[38] The Gambler's Ruin. https://www.mathpages.com/home/kmath084/kmath084.htm

[39] Mark J Barlow. Mathematical Methods in Actuarial Sciences. Mathematics 3rd Year Project Spring 2007/2008, School of Mathematical Sciences, University of Nottingham.

[40] Adams, S., 2015. MA3H2 Markov Processes and Percolation theory. update.

[41] Weisstein, Eric W. "Lattice Path." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/LatticePath.html

[42] Abel, Zachary. Multivariate generating functions and other tidbits. Mathematical Reflections, March 2006.

[43] Math Stackexchange. https://math.stackexchange.com/questions/314788/let-p-be-an-odd-prime-number-how-many-p-element-subsets-of-1-2-3-4-ldo

[44] Sedgewick, Robert. 1.4 Compositions and Partitions. Lecture 1, Combinatorial structures and OGFs.

[45] Min Wang, Sergey Utev: Christmas trees and option pricings. (paper in preparation).

[46] Federico, Lecture notes of Matroids. http://math.sfsu.edu/federico/Clase/Matroids /LectureNotes/lectures1-25.pdf

[47] Goemans, Lecture 8 Matroids, Lecture notes of Advanced Combinatorial Optimization. https://math.mit.edu/ goemans/18438F09/lec8.pdf

[48] http://math.mit.edu/ goemans/18453S17/matroid-notes.pdf

[49] http://www-math.mit.edu/ goemans/18433S09/matroid-notes.pdf

[50] https://www.whitman.edu/Documents/Academics/Mathematics/hillman.pdf

[51] https://www.youtube.com/watch?v=-tj0T5WihzI

[52] Okstate. https://math.okstate.edu/people/binegar/3013-S99/3013-l10.pdf

[53] Chekuri, Chandra. Introduction to Matroids, CS 598CSC: Combinatorial Optimization https://courses.engr.illinois.edu/cs598csc/sp2010/Lectures/Lecture14.pdf

[54] Saurabh, Saket. Lecture 13, Advanced Graph Algorithms. https://kam.mff.cuni.cz/ ashutosh/lectures/lec13.pdf

[55] Jonathan Hirata. NOTES ON MATCHING, MIT. http://www-math.mit.edu/ djk/18.310/Lecture-Notes/MatchingProblem.pdf

[56] Paul E. Black and Alen Lovrencic, "subgraph", in Dictionary of Algorithms and Data Structures [online], Paul E. Black, ed. 17 December 2004. (accessed TODAY) Available from: https://www.nist.gov/dads/HTML/subgraph.html

[57] Arobinda Gupta, Lecture slide of Matching and Covering, CS60047: Advanced Graph Theory http://cse.iitkgp.ac.in/ agupta/graph/Matching.pdf

[58] Bipartite Graphs/Matching (Intro)-Tutorial 12 D1 Edexcel https://www.youtube.com/watch?v=JpapV5DrBek

[59] Maximum Matching Algorithm - Tutorial 13 D1 Edexcel A-Level https://www.youtube.com/watch?v=gbasc4F-7hk

[60] Bonin, J., de Mier, A. and Noy, M., 2003. Lattice path matroids: enumerative aspects and Tutte polynomials. Journal of Combinatorial Theory, Series A, 104(1), pp.63-94.

[61] Bonin, J.E., 2010. An introduction to transversal matroids.

[62] Federico Ardila. Tutte polynomials in combinatorics and geometry, Modern Math Workshop SACNAS National Conference, San Antonio, Texas, 2 de octubre de 2013 http://math.sfsu.edu/federico/Talks/SACNAStutte.pdf

[63] Porter, A.M. and Balof, B., 2015. The Tutte polynomial and applications.

[64] Merino, C., Ramírez-Ibáñez, M. and Rodríguez-Sánchez, G., 2012. The Tutte polynomial of some matroids. International Journal of Combinatorics, 2012.

[65] Lin, B., 2013. An Introduction of Tutte Polynomial.

[66] Brandt, M., 2015. The Tutte Polynomial.