APPLICATIONS OF SIEVE METHODS IN ENUMERATIVE COMBINATORICS

RISHABH VENKATARAMANI

Abstract. Sieve methods are very useful tools to locate elements of a set and calculate the cardinalities of sets. This paper will show many forms of sieve methods and their applications in combinatorics. Most of the sieve methods are in conjunction with another combinatorial method known as the *Principle of Inclusion-Exclusion*. This principle will help us identify the usefulness of sieve methods and their applications, such as restricted permutations, ferrers boards, and involutions.

1. Introduction

Sieve methods are an age-old mathematical technique used to "sieve" through sets, in order to extract specific elements. Essentially, these sieves separate out the unwanted elements, extracting patterns and structures. Sieve methods are applicable in number theory, cryptography, combinatorics, and even computer science. In this paper, we will discuss predominantly the usefulness of sieve methods in (enumerative) combinatorics.

There are two main types of sieve methods:

(1) Approximating with an overcount and then subtracting off an overcounted approximation of our error, then overcounting this error, subtracting it off, and so on, until we have a correct count of the objects. This method is essentially the *Principle of Inclusion-Exclusion*.
(2) The elements of the larger set can be weighted, in such a way that unwanted elements cancel out, leaving us with our original set.

This paper will focus on the first type, though we will touch upon the second type.

One of the most well-known, if not the most well-known sieve is the *Sieve of Eratosthenes*. This is a method of finding the number of primes of a given set. The procedure for this comes out to be something of this sort:

Suppose we have a list of numbers $(1, 2, \ldots, n)$. We want to find the number of prime numbers in this list. We can do this in a systematic approach instead of testing out every single number, which would take a considerable amount of time, especially for large n. Notice that 1 is not a prime number, but 2 is (which we can find out through testing). Then, any proper multiples (multiples that do not equal 2) of 2 cannot be prime. Therefore, we can remove them from our list. The next number in our list that has not been crossed out must be a prime number, and the (proper) multiples of this prime number may be crossed off as well. We can continue this until we are left with our primes. This is the essence of the Sieve of Eratosthenes.

Something to note is that when we are supposedly crossing out multiples of 2 primes p and q, we are crossing the numbers which are multiples of both primes p and q twice. This looks very much like a job for Inclusion-Exclusion.

Date: July 17, 2023.

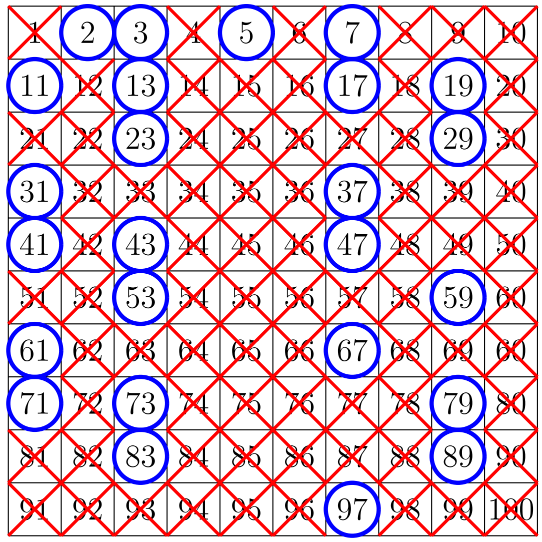

Figure 1. Sieve of Eratosthenes (1-100)

In this paper, we will explore sieve methods similar to the Sieve of Eratosthenes, which are woven through the intricacies of the Principle of Inclusion-Exclusion. We will talk bout many of the applications of the Principle of Inclusion-Exclusion. After going over some background, we will explicitly state the Principle of Inclusion-Exclusion and then prove it. We will discuss some simple examples of the use of the Principle of Inclusion-Exclusion, and then discuss different combinatorial structures, which include ferrers boards and involutions. Finally, we will branch into a more algebraic approach and show how the Principle of Inclusion-Exclusion can be expressed in terms of matrices and determinants.

2. Preliminaries

Throughout this paper $[n]$ will be defined as the set $\{1, 2, \dots, n\}$.

Definition 2.1. Let $[n] \times [n]$ denote the set of ordered pairs of (a, b), such that $1 \leq a, b \leq n$.

Therefore, $[n] \times [n]$, would be the set of ordered pairs $\{[(1,1),(1,2),\dots,(1,n)], [(2,1),(2,2), \dots,(2,n)], \dots, [(n,1),(n,2),\dots,(n,n)]\}$.

2.1. Generating Functions and Power Series.

One of the most convenient ways to count objects is to represent them as a power series $p(n) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x^1 + a_2 x^2 + \cdots$. This is what is known as an ordinary generating function.

There are two main types of generating functions, *ordinary generating functions*, like the one we saw prior, and *exponential generating functions*. For $\mathbb{N}$, we can represent the ordinary generating function $f(n)$ as the formal power series:

$$\sum_{n \geq 0} f(n) x^n.$$

The exponential generating function of $f(n)$ can be represented as the formal power series

$$\sum_{n \geq 0} f(n) \frac{x^n}{n!}.$$

Note that all *formal* means is that we do not care whether the function converges or not, which is irrelevant to the content of this paper.

3. The Principle of Inclusion-Exclusion

Put in simple terms the Principle of Inclusion-Exclusion (PIE) is a combinatorial argument used to find the cardinality of the union of sets by adding and subtracting off the intersections of the subsets. We can use the following Venn diagram to visualize this.

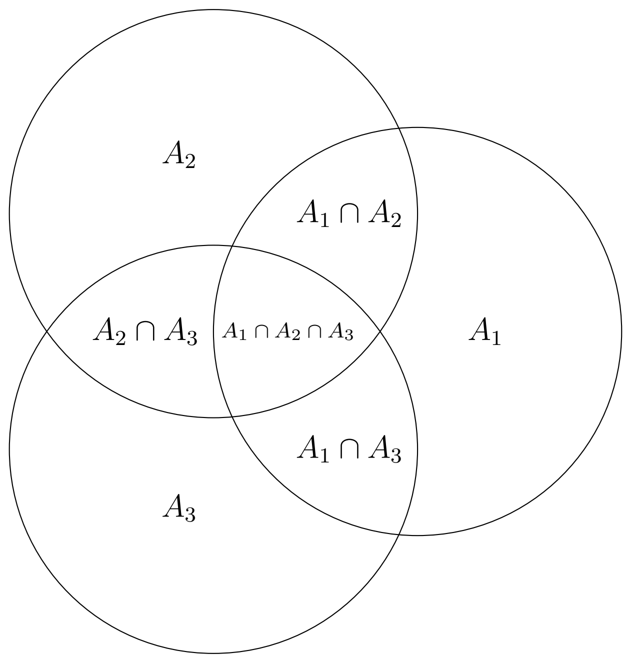

Figure 2. Venn Diagram of PIE for 3 sets

Suppose we have sets A_1, A_2, and A_3 that are subsets of X. Suppose we are trying to find $|X|$, the cardinality of the union of the three subsets. Note that the expression

$$|A_1| \cup |A_2| \cup |A_3|$$

has overcounted each $A_1 \cap A_2$, $A_2 \cap A_3$, and $A_1 \cap A_3$ twice, and $A_1 \cap A_2 \cap A_3$ thrice. Therefore, we must subtract each pairwise intersection and add back the intersection of all three sets.

Theorem 3.1. *Suppose we have subsets $A_1, A_2, \ldots, A_n$ of set X. Let X be the union of the sets A_i for integers $1 \leq i \leq n$. $|X|$ defined as $|A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n|$. Then, $|X|$ can be expressed as*

$$|X| = \sum_{i=1}^{n} |A_i| - \sum_{1 \leq i \leq j \leq n} |A_i \cap A_j| + \cdots + (-1)^{n-1} |A_1 \cap A_2 \cap \cdots \cap A_n| \tag{3.1}$$

Proof. Suppose we have an element $j \in X$, that is in k of the m subsets. Suppose we assume that j is in subsets $A_1, \ldots A_k$. This means that j is not in the rest of the subsets $A_{k+1}, \ldots, A_m$.

The following is how many times j is counted:

$$\begin{gathered}\binom{k}{1} \text{ times in } A_1, \ldots, A_k \\ -\binom{k}{2} \text{ times in } A_a \cap A_b \\ +\binom{k}{3} \text{ times in } A_a \cap A_b \cap A_c \\ \vdots \\ +(-1)^{k-1}\binom{k}{k} \text{ times in } A_1 \cap A_2 \cdots \cap A_k \\ = \binom{k}{0} - (1-1)^k = 1.\end{gathered}$$

Therefore, every element $j \in X$ is counted once and only once in the set X for all m, as desired. ■

The following is a consequence of PIE. Let A be a finite set and let S be a set of properties that may or may not be satisfied by set A. For any $T \subseteq S$, we can denote $f_=(T)$ to be the number of objects in A that exactly satisfy the properties in T. This means that A fails to satisfy any of the properties of $S - T$. The number of objects in A that at least satisfy the properties in T can be denoted by $f_\leq(T)$. If Y is the set that ranges over all the subsets of T, it can be seen from this that

$$f_\geq(T) = \sum_{U \supseteq T} f_=(U). \tag{3.2}$$

Therefore,

$$f_=(U) = \sum_{U \supseteq T} (-1)^{|U-T|} f_\geq(T). \tag{3.3}$$

More importantly, the number of objects satisfying none of the properties is seen as

$$f_=(\varnothing) = \sum_{U \supseteq T} (-1)^{|U|} f_\geq(T). \tag{3.4}$$

In part, this is useful since many times it is much harder to count the number of objects that do satisfy a set of properties. Using this identity, we can count the number of objects that satisfy none of the properties and are in turn counting the number of objects that do satisfy the properties.

We can also think of the number of objects in A that contain at most the properties in T denoted by $f_\leq(T)$. Then we get

$$f_\leq(T) = \sum_{U \subseteq T} f_=(U), \tag{3.5}$$

and

$$f_=(U) = \sum_{U \subseteq T} (-1)^{|T-U|} f_\leq(T). \tag{3.6}$$

3.1. **Euler's Totient (ϕ) Function.** PIE, although in the domain of (enumerative) combinatorics, is widely applicable to other branches and forms of mathematics (it is simply adding and subtracting). One such application is in the realm of number theory with Euler's totient function. This is a useful method for determining the number of relatively prime integers that are less than an integer n.

Theorem 3.2. *Let n be a positive integer greater than 1, and let $\phi(n)$ be the function defined by $\phi(n) = \#\{k, 1 \leq k \leq n, gcd(n,k) = 1\}$. If $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$, then*

$$\phi(n) = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\left(1 - \frac{1}{p_3}\right)\cdots\left(1 - \frac{1}{p_k}\right). \tag{3.7}$$

Proof. For each $i \in 1, 2, \ldots, k$, let A_i be the set of integers $m \leq n$ that is divisible by the prime p_i. Then,

$$\phi(n) = n - \left|\bigcup_{i=1}^{k} A_i\right|.$$

Using PIE, we can expand $|A_1 \cup A_2 \cup \cdots \cup A_k|$ to $I_1 - I_2 + \cdots + (-1)^k I_{k-1} = S$, where I_1 is the sum of the intersections of the sets. Each I_i contains the multiples of the primes $p_{1_j}, p_{2_j}, \cdots, p_{k_j}$, thus the sum S can be expressed as

$$n\left(\left(\frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_k}\right) - \left(\frac{1}{p_1 p_2} + \frac{1}{p_2 p_3} + \cdots + \frac{1}{p_{k-1} p_k}\right) + \cdots + \left((-1)^{k-1}\frac{1}{p_1 p_2 \cdots p_k}\right)\right).$$

Plugging this back into 3.1, the equation becomes

$$n\left(1 - \left(\frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_k}\right) + \left(\frac{1}{p_1 p_2} + \frac{1}{p_2 p_3} + \cdots + \frac{1}{p_{k-1} p_k}\right) + \cdots + \left((-1)^{k}\frac{1}{p_1 p_2 \cdots p_k}\right)\right),$$

which simplifies to

$$n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\left(1 - \frac{1}{p_3}\right)\cdots\left(1 - \frac{1}{p_k}\right).$$

■

This is one of the many applications of PIE and sieve methods in general, the following sections will provide us with that insight.

4. Restriction and Permutation

4.1. **Derangements.** A very straightforward application of PIE is the *derangement problem.* Let $\mathfrak{S}_n$ denote the set of permutations of S. Then, how many $\omega \in \mathfrak{S}_n$ are there such that $\omega(i) \neq i$? This type of permutation is what is known as a derangement. An example of a derangement is the number of permutations of n objects into n boxes. Let $D(n)$ denote the number of ways to arrange n objects in n boxes such that no object n_i is put in the ith box. We now see the power of equation (3.4). By fixing an object and permuting the rest of the objects, we are in effect finding the number of permutations that do not satisfy the conditions (which we can then use to calculate the number of objects that do satisfy the properties). Using equation (3.4), we can write $D(n) = f_=(\varnothing) = \sum_{U \supseteq S} (-1)^{|U|} f_{\geq}(S)$, where S is the set of properties that $\omega(i) \neq i$ for all $i = 1, 2, \ldots n$. $\omega(i)$ maps the objects to the boxes. It can be seen from this that $D(n) = \sum_{U \supseteq S} (-1)^{|U|}(n - |U|)!$. Setting k equal to $|S|$, we get the following equation

$$D(n) = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} k!, \tag{4.1}$$

which can be rewritten as

$$D(n) = n!\left(1 - \frac{1}{2!} + \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}\right), \tag{4.2}$$

Taking the limit as n approaches infinity of the previous equation yields an asymptotic formula that may sometimes be preferred over the closed form we have just derived:

$$\lim_{n\to\infty} D(n) = \frac{n!}{e}. \tag{4.3}$$

This equation also gives the closest integer to $D(n)$ for all $n \geq 0$. Furthermore, we can express the closed form as a generating function as follows:

$$\sum_{n\geq 0} D(n)\frac{x^n}{n!} = \frac{1}{e^x(1-x)} \tag{4.4}$$

Permutations of this sort are known by the general category of permutations with restricted positions. In the previous example of derangements, our restriction was $\omega(i) \neq i$, but many such restrictions may arise.

4.2. **Boards and Rook Polynomials.** The terminology behind boards, rook numbers, rook polynomials, etc. comes from the game of chess, all of which provide pathways to restricted positions.

Definition 4.1 (Board)**.** Define a Board $B \subseteq [n] \times [n]$.

Definition 4.2 (Graph)**.** Define the graph $G(\omega)$ of $\omega \in \mathfrak{S}_n$ to be $G(i, \omega(i)) : i \in [n])$.

Definition 4.3 (Rook Placement)**.** $N_j = |\{\omega \in \mathfrak{S}_n : j = |(B \cap G(w))|\}|$

Definition 4.4 (Rook Number)**.** Define r_k to be the number of k-subsets such that no two elements have a single coordinate in common.

This is corresponding to placing k rooks onto a board B such that no two rooks attack each.

We also define a rook polynomial by

$$r_B(x) = \sum_k r_k x^k.$$

With all of this terminology set, we are ready for the following theorem.

Theorem 4.5.

$$N_n(x) = \sum_{k=0}^{n} r_k (n-k)!(x-1)^k \tag{4.5}$$

Proof. Note that N_j is simply the number of ways to place n rooks on $[n] \times [n]$ such that exactly j of these rooks are on the board. ■

Example. Consider the board $B = \{(1,1), (2,2), (3,3), (3,4), (4,4)\}$. Notice that we can find each N_j. $N_0 = 6$, $N_1 = 9$, $N_2 = 7$, $N_3 = 1$, $N_4 = 1$. From this, we can find out that the rook polynomial $r_B(x)$ of this board is

$$r_B(x) = 1 + 5x + 8x^2 + 5x^3 + x^4$$

Each term in the polynomial $r_B(x)$ corresponds to placing k rooks into the board B. For example, the term $5x^3$ represents the 5 ways to place 3 non-attacking rooks onto the board B.

Notice how much simpler this is than manually trying to calculate the number of ways to place k rooks onto a board B

Derangements, from the previous section, can also be described as computing N_0 from the board $\{(1,1),(2,2),\dots,(n,n)\}$. Plugging $r_k = \binom{n}{k}$ into the equation $N_n(x) = \sum_{k=0}^{n} r_k(n-k)!(x-1)^k$, gives us

$$N_0 = \sum_{k=0}^{n} (-1)^k \frac{n!}{k!}$$

4.3. **Ménage Problem.** The *ménage problem* asks for the number of ways to seat couples around a circular dining table such that men and women alternate and no one sits next to their partner (ménage in French means household, referring to the couples).

In other words, this is asking for $M(n)$ with the restriction that $\omega(i) \neq i, i+1 \pmod{n}$. We are essentially being asked to find N_0 of the board

$$\{(1,1),(2,2),\dots,(n,n),(n,1),(1,2),(2,3),\dots,(n-1,n)\}$$

which accounts for our restriction of $\omega(i) \neq i, i+1 \pmod{n}$. This is simply counting the number of ways to pick k points from a group of $2n$ points in a circle, such that no two are consecutive.

Theorem 4.6. *The number of ways to choose k points from a collection of $2n$ points in a circle such that no two are consecutive is given by the formula*

$$\frac{2n}{2n-k}\binom{2n-k}{k}$$

Proof. Start by thinking about how many ways to choose if the point a_i isn't chosen. We can picture this as having $2n-k$ points, containing a_i, and inserting k new points into the $2n-k$ spaces available (this is an example of stars and bars). This can be done in $\binom{2n-k}{k}$. If a_i is to be chosen, we arrange $2n-k-1$ points around the circle and choose one of them to be a_i and "mark" it as chosen. Then we have to insert $k-1$ points in $\binom{2n-k-1}{k-1}$. Summing these two cases gives us

$$\binom{2n-k}{k} + \binom{n-k-1}{k-1} = \frac{2n}{2n-k}\binom{2n-k}{k}$$

■

Note that this proof is a purely bijective proof that we desire.

5. Ferrers Boards

This is a special class of boards known as *ferrers boards*, but to gain an appreciation for their usefulness, we must first understand what they represent.

5.1. **Partitions.** A *partition*, or more specifically an integer partition is very similar to what it sounds like. It is the number of ways to split a number n into positive integers that add up to n. For example for $n = 4$, we get

$$1 + 1 + 1 + 1$$
$$1 + 1 + 2$$
$$2 + 2$$
$$1 + 3$$
$$4$$

The order of the sum does not matter, but usually, it is given from least to greatest. A simplified way of writing the partitioned sum, is using the following notation

1. If any number x is repeated n number of times, then represent it as x^n
2. List out the integers, in accordance with rule 1

From that we get:

$$1 + 1 + 1 + 1 = (1, 1, 1, 1) = \boxed{(1^4)}$$
$$1 + 1 + 2 = (1, 1, 2) = \boxed{(1^2, 2)}$$
$$2 + 2 = (2, 2) = \boxed{(2^2)}$$
$$1 + 3 = \boxed{(1, 3)}$$
$$4 = \boxed{(4)}$$

Ferrers boards are used to describe the partitions. Each value is depicted by the number of boxes in each row. For example, this is the Ferrers board for $1 + 1 + 2$

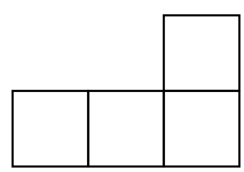

Figure 3. Ferrers board of size $(1^2, 2)$

As you can see, each column matches with the number in the partition. This is also partly why it is slick to have the sums from least to greatest.

5.2. **Stirling Numbers.** *Stirling numbers of a second kind* are of the form $S(n, k)$, and these numbers posses many interesting properties.

Definition 5.1. A *Stirling number of the second kind* is the number of partitions of an n-element set into k non-empty disjoint subsets.

Definition 5.2. Let $A = A_1, A_2, \ldots, A_k$. If $A_1 \cup A_2 \cup \cdots \cup A_k = [n]$, then $S(n, k)$ is the number of ways to partition $[n]$ into k groups. (Note that A_i must not be empty and must be disjoint)

Proposition 5.3.
$$S(n,k) = \sum_{1\le a_1\le a_2\le\cdots\le a_{n-k}\le k} a_1a_2\cdots a_{n-k}. \tag{5.1}$$

This follows straight from our definition of Stirling numbers. We will not prove it in this paper, as it is not essential to the core understanding.

Proposition 5.4.
$$S(n,2) = 2^{n-1} - 1 \tag{5.2}$$
for $n \ge 1$*.*

Proof. We can choose some element k to be in set A and the rest of the $n-1$ elements have 2 choices for which subsets it goes into, either set A or B. Therefore, there are 2^{n-1} to distribute the rest of the $n-1$ elements. However, we have over counted, since set B cannot be empty. Therefore, the total number of arrangements is $2^{n-1}-1$. ■

Note that there isn't a simple closed form for Stirling numbers in general, but this is a special case for *Stirling numbers of the second kind.*

Example. $S(5,3) = (1\cdot1)+(1\cdot2)+(1\cdot3)+(2\cdot2)+(2\cdot3)+(3\cdot3) = 1+2+3+4+6+9 = 25$

An interesting property of these numbers is their recurrence:

Proposition 5.5.
$$S(n,k) = S(n-1,k-1) + k\cdot S(n-1,k) \tag{5.3}$$

Proof. Let us partition $[n-1]$ into subsets $a_1, a_2, \ldots, a_{k-1}$. We have n left over, and we can partition $[n]$ into k subsets with n being its own subset. Similarly, if we partition $[n-1]$ into k different subsets, then the remaining n has k subsets to choose from. This constructs a bijection between the right-hand side and the left-hand side. ■

5.3. **Rook Numbers.** Given integers $0 \le b_1 \le b_2 \le \cdots \le b_n$, the Ferrers board F is defined as
$$F = \{(i,j) : 1 \le i \le n, 1 \le j \le b_i\}.$$

Ferrers boards, like before using cartesian coordinates.

Theorem 5.6. *Let* $(b_1, b_2, \ldots, b_n)$ *denote the Ferrers board* F*.* $\sum r_k x^k$ *is the rook polynomial for the Ferrers board* F*. If* $s_i = b_i - i + 1$*, then*
$$\sum_k r_k(x)_{m-k} = \prod_{i=1}^{n}(x+s_i) \tag{5.4}$$

Proof. To place k rooks on board B', where B' is the Ferrers board of shape $\{x+b_1, x+b_2, \ldots, x+b_n\}$. Therefore $B' = B \cup C$ where C is the rectangle of dimension $x \times n$ directly underneath B.

We can count the rook number $r_k(x)_{m-k}$ in two different ways:

(1) We can place k of the m rooks B order in some k out of the total n columns, and then place the remaining $m-k$ rooks in C, in the remaining $n-k$ columns, such that no rooks attack each other. Placing rooks on B and C is equal to r_k and $(x)_{m-k}$ by definition, getting
$$\sum_k r_k(x)_{m-k}.$$

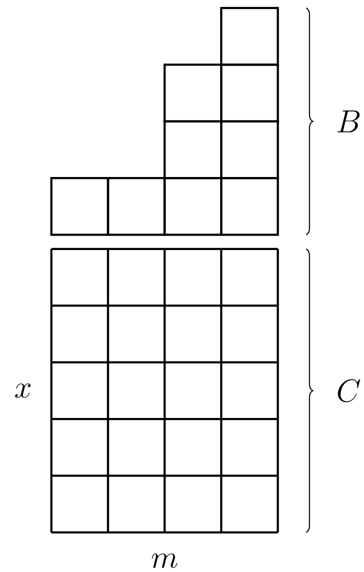

Figure 4. Board B'

(2) There are $x + b_i$ ways to place a rook in the column a_i of B, and $x + b_{i-1} - 1$ ways to place a rook in the $(i+1)th$ column, such that it does not attack the first rook placed, giving us

$$\prod_{i=1}^{n}(x + s_i)$$

Since both of these cases are different ways to count the same permutation, they are equal, satisfying 5.4 as desired. ■

The following is a direct consequence of 5.6.

Corollary 5.7. *Let T be the triangular board $\{(0, 1, 2, \ldots, m-1)\}$. Then $r_k = S(m, m-k)$.*

Proof. From 5.6, and from the fact that $s_i = 0$ (since $b_i = i - 0$), we have that $x^m = \sum r_k \cdot (x)_{m-k}$. We can see that from previous bit on Stirling numbers (of the second kind), $r_k = S(m, m-k)$. ■

That was a simple plug-and-chug proof. The following is a combinatorially inclined proof.

Proof. Suppose we have a rook on (i, j), we define i and j to be in the same block of the partition. We want to create a partition of $[n]$ into $n - k$ blocks to place k rooks on board B, where $B = \{(i, j) : 1 \leq i \leq m, 1 \leq j \leq i\}$. Notice that this corresponds to $r_k = S(m, m-k)$. ■

For example, if we had rooks on $(1, 3), (2, 4), (4, 5), (8, 9)$, then we have the blocks: $\{1, 3\}$, $\{2, 4, 5\}$, $\{6\}$, $\{7\}, \{8, 9\}$.

Theorem 5.8. *For nonnegative integers $0 \leq b_1 \leq b_2 \leq \cdots \leq b_m$. We let $f(b_1, b_2, \ldots, b_m)$ denote the number of Ferrers boards with every column occupied such that it has the same rook polynomial as Ferrers board of shape $(b_1, b_2, \ldots, b_m)$. Adding enough initial 0's to $b_1, b_2, \ldots, b_m$ to get a shape $(c_1, c_2, \ldots, b_n) = (0, 0, \ldots, 0, b_1, \ldots, b_m)$. From this, $s_1 = 0$*

and $s_i \leq 0 for 2 \leq i \leq n$ *from* $s_i = b_i - i + 1$*. For every* i*, suppose* a_i *of the* s_j*'s are equal to* $-i$*, so that* $\sum_{i\geq 1} = m + 1$*. Then,*

$$f(b_1, b_2, \ldots, b_m) = \prod_{i\geq 1} \left(\frac{a_i + a_{i+1} - 1}{a_{i+1}} \right).$$

This theorem is quite a mouthful to understand, so before the proof, we will do a quick example to get an intuition of the theorem.

Example. We want to find the number of Ferrers boards with the same rook polynomial as the triangular board $(1, 2, \ldots, n)$. We can add a 0 to the beginning, such that $s_i = -1$ for $i \geq 2$, and $s_1 = 0$. This is necessary since we cannot have $s_i = 0$ (for all values of i). Therefore, there are no other Ferrers boards, with the same rook polynomial as the triangular board.

Proof. ■

Theorem 5.9. *For* $x \in \mathbb{R}$*,*

$$x^n = \sum_{k=0}^{n} S(n, k) \cdot (x)_k \tag{5.5}$$

6. Involutions

An involution is any set (or function) that maps to itself. In algebraic terms, it is a function that is its own inverse. In combinatorics, it is any mapping that results in a bijection.

Definition 6.1. An involutory function is a function $f : Y \to Y$, such that $f(f(x)) = x$ for all $x \in Y$.

Example. An example of a non-trivial involution is the function $f(x) = \frac{1}{x}$, resulting in $\frac{1}{x} \to x$. Also note that $f(f(x)) = x$, satisfying our condition.

Proposition 6.2. *Let there exist some set* Y*, and let* n *be a non-negative integer. Then,*

$$\sum_{|Y|=2n} f_{\geq}(Y) = \sum_{|Y|=2n+1} f_{\geq}(Y) + f_{=}(\varnothing). \tag{6.1}$$

Proof. Note that $f_{\geq}(Y)$ denotes the number of objects that have *at least* the properties in Y and $f_{=}(Y)$ denotes the number of objects that have *exactly* the properties in Y. The left-hand side of this equation is the cardinality of the set of objects $x \in A$, such that $Z \subseteq Y$ and $|Y|$ is even. Denote this set as M. The right-hand side of this equation is the sums of the cardinalities of the sets S and T, where S is the set of objects that do not satisfy the properties, and T is the set of objects $x' \in A$, such that $Z' \subseteq Y$ and $|Y|$ is odd. Mapping $\delta : S \cup T \to M'$, $\delta = (x, \varnothing, \varnothing)$. Notice that the δ^{-1} is a bijection to δ, therefore we have our desired bijective proof. ■

Definition 6.3. Fix(f) is the number of fixed points of the involution $f : Y \to Y$. Therefore, $f(x) = x$, where $x \in Y$

We get a very interesting property from fixed points and involutions is the following.

Proposition 6.4. *Suppose* $f : Y \to Y$ *is an involution, with some finite set* Y*. Then,* $|Y| \equiv |Fix(f)| \pmod 2$

Proof. Define $S_x = \{x, f(x)\}$. Using this, we can partition Y into disjoint subsets containing 1 or 2 elements, yielding the number of fixed points of size equal to the number of S_x's that have a size of 1. If we denote the number of fixed points as m, and the number of S_x's of size 2 as n, we get that $|Y| \equiv m + 2n \pmod 2$. Since $2n \equiv 0 \pmod 2$, we get $m \equiv |Y| \pmod 2$, satisfying the condition. ■

This can be used to prove many consequences such as Fermat's theorem on the sum of two squares. Though, this is beyond the scope of this paper.

Proposition 6.5. *Suppose $f : Y \to Y$ is an involution, with some finite set Y. Then,* $|Y| \equiv |Fix(f)| \pmod 2$

6.1. **The Involution Principle.** Suppose there is some finite set X, which is partitioned into two disjoint subsets X^+ and X^-. These are called the "positive" and "negative" parts of the set X respectively.

Then, $X = X^+ \cup X^-$. τ is an involution on X such that if $\tau(X) = y$ and $x \neq y$, then one of x and y must be in X^+ and X^-. In other words, both x and y must be in different subsets and both must be in one of the subsets (since they are disjoint) and if $\tau(x) = x$, then $x \in X^+$.

Notice that for each non-fixed point, we have one in X^+ and in X^-. This can be described by the weight function

$$\omega(x) = \begin{cases} 1, & x \in X^+ \\ -1, & x \in X^- \end{cases} .$$

We can see from this can $|\mathrm{Fix}(\tau)| = |X^+| - |X^-|$.

Now suppose that we have another set Y, that is broken up into subsets Y^+ and Y^-, analogous to X. Let σ be the involution on Y, analogous to τ. Suppose we also have a bijection $f : X \to Y$. The result of this bijection is that $|\mathrm{Fix}(\tau)| = |\mathrm{Fix}(\sigma)|$. This is what is known as the *involution principle.*

We wish to find a bijection between $\mathrm{Fix}(\tau)$ and $\mathrm{Fix}(\sigma)$, that results from our original bijection.

6.2. **Rogers-Ramanujan Identity.** The involution principle (also known as the *Garsia-Milne involution principle*), was used to construct a bijection for the Rogers-Ramanujan identity.

Theorem 6.6. *The partitions of n with parts congruent to 1 or* $4 \pmod 5$ *are equal to (i.e. a bijection to) partitions of n with parts differing by at least 2.*

We will let $\mathbf{D}$ denote the class of partitions with parts different by at least 2, and E denote the class of partitions with parts congruent to 1 or 4 (mod 5). Each E_i will denote parts congruent to $i \pmod 5$. For example, E_2 will denote parts congruent to 2 (mod 5). $\mathbf{E} = E_1 \times E_2 \times E_3 \times E_4 \times E_5$. $\mathbf{E}'$ or E_i' will denote that the parts must be distinct.

The idea of this proof was to use 3 different bijections to prove the identity. One of them was by Schur, and the other two were thought up by A. M. Garsia and S. C. Milne. Notice that we want a weight-preserving bijection between the class $\mathbf{D}$ and the class $E_1 \times E_4$.

6.2.1. *The Schur Involution.*

Definition 6.7 (Special Schur Pair of Type A)**.** Given a positive integer n, the partitions

$$(1)\ \mathbf{E}' = (2n, 2n-1, \ldots, n+1), \tag{6.2}$$

$$(2)\ \mathbf{D} = (2n-1, 2n-3, \ldots, 3, 1)$$

will be denoted by SSA(n)

Definition 6.8 (Special Schur Pair of Type B)**.** Given an non-negative integer m, the partitions

$$(1)\ \mathbf{E}' = (2m-1, 2m-2 \ldots, m), \tag{6.3}$$

$$(2)\ \mathbf{D} = (2m-1, 2m-3, \ldots, 3, 1)$$

will be denoted by SSB(m)

Note that since m can be equal to 0, SSB(0) $= (\varnothing, \varnothing)$.

Now let $\mathbf{B} = \mathbf{E}' \times \mathbf{D} \times \mathbf{E}$. The positive and negative of $\mathbf{B}$ will be denoted as $\mathbf{B}^+$ and $\mathbf{B}^-$ respectively. Suppose we have a subset of B with the form (SSA(n), $\mathbf{E}$) or (SSB(m), $\mathbf{E}$). This can be denoted as $\mathbf{SS} \times \mathbf{E}$, known as the *Schur special set.* Similarly, the positive and negative parts of $\mathbf{SS}^+ \times \mathbf{E}$ will be denoted as $\mathbf{SS}^- \times \mathbf{E}$ respectively.

The Schur bijection is a permutation of $\mathbf{B}$ onto itself, which we will denote as $S(\beta)$. $S(\beta)$ satisfies the following properties:

(1) $S(\beta)$ operates only on the first two components of a triplet $(\mathbf{E}', D, \mathbf{E})$
(2) $S(\beta)$ is an involution
(3) Fix($S(\beta)$ are the elements of $\mathbf{SS} \times \mathbf{E}$.
(4) $S(\beta)$ bijectively interchanges the sets $\mathbf{B}^+$ - $\mathbf{SS}^+ \times \mathbf{E}$ and $\mathbf{B}^-$ - $\mathbf{SS}^- \times \mathbf{E}$.

We can represent this using the following picture.

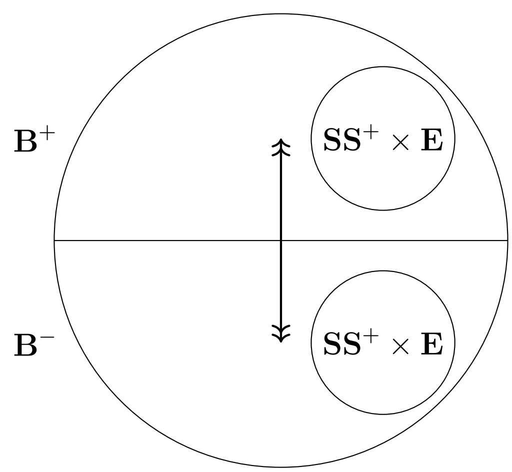

Figure 5. $S(\beta)$ Bijection

This is just the introduction to the Rogers-Ramanujan bijection and the Schur Involution. Check out [GM81] for further exploration on the bijections leading up to the proof of the Rogers-Ramanujan Identity.

7. Matrices and Determinants

7.1. Calculus of Finite Differences.

Definition 7.1. (Characteristic) Let 1 be the multiplicative identity and let 0 be the additive identity. Define the *characteristic* of the field K to be the smallest positive integer n such that the sum of 1 added to itself n times equals zero. If no such n exists then, the characteristic is 0.

Let there be a function $f : \mathbb{Z} \to K$, where K is a field with characteristic 0. Define the function $\Delta f(n) = f(n+1) - f(n)$. This is known as the first difference of f. Δ is known as the (first) *difference operator*. We can get the kth difference operator by the equation

$$\Delta^k f = \Delta(\Delta^{k-1} f).$$

If we define $Ef(n) = f(n+1)$ for the operator E, it can be seen that $\Delta = E - 1$. Then,

$$\begin{aligned}\Delta^k f(n) &= (E-1)^k f(n) \\ &= \sum_{i=0}^{k} (-1)^{k-1} \binom{k}{i} E^i f(n) \\ &= \sum_{i=0}^{k} (-1)^{k-1} \binom{k}{i} f(n-1) \\ &= \sum_{i=0}^{k} (-1)^{k-1} \binom{k}{i} E^i f(n) \end{aligned} \tag{7.1}$$

For $n = 0$, we get the equation

$$\Delta^k f(0) = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} f(i), \tag{7.2}$$

giving us a closed form for $\Delta^k f(0)$ in terms of $f(i)$. Note that we can express $f(n)$ in terms of $f(0)$ as,

$$f(n) = \sum_{i=0}^{n} \binom{n}{i} \Delta^i f(0). \tag{7.3}$$

7.2. Inclusion-Exclusion Revisited.

It is important to note that the Principle of Inclusion-Exclusion can be described in a more algebraic way than the

Suppose a function $f_=$ satisfies $f_=(T) = f_=(T')$, when $|T| = |T'|$. Let us set $a(n-i) = f_=(T)$ and $b(n-i) = f_\leq(T)$, when $|T| = i$.

Therefore, we get the following equations, analogous to equations 3.2 and 3.3 respectively,

$$b(m) = \sum_{i=0}^{m} \binom{m}{i} a(i), \tag{7.4}$$

$$a(m) = \sum_{i=0}^{m} \binom{m}{i} (-1)^{m-i} b(i), \tag{7.5}$$

for $0 \leq m \leq n$.

Using calculus of finite differences, this is

$$a(m) = \Delta^m b(0),$$

for $0 \leq m \leq n$.

This can be represented by a matrix A and its inverse A'. Suppose matrix A is a $(n+1) \times (n+1)$ matrix with entries (i, j) satisfying $(i, j) = \binom{j}{i}$ for $0 \leq i, j \leq n$. Then, A' must have entries (i, j) satisfying $(i, j) = (-1)^{j-i}\binom{j}{i}$ for $0 \leq i, j \leq n$.

Example. $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ since for (0, 1) and (2, 1), $j - i$ is odd resulting in $(-1)^{j-i}$ = -1.

Proposition 7.2. *Let S_n be the set of n properties that elements of B_n may or may not have, for $n \in \mathbb{N}$. Let's suppose that for every $T \subseteq S_n$, the number $x \in f \leq (B_n)$ for the function f depends on $|T|$ and not on n. If we let $b(n) = |B_n|$, and let $a(n) = f_=(\varnothing)$ (i.e. the number of objects in B_n that have none of the properties in S_n), then $a(n) = \Delta^n b(0)$.*

7.3. **Descents.** Define the *descent* of $\omega = \omega_1\omega_2 \cdots \omega_n$ for $1 \leq i \leq n-1$, to be $\omega_i \geq \omega_{i-1}$.

Definition 7.3. (Descent Set) The *descent set* $D(\omega)$ is defined as $D(\omega) = \{i : \omega_i \geq \omega_{i+1}\} \subseteq [n-1]$.

Suppose $S \subseteq [n-1]$. Then, let

$$\alpha(S) = |\{\omega \in \mathfrak{S}_n \ : D(\omega) \subseteq S\}, \tag{7.6}$$

$$\beta(S) = |\{\omega \in \mathfrak{S}_n \ : D(\omega) = S\}. \tag{7.7}$$

Analogous to 3.5 and 3.6, we can express $\alpha(S)$ and $\beta(S)$ as

$$\alpha(S) = \sum_{T \subseteq S} \beta(S), \tag{7.8}$$

and

$$\beta(S) = \sum_{T \subseteq S} (-1)^{|S-T|}\alpha(S), \tag{7.9}$$

or more specifically,

$$\alpha_n(S) = \sum_{T \subseteq S} \beta_n(S),$$

and

$$\beta_n(S) = \sum_{T \subseteq S} (-1)^{|S-T|}\alpha_n(S).$$

Proposition 7.4. *For some $S \subseteq [n-1]$, let $S = \{s_1, s_2, \ldots, s_k\}$. Then,*

$$\alpha(S) = \binom{n}{s_1, s_2 - s_1, s_3 - s_2, \ldots n - s_k}. \tag{7.10}$$

Proof. It can be see that we can choose $\omega_1 \leq \omega_2 \leq \cdots \omega_{s_1}$ in $\binom{n}{s_1}$ ways. Essentially, we can choose $\omega_{s_i+1} \leq \omega_{s_i+2} \leq \cdots \omega_{s_{i+1}}$ in $\binom{n-s_i}{s_{i+1}-s_i}$ ways. Since $\alpha(S)$ is the product of these $\binom{n-s_i}{s_{i+1}-s_i}$'s, we get the multinomial

$$\alpha(S) = \binom{n}{s_1, s_2 - s_1, s_3 - s_2, \ldots n - s_k},$$

as desired. ∎

From our previous results, we get that

$$\beta_n(S) \sum_{1 \leq i_1 \leq \cdots \leq i_j \leq k} (-1)^{k-j} \binom{n}{s_{i_1}, s_{i_2} - s_{i_1}, \ldots, n - s_{i_j}}. \tag{7.11}$$

This is not the only way 7.11 can be written as. Suppose we have the function defined by $[0, k+1] \times [[0, k+1]$, so that $f(i,i) = 1$, and $f(i,j) = 0$ for $i \leq j$. Then the sum is

$$A_k = \sum_{1 \leq i_1 \leq \cdots \leq i_j \leq k} (-1)^{k-j} f(0, i_1) f(i_1, i_2) \cdots f(i_j, k+1) \tag{7.12}$$

The terms in this sum are just the nonzero terms in the expansion of the determinant of $(i,j) \to f(i, j+1), (i,j) \in [0,k] \times [0,k]$. If $f(i,j) = \frac{1}{(s_j - s_i)!}$, then we get that

$$\beta_n(S) = n! \begin{vmatrix} a_1 & a_2 & a_3 & \cdots \\ a_i & a_j & a_k & \cdots \\ a_p & a_q & a_r & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} \tag{7.13}$$

where each term a_i is $\frac{1}{(s_{j+1} - s_i)!}$, for $(i,j) \in [0,k] \times [0,k]$

In other words,

$$\beta_n(S) = n! \cdot det \left[\frac{1}{(s_{j+1} - s_i)!} \right] = det \left[\binom{n - s_i}{s_{j+1} - s_i} \right], \tag{7.14}$$

where "det" denotes the determinant of the matrix, and $(i,j) \in [0,k] \times [0,k]$

Definition 7.5. The *q-analogue* of a theorem is a generalization using a parameter q, such that it returns the original theorem, identity or expression in the limit as $q \to 1$

Using this definition, we can now obtain a q-analogue of the previous work. We need some $s(\omega); \omega \in \mathfrak{S}_n$, with $D(\omega) \subseteq S$ such that

$$\sum_{\omega \in \mathfrak{S}_n} q^{s(\omega)} = \binom{n}{s_1, s_2 - s_1, \ldots, n - s_k},$$

where $S = 1 \leq s_1 \leq s_2 \leq \cdots \leq s_k \leq n-1$

Proposition 7.6. *Let S be a set of properties $\{P_1, P_2 \ldots P_n\}$ and let $T \subseteq S$, where $T = \{P_{s_1}, P_{s_2} \ldots P_{s_k}\}$. Suppose that*

$$f_{\leq}(T) = h(n)[e(s_0, s_1)][e(s_1, s_2)] \cdots [e(s_k, s_{k+1}])$$

for a function $h(n)$ and e in $\mathbb{N} \times \mathbb{N}$. We set $s_0 = 0, s_{k+1} = n, e(i,i) = 1$ and $e(i,j) = 0$ for $j \leq i$. This results in

$$f_{=}(T) = h(n) det[e(s_i, s_{j+1})] \tag{7.15}$$

from 0 to k.

7.4. **Lattice Paths.** Let $S \subseteq \mathbb{Z}^d$, where a lattice path L in $\mathbb{Z}^d$ has length of k, with steps in S. Let the sequence of the lattice path be $\{s_0, s_1, \dots, s_k\}$, so that any $s_{i+1} - s_i$ is a step in S. The following figure is the lattice path from $(0,0)$ to $(6,6)$, with length 12.

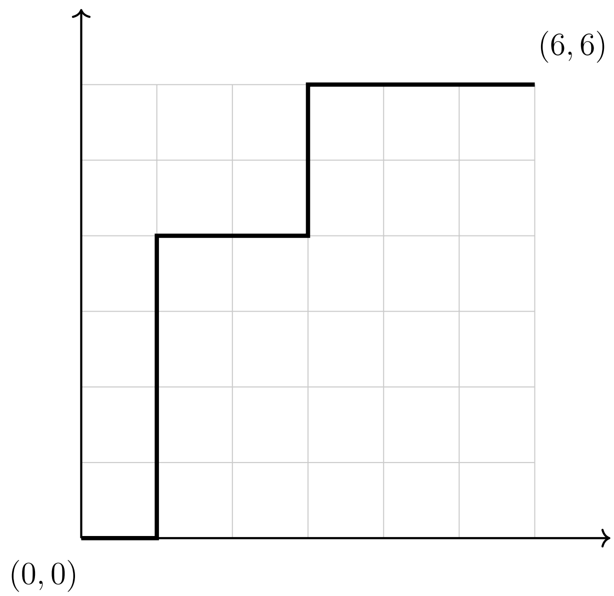

Figure 6. Lattice Path from (0,0) to (6,6)

Proposition 7.7. *Let $v = (a_1, \dots, a_d) \in \mathbb{N}^d$, and let e_i be the ith unit vector in $\mathbb{Z}^d$. The number of lattice paths in $\mathbb{Z}^d$ (from the origin to the point v) is given by $\binom{a_1+\cdots+a_d}{a_1,\dots,a_d}$.*

Proof. The sequences of steps are just $a_i e_i$'s. Since we would like to find the permutations of these steps $\mathfrak{S}_n$, we get

$$\mathfrak{S}_n = \binom{a_1 + \cdots + a_d}{a_1, \dots, a_d}$$

■

Definition 7.8. (Tuple) A tuple is a finite sequence of objects. An n-tuple is a tuple of n elements, where n is a non-negative integer.

Definition 7.9. An n-path is an n-tuple $\mathbf{L} = (L_1, L_2, \dots, L_n)$ of lattice paths.

Example. The lattice path [(1, 2), (2, 2), (2, 4), (3, 4), (3, 5), (4, 5), (4, 1)] can be drawn as

Definition 7.10. For $\alpha, \beta, \gamma, \delta \in \mathbb{N}^b$, let $\mathbf{L}$ be of type $(\alpha, \beta, \gamma, \delta)$ if and only if $\mathbf{L}$ goes from (β_i, δ_i) to (α_i, γ_i).

(n-path intersection)

Definition 7.11. An n-path is considered *intersecting* if L_i and L_j have some point(s) in common, where $i \neq j$.

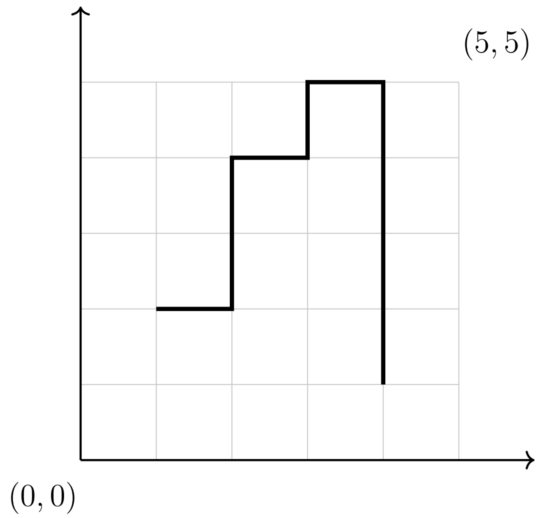

Figure 7. Example: Lattice Path from (1,2) to (4, 1)

Definition 7.12. (Weight) The *weight* of L, denoted as $\Lambda(L)$ is defined to be the product of the horizontal steps of L.

Suppose we have some $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_n\} \in \mathbb{N}^n$, and $\omega \in \mathfrak{S}_n$, then $\omega(\alpha) = \{\alpha_{\omega(1)}, \ldots, \alpha_{\omega(n)}\}$. Let the path from (β_i, δ_i) to (α_i, γ_i) be in the set of all n-paths of type $(\alpha, \beta, \gamma, \delta)$, denoted by $\mathcal{A}$. If we let $m = \alpha_i - \beta_i$. Then, there is exactly one horizontal step from $(j + \beta_i - 1, k_j)$ to $(j + \beta_i, k_j)$ (i.e. a horizontal step of length 1) for $1 \le j \le n$ and k_r, where $\delta_i \le k_r \le \gamma_i$ and $k_r \ge k_{r+1}$. If $h(m; \gamma_i, \delta_i) = \sum x_{k_1} \cdots x_{k_m}$, it can be seen that

$$A(\alpha, \beta, \gamma, \delta) = \prod_{i=1}^{n} h(\alpha_i - \beta_i, \gamma_i, \delta_i), \tag{7.16}$$

where $A(\alpha, \beta, \gamma, \delta)$ is the sum of the weights of $\mathcal{A}(\alpha, \beta, \gamma, \delta)$. We get 7.16 as the result because we are summing over the multiple possible sequences of k_r.

Theorem 7.13. *Let $\mathcal{B}(\alpha, \beta, \gamma, \delta)$ be the set of n-paths are a non-intersecting, and let $B(\alpha, \beta, \gamma, \delta)$ be the sum of the weights of $\mathcal{B}(\alpha, \beta, \gamma, \delta)$. Suppose we have $(\alpha, \beta, \gamma, \delta) \in \mathbb{N}^n$ such that for $\omega \in \mathfrak{S}_n, \beta(\omega(\alpha), \beta, \gamma, \omega(\delta)$, for $\omega \neq I$, where I is the identity permutation. Then,*

$$B(\alpha, \beta, \gamma, \delta) = det[h(\alpha_j - \beta_i, \gamma_i, \delta_j)] \tag{7.17}$$

Proof. Let $A(\omega) = A(\omega(\alpha), \beta, \gamma, \omega(\delta))$. We will construct a bijection from $\mathbf{L}$ to $\mathbf{L}'$

(1) $L'' = L$
(2) $\Lambda(L') = \Lambda(L)$
(3) If $L \in A_u$ and $L' \in A_v$ then sgn u = -sgn v.

"sgn" is known as *signum* and it is a function that indicates the sign of a number. It is defined as 1 for positive numbers, 0 for zero, and -1 for negative numbers.

If we group terms in 7.17 into $(\mathbf{L}, \mathbf{L}')$ of intersecting n-paths, then we get that all unnecessary terms cancel revealing $B(\alpha, \beta, \gamma, \delta)$. ∎

8. Further Enquiries

The story of sieve methods does not end here. As discussed, sieve methods have a very close tie with the Principle of Inclusion-Exclusion, where instead of directly counting the objects, we construct a system to strategically eliminate objects that do not satisfy our conditions. Similarly to this, sieve methods have applications in probabilistic methods, where instead of relying on randomness, we can again systematically remove objects that do not meet our conditions. This provides us with asymptotic expressions, alike the formula we derived for derangements. Sieve methods also are very useful in filtering through sets and objects providing us with specific combinatorial structures and patterns. An example of this is the Turán sieve, as it is derived directly from the Principle of Inclusion-Exclusion. In essence, the Turán sieve estimates the size of sets of positive integers, expressed by congruences. For further knowledge on sieve methods, check out [LM05], and for more thoughts on enumerative combinatorics, check out [Sta11, SF99].

Acknowledgements

The author would like to thank Simon Rubinstein-Salzedo and Andrew Lin for making this paper possible.

References

[GM81] Adriano M Garsia and Stephen C Milne. A rogers-ramanujan bijection. *Journal of Combinatorial Theory, Series A*, 31(3):289–339, 1981.

[LM05] Yu-Ru Liu and M Ram Murty. Sieve methods in combinatorics. *Journal of Combinatorial Theory, Series A*, 111(1):1–23, 2005.

[SF99] Richard P Stanley and Sergey Fomin. Enumerative combinatorics. vol. 2, volume 62 of. *Cambridge studies in advanced mathematics*, page 25, 1999.

[Sta11] Richard P Stanley. Enumerative combinatorics volume 1 second edition. *Cambridge studies in advanced mathematics*, 2011.

Euler Circle, Folsom, CA 95630
Email address: rven2308@gmail.com