Lecture Notes on Discrete Mathematics

A. K. Lal

September 26, 2012

Contents

Chapter 1

Preliminaries

We will use the following notation throughout these notes.

1. The empty set, denoted $\emptyset$, is a set that has no element.
2. $\mathbb{N} := \{0, 1, 2, \ldots\}$, the set of Natural numbers;
3. $\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, the set of Integers;
4. $\mathbb{Q} := \{\frac{p}{q} : p, q \in \mathbb{Z},\ q \neq 0\}$, the set of Rational numbers;
5. $\mathbb{R} :=$ the set of Real numbers; and
6. $\mathbb{C} :=$ the set of Complex numbers.

For the sake of convenience, we have assumed that the integer 0, is also a natural number. This chapter will be devoted to understanding set theory, relations, functions and the principle of mathematical induction. We start with basic set theory.

1.1 Basic Set Theory

We have already seen examples of sets, such as $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ at the beginning of this chapter. For example, one can also look at the following sets.

Example 1.1.1. *1.* $\{1, 3, 5, 7, \ldots\}$*, the set of odd natural numbers.*

2. $\{0, 2, 4, 6, \ldots\}$*, the set of even natural numbers.*

3. $\{\ldots, -5, -3, -1, 1, 3, 5, \ldots\}$*, the set of odd integers.*

4. $\{\ldots, -6, -4, -2, 0, 2, 4, 6, \ldots\}$*, the set of even integers.*

5. $\{0, 1, 2, \ldots, 10\}$.

6. $\{1, 2, \ldots, 10\}$.

7. $\mathbb{Q}^+ = \{x \in \mathbb{Q} : x > 0\}$, *the set of positive rational numbers.*

8. $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$, *the set of positive real numbers.*

9. $\mathbb{Q}^* = \{x \in \mathbb{Q} : x \neq 0\}$, *the set of non-zero rational numbers.*

10. $\mathbb{R}^* = \{x \in \mathbb{R} : x \neq 0\}$, *the set of non-zero real numbers.*

We observe that the sets that appear in Example 1.1.1 have been obtained by picking certain elements from the sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. These sets are example of what are called "subsets of a set", which we define next. We also define certain operations on sets.

Definition 1.1.2 (Subset, Complement, Union, Intersection)**.** *1. Let A be a set. If B is a set such that each element of B is also an element of the set A, then B is said to be a subset of the set A, denoted $B \subseteq A$.*

2. *Two sets A and B are said to be equal if $A \subseteq B$ and $B \subseteq A$, denoted $A = B$.*

3. *Let A be a subset of a set Ω. Then the complement of A in Ω, denoted A', is a set that contains every element of Ω that is not an element of A. Specifically, $A' = \{x \in \Omega : x \notin A\}$.*

4. *Let A and B be two subsets of a set Ω. Then their*

   (a) *union, denoted $A \cup B$, is a set that exactly contains all the elements of A and all the elements of B. To be more precise, $A \cup B = \{x \in \Omega : x \in A \text{ or } x \in B\}$.*

   (b) *intersection, denoted $A \cap B$, is a set that exactly contains those elements of A that are also elements of B. To be more precise, $A \cap B = \{x \in \Omega : x \in A \text{ and } x \in B\}$.*

Example 1.1.3. *1. Let A be a set. Then $A \subseteq A$.*

2. *The empty set is a subset of every set.*

3. *Observe that $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$.*

4. *As mentioned earlier, all examples that appear in Example 1.1.1 are subsets of one or more sets from $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$.*

5. *Let A be the set of odd integers and B be the set of even integers. Then $A \cap B = \emptyset$ and $A \cup B = \mathbb{Z}$. Thus, it also follows that the complement of A, in $\mathbb{Z}$, equals B and vice-versa.*

6. *Let $A = \{\{b, c\}, \{\{b\}, \{c\}\}\}$ and $B = \{a, b, c\}$ be subsets of a set Ω. Then $A \cap B = \emptyset$ and $A \cup B = \{a, b, c, \{b, c\}, \{\{b\}, \{c\}\}\}$.*

Definition 1.1.4 (Cardinality)**.** *A set A is said to have finite cardinality, denoted $|A|$, if the number of distinct elements in A is finite, else the set A is said to have infinite cardinality.*

Example 1.1.5. *1. The cardinality of the empty set equals* 0. *That is,* $|\emptyset| = 0$.

2. *Fix a positive integer* n *and consider the set* $A = \{1, 2, \ldots, n\}$. *Then* $|A| = n$.

3. *Let* $S = \{2x \in \mathbb{Z} : x \in \mathbb{Z}\}$. *Then* S *is the set of even integers and it's cardinality is infinite.*

4. *Let* $A = \{a_1, a_2, \ldots, a_m\}$ *and* $B = \{b_1, b_2, \ldots, b_n\}$ *be two finite subsets of a set* Ω, *with* $|A| = m$ *and* $B| = n$. *Also, assume that* $A \cap B = \emptyset$. *Then, by definition it follows that*

$$A \cup B = \{a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_n\}$$

and hence $|A \cup B| = |A| + |B|$.

5. *Let* $A = \{a_1, a_2, \ldots, a_m\}$ *and* $B = \{b_1, b_2, \ldots, b_n\}$ *be two finite subsets of a set* Ω. *Then* $|A \cup B| = |A| + |B| - |A \cap B|$. *Observe that Example 1.1.5.4 is a particular case of this result, when* $A \cap B = \emptyset$.

6. *Let* $A = \{\{a_1\}, \{a_2\}, \ldots, \{a_m\}\}$ *be a subset of a set* Ω. *Now choose an element* $a \in \Omega$ *such that* $a \neq a_i$, *for any* $i, 1 \leq i \leq n$. *Then verify that the set* $B = \{S \cup \{a\} : S \in A\}$ *equals* $\{\{a, a_1\}, \{a, a_2\}, \ldots, \{a, a_m\}\}$. *Also, observe that* $A \cap B = \emptyset$ *and* $|B| = |A|$.

Exercise 1.1.6. *1. Does there exist unique sets* X *and* Y *such that* $X - Y = \{1, 3, 5, 7\}$ *and* $Y - X = \{2, 4, 8\}$*?*

2. *In a class of* 60 *students, all the students play either football or cricket. If* 20 *students play both football and cricket, determine the number of players for each game if the number of students who play football is*

   (a) 14 *more than the number of students who play cricket.*

   (b) *exactly* 5 *times more than the number of students who play only cricket.*

   (c) *a multiple of* 2 *and* 3 *and leaves a remainder of* 3 *when divided by* 5.

   (d) *is a factor of* 90 *and the number of students who play cricket is a factor of* 70.

Definition 1.1.7 (Power Set)**.** *Let* A *be a subset of a set* Ω. *Then a set that contains all subsets of* A *is called the power set of* A *and is denoted by* $\mathcal{P}(A)$ *or* 2^A.

Example 1.1.8. *1. Let* $A = \emptyset$. *Then* $\mathcal{P}(\emptyset) = \{\emptyset, A\} = \{\emptyset\}$.

2. *Let* $A = \{\emptyset\}$. *Then* $\mathcal{P}(A) = \{\emptyset, A\} = \{\emptyset, \{\emptyset\}\}$.

3. *Let* $A = \{a, b, c\}$. *Then* $\mathcal{P}(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$.

4. *Let* $A = \{\{b, c\}, \{\{b\}, \{c\}\}\}$. *Then* $\mathcal{P}(A) = \{\emptyset, \{\{b, c\}\}, \{\{\{b\}, \{c\}\}\}, \{\{b, c\}, \{\{b\}, \{c\}\}\}\ \}$.

1.2 Properties of Integers

Axiom 1.2.1 (Well-Ordering Principle)**.** *Every non-empty subset of natural numbers contains its least element.*

We will use Axiom 1.2.1 to prove the weak form of the principle of mathematical induction. The proof is based on contradiction. That is, suppose that we need to prove that "whenever the statement P holds true, the statement Q holds true as well". A proof by contradiction starts with the assumption that "the statement P holds true and the statement Q does not hold true" and tries to arrive at a contradiction to the validity of the statement P being true.

Theorem 1.2.2 (Principle of Mathematical Induction: Weak Form)**.** *Let $P(n)$ be a statement about a positive integer n such that*

1. *$P(1)$ is true, and*

2. *$P(k+1)$ is true whenever one assumes that $P(k)$ is true.*

Then $P(n)$ is true for all positive integer n.

Proof. On the contrary, assume that there exists $n_0 \in \mathbb{N}$ such that $P(n_0)$ is not true. Now, consider the set

$$S = \{m \in \mathbb{N} \,:\, P(m) \text{ is false }\}.$$

As $n_0 \in S$, $S \neq \emptyset$. So, by Well-Ordering Principle, S must have a least element, say N. By assumption, $N \neq 1$ as $P(1)$ is true. Thus, $N \geq 2$ and hence $N - 1 \in \mathbb{N}$.

Therefore, the assumption that N is the least element in S and S contains all those $m \in \mathbb{N}$, for which $P(m)$ is false, one deduces that $P(N-1)$ holds true as $N - 1 < N \leq 2$. Thus, the implication "$P(N-1)$ is true" and Hypothesis 2 imply that $P(N)$ is true.

This leads to a contradiction and hence our first assumption that there exists $n_0 \in \mathbb{N}$, such that $P(n_0)$ is not true is false. ■

Example 1.2.3. *1. Prove that* $1 + 2 + \cdots + n = \dfrac{n(n+1)}{2}$.

Solution: *Verify that the result is true for $n = 1$. Hence, let the result be true for n. Let us now prove it for $n+1$. That is, one needs to show that* $1 + 2 + \cdots + n + (n+1) = \dfrac{(n+1)(n+2)}{2}$.

Using Hypothesis 2,

$$1 + 2 + \cdots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{n+1}{2}(n+2).$$

Thus, by the principle of mathematical induction, the result follows.

2. *Prove that* $1^2 + 2^2 + \cdots + n^2 = \dfrac{n(n+1)(2n+1)}{6}$.

Solution: *The result is clearly true for* $n = 1$. *Hence, let the result be true for* n *and one needs to show that* $1^2 + 2^2 + \cdots + n^2 + (n+1)^2 = \dfrac{(n+1)(n+2)\,(2(n+1)+1)}{6}$.

Using Hypothesis 2,

$$\begin{aligned} 1^2 + 2^2 + \cdots + n^2 + (n+1)^2 &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \\ &= \frac{n+1}{6}\,(n(2n+1) + 6(n+1)) \\ &= \frac{n+1}{6}\left(2n^2 + 7n + 6\right) = \frac{(n+1)(n+2)\,(2n+3)}{6}. \end{aligned}$$

Thus, by the principle of mathematical induction, the result follows.

3. *Prove that for any positive integer* n, $1 + 3 + \cdots + (2n-1) = n^2$.

Solution: *The result is clearly true for* $n = 1$. *Let the result be true for* n. *That is,* $1 + 3 + \cdots + (2n-1) = n^2$. *Now, we see that*

$$1 + 3 + \cdots + (2n-1) + (2n+1) = n^2 + (2n+1) = (n+1)^2.$$

Thus, by the principle of mathematical induction, the result follows.

4. **AM-GM Inequality:** *Let* $n \in \mathbb{N}$ *and suppose we are given real numbers* $a_1 \geq a_2 \geq \cdots \geq a_n \geq 0$. *Then*

$$\textit{Arithmetic Mean (AM)} := \frac{a_1 + a_2 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 \cdot a_2 \cdot \cdots \cdot a_n} =: \textit{ (GM) Geometric Mean.}$$

Solution: *The result is clearly true for* $n = 1, 2$. *So, we assume the result holds for any collection of* n *non-negative real numbers. Need to prove* $AM \geq GM$, *for any collections of non-negative integers* $a_1 \geq a_2 \geq \cdots \geq a_n \geq a_{n+1} \geq 0$.

So, let us assume that $A = \dfrac{a_1 + a_2 + \cdots + a_n + a_{n+1}}{n+1}$. *Then, it can be easily verified that* $a_1 \geq A \geq a_{n+1}$ *and hence* $a_1 - A,\ A - a_{n+1} \geq 0$. *Thus,* $(a_1 - A)(A - a_{n+1}) \geq 0$. *Or equivalently,*

$$A(a_1 + a_{n+1} - A) \geq a_1 a_{n+1}. \tag{1.1}$$

Now, let us assume that the AM-GM inequality holds for any collection of n *non-negative numbers. Hence, in particular, for the collection* $a_2, a_3, \ldots, a_n, a_1 + a_{n+1} - A$. *That is,*

$$AM = \frac{a_2 + \cdots + a_n + (a_1 + a_{n+1} - A)}{n} \geq \sqrt[n]{a_2 \cdots a_n \cdot (a_1 + a_{n+1} - A)} = \ GM. \tag{1.2}$$

But $\dfrac{a_2 + a_3 + \cdots + a_n + (a_1 + a_{n+1} - A)}{n} = A$. *Thus, by Equation (1.1) and Equation (1.2), one has*

$$A^{n+1} \geq (a_2 \cdot a_3 \cdot \cdots \cdot a_n \cdot (a_1 + a_{n+1} - A)) \cdot A \geq (a_2 \cdot a_3 \cdot \cdots \cdot a_n)\, a_1 a_{n+1}.$$

Therefore, we see that by the principle of mathematical induction, the result follows.

5. *Fix a positive integer n and let A be a set with $|A| = n$. Then prove that $\mathcal{P}(A) = 2^n$.*

   Solution: *Using Example 1.1.8, it follows that the result is true for $n = 1$. Let the result be true for all subset A, for which $|A| = n$. We need to prove the result for a set A that contains $n + 1$ distinct elements, say $a_1, a_2, \ldots, a_{n+1}$.*

   Let $B = \{a_1, a_2, \ldots, a_n\}$. Then $B \subseteq A$, $|B| = n$ and by induction hypothesis, $|\mathcal{P}(B)| = 2^n$. Also, $\mathcal{P}(B) = \{S \subseteq \{a_1, a_2, \ldots, a_n, a_{n+1}\} : a_{n+1} \notin S\}$. Therefore, it can be easily verified that

$$\mathcal{P}(A) = \mathcal{P}(B) \cup \{S \cup \{a_{n+1}\} : S \in \mathcal{P}(B)\}.$$

   Also, note that $\mathcal{P}(B) \cap \{S \cup \{a_{n+1}\} : S \in \mathcal{P}(B)\} = \emptyset$, as $a_{n+1} \in S$, for all $S \in \mathcal{P}(B)$. Hence, using Examples 1.1.5.4 and 1.1.5.6, we see that

$$|\mathcal{P}(A)| = |\mathcal{P}(B)| + |\{S \cup \{a_{n+1}\} : S \in \mathcal{P}(B)\}| = |\mathcal{P}(B)| + |\mathcal{P}(B)| = 2^n + 2^n = 2^{n+1}.$$

   Thus, the result holds for any set that consists of $n + 1$ distinct elements and hence by the principle of mathematical induction, the result holds for every positive integer n.

We state a corollary of the Theorem 1.2.2 without proof. The readers are advised to prove it for the sake of clarity.

Corollary 1.2.4 (Principle of Mathematical Induction)**.** *Let $P(n)$ be a statement about a positive integer n such that for some fixed positive integer n_0,*

1. *$P(n_0)$ is true,*
2. *$P(k + 1)$ is true whenever one assumes that $P(n)$ is true.*

Then $P(n)$ is true for all positive integer $n \geq n_0$.

We are now ready to prove the strong form of the principle of mathematical induction.

Theorem 1.2.5 (Principle of Mathematical Induction: Strong Form)**.** *Let $P(n)$ be a statement about a positive integer n such that*

1. *$P(1)$ is true, and*
2. *$P(k + 1)$ is true whenever one assume that $P(m)$ is true, for all m, $1 \leq m \leq k$.*

Then, $P(n)$ is true for all positive integer n.

Proof. Let $R(n)$ be the statement that "the statement $P(m)$ holds, for all positive integers m with $1 \leq m \leq n$". We prove that $R(n)$ holds, for all positive integers n, using the weak-form of mathematical induction. This will give us the required result as the statement "$R(n)$ holds true" clearly implies that "$P(n)$ also holds true".

As the first step of the induction hypothesis, we see that $R(1)$ holds true (already assumed in the hypothesis of the theorem). So, let us assume that $R(n)$ holds true. We need to prove that $R(n+1)$ holds true.

The assumption that $R(n)$ holds true is equivalent to the statement "$P(m)$ holds true, for all $m,\ 1 \leq m \leq n$". Therefore, by Hypothesis 2, $P(n+1)$ holds true. That is, the statements "$R(n)$ holds true" and "$P(n+1)$ holds true" are equivalent to the statement "$P(m)$ holds true, for all $m,\ 1 \leq m \leq n+1$". Hence, we have shown that $R(n+1)$ holds true. Therefore, we see that the result follows, using the weak-form of the principle of mathematical induction. ■

We state a corollary of the Theorem 1.2.5 without proof.

Corollary 1.2.6 (Principle of Mathematical Induction)**.** *Let $P(n)$ be a statement about a positive integer n such that for some fixed positive integer n_0,*

1. *$P(n_0)$ is true,*
2. *$P(k+1)$ is true whenever one assume that $P(m)$ is true, for all $m,\ n_0 \leq m \leq k$.*

Then $P(n)$ is true for all positive integer $n \geq n_0$.

Remark 1.2.7 (Pitfalls)**.** *Find the error in the following arguments:*

1. *If a set of n balls contains a green ball then all the balls in the set are green.*
   Solution: *If $n = 1$, we are done. So, let the result be true for any collection of n balls in which there is at least one green ball.*

   So, let us assume that we have a collection of $n+1$ balls that contains at least one green ball. From this collection, pick a collection of n balls that contains at least one green ball. Then by the induction hypothesis, this collection of n balls has all green balls.

   Now, remove one ball from this collection and put the ball which was left out. Observe that the ball removed is green as by induction hypothesis all balls were green. Again, the new collection of n balls has at least one green ball and hence, by induction hypothesis, all the balls in this new collection are also green. Therefore, we see that all the $n+1$ balls are green. Hence the result follows by induction hypothesis.

2. *In any collection of n lines in a plane, no two of which are parallel, all the lines pass through a common point.*
   Solution: *If $n = 1, 2$ then the result is easily seen to be true. So, let the result be true for any collection of n lines, no two of which are parallel. That is, we assume that if we are given any collection of n lines which are pairwise non-parallel then they pass through a common point.*

   Now, let us consider a collection of $n+1$ lines in the plane. We are also given that no two lines in this collection are parallel. Let us denote these lines by $\ell_1, \ell_2, \ldots, \ell_{n+1}$. From this

collection of lines, let us choose the subset $\ell_1, \ell_2, \ldots, \ell_n$, consisting of n lines. By induction hypothesis, all these lines pass through a common point, say P, the point of intersection of the lines ℓ_1 and ℓ_2. Now, consider the collection $\ell_1, \ell_2, \ldots, \ell_{n-1}, \ell_{n+1}$. This collection again consists of n non-parallel lines and hence by induction hypothesis, all these lines pass through a common point. This common point is P itself, as P is the point of intersection of the lines ℓ_1 and ℓ_2. Thus, by the principle of mathematical induction the proof of our statement is complete.

3. *Consider the polynomial $f(x) = x^2 - x + 41$. Check that for $1 \leq n \leq 40$, $f(n)$ is a prime number. Does this necessarily imply that $f(n)$ is prime for all positive integers n?* Check that $f(41) = 41^2$ and hence $f(41)$ is not a prime. Thus, the validity is being negated using the proof technique "disproving by counter-example".

Exercise 1.2.8. 1. *Prove that $n(n+1)$ is even for all $n \in \mathbb{N}$.*

2. *Prove that 3 divides $n^{16} - 2n^4 + n^2$, for all $n \in \mathcal{N}$.*

3. *Prove that 3 divides $n^4 - 4n^2$.*

4. *Prove that 5 divides $n^5 - n$, for all $n \in \mathcal{N}$.*

5. *Prove that 6 divides $n^3 - n$ for all $n \in \mathbb{N}$.*

6. *Prove that 7 divides $n^7 - n$, for all $n \in \mathcal{N}$.*

7. *Prove that 9 divides $2^{2n} - 3n - 1$.*

8. *Prove that 12 divides $2^{2n+2} - 3n^4 + 3n^2 - 4$.*

9. *Determine $1 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 + \cdots + (n-1) \cdot n$.*

10. *Prove that for all $n \geq 32$, there exist non-negative integers x and y such that $n = 5x + 9y$.*

11. *Prove that for all $n \geq 40$, there exist non-negative integers x and y such that $n = 5x + 11y$.*

12. *Let $x \in \mathbb{R}$ with $x \neq 1$. Then prove that $1 + x + x^2 + \cdots + x^n = \dfrac{x^{n+1} - 1}{x - 1}$.*

13. *Let $a, a+d, a+2d, \ldots, a+(n-1)d$ be the first n terms of an arithmetic progression. Then prove that $S = \sum_{i=0}^{n-1} (a + id) = \dfrac{n}{2}(a + nd)$.*

14. *Let $a, ar, ar^2, \ldots, ar^{n-1}$ be the first n terms of a geometric progression, with $r \neq 1$. Then prove that $S = \sum_{i=0}^{n-1} ar^i = a\dfrac{r^n - 1}{r - 1}$.*

15. *Prove that $1^3 + 2^3 + \cdots + n^3 = \left(\dfrac{n(n+1)}{2}\right)^2$.*

16. *Determine* $1 \cdot 2 \cdot 3 + 2 \cdot 3 \cdot 4 + 3 \cdot 4 \cdot 5 + \cdots + (n-1) \cdot n \cdot (n+1)$.

17. *Determine* $1 \cdot 3 \cdot 5 + 2 \cdot 4 \cdot 6 + \cdots + n \cdot (n+2) \cdot (n+4)$.

In the next few pages, we will try to study properties of integers that will be required later. We start with a lemma, commonly known as the "division algorithm". The proof again uses the technique "proof by contradiction".

Lemma 1.2.9 (Division Algorithm)**.** *Let a and b be two integers with $b > 0$. Then there exist unique integers q, r such that $a = qb + r$, where $0 \leq r < b$. The integer q is called the quotient and r, the remainder.*

Proof. Consider the set $S = \{a + bx : x \in \mathbb{Z}\} \cap \mathbb{N}$. Clearly, $a \in S$ and hence S is a non-empty subset of $\mathbb{N}$. Therefore, by Well-Ordering Principle, S contains its least element, say s_0. That is, there exists $x_0 \in \mathbb{Z}$, such that $s_0 = a + bx_0$. We claim that $0 \leq s_0 < b$.

As $s_0 \in S \subset \mathbb{N}$, one has $s_0 \geq 0$. So, let if possible assume that $s_0 \geq b$. This implies that $s_0 - b \geq 0$ and hence $s_0 - b = a + b(x_0 - 1) \in S$, a contradiction to the assumption that s_0 was the least element of S. Hence, we have shown the existence of integers q, r such that $a = qb + r$ with $0 \leq r < b$.

Uniqueness: Let if possible q_1, q_2, r_1 and r_2 be integers with $a = q_1 b + r_1 = q_2 b + r_2$, with $0 \leq r_1 \leq r_2 < b$. Therefore, $r_2 - r_2 \geq 0$ and thus, $0 \leq (q_1 - q_2)b = r_2 - r_1 < b$. Hence, we have obtained a multiple of b that is strictly less than b. But this can happen only if the multiple is 0. That is, $0 = (q_1 - q_2)b = r_2 - r_1$. Thus, one obtains $r_1 = r_2$ and $q_1 = q_2$ and the proof of uniqueness is complete.

This completes the proof of the lemma. ∎

Definition 1.2.10 (Greatest Common Divisor)**.**

1. *An integer a is said to* divide *an integer b, denoted $a|b$, if $b = ac$, for some integer c. Note that c can be a negative integer.*

2. Greatest Common Divisor: *Let $a, b \in \mathbb{Z} \setminus \{0\}$. Then the greatest common divisor of a and b, denoted $\gcd(a, b)$, is the largest positive integer c such that*

   (a) *c divides a and b, and*

   (b) *if d is any positive integer dividing a and b, then d divides c as well.*

3. Relatively Prime/Coprime Integers: *Two integers a and b are said to be relatively prime if $\gcd(a, b) = 1$.*

Theorem 1.2.11 (Euclid's Algorithm)**.** *Let a and b be two non-zero integers. Then there exists an integer d such that*

1. *$d = \gcd(a, b)$, and*

2. *there exist integers x_0, y_0 such that $d = ax_0 + by_0$.*

Proof. Consider the set $S = \{ax + by : x, y \in \mathbb{Z}\} \cap \mathbb{N}$. Then, either $a \in S$ or $-a \in S$, as exactly one of them is an element of $\mathbb{N}$ and both $a = a \cdot 1 + b \cdot 0$ and $-a = a \cdot (-1) + b \cdot 0$ are elements of the set $\{ax + by : x, y \in \mathbb{Z}\}$. Thus, S is non-empty subset of $\mathbb{N}$. So, by Well-Ordering Principle, S contains its least element, say d. As $d \in S$, there exist integers x_0, y_0 such that $d = ax_0 + by_0$.

We claim that d obtained as the least element of S also equals $\gcd(a, b)$. That is, we need to show that d satisfies both the conditions of Definition 1.2.10.2.

We first show that $d|a$. By division algorithm, there exist integers q and r such that $a = dq+r$, with $0 \leq r < d$. Thus, we need to show that $r = 0$. On the contrary, assume that $r \neq 0$. That is, $0 < r < d$. Then by definition, $r \in \mathbb{N}$ and $r = a-dq = a-q\cdot(ax_0+by_0) = a\cdot(1-qx_0)+b\cdot(-qy_0) \in \{ax + by : x, y \in \mathbb{Z}\}$. Hence, $r \in S$ and by our assumption $r < d$. This contradicts the fact that d was the least element of S. Thus, our assumption that $r \neq 0$ is false and hence $a = dq$. This implies that $d|a$. In a similar way, it can be shown that $d|b$.

Now, assume that there is an integer c such that c divides both a and b. We need to show that $c|d$. Observe that as c divides both a and b, c also divides both ax_0 and by_0 and hence c also divides $ax_0 + by_0 = d$. Thus, we have shown that d satisfies both the conditions of Definition 1.2.10.2 and therefore, the proof of the theorem is complete. ■

The above theorem is often stated as "the $\gcd(a, b)$ is a linear combination of the numbers a and b". To proceed further, we need the following definitions.

Example 1.2.12. *1. Consider two integers, say* 155 *and* -275*. Then, by division algorithm, one obtains*

$$\begin{aligned} -275 &= (-2) \cdot 155 + 35 & \qquad 155 &= 4 \cdot 35 + 15 \\ 35 &= 2 \cdot 15 + 5 & \qquad 15 &= 3 \cdot 5. \end{aligned}$$

Hence, $5 = \gcd(155, -275)$ *and* $5 = 9 \cdot (-275) + 16 \cdot 155$*, as*

$$5 = 35-2\cdot 15 = 35-2(155-4\cdot 35) = 9\cdot 35-2\cdot 155 = 9(-275+2\cdot 155)-2\cdot 155 = 9\cdot(-275)+16\cdot 155.$$

Also, note that $275 = 5\cdot 55$ *and* $155 = 5\cdot 31$ *and thus,* $5 = (9+31x)\cdot(-275)+(16+55x)\cdot 155$*, for all* $x \in \mathbb{Z}$*. Therefore, we see that there are infinite number of choices for the pair* $(x, y) \in \mathbb{Z}^2$*, for which* $d = ax + by$*.*

2. In general, given two non-zero integers a *and* b*, we can use the division algorithm to get* $\gcd(a, b)$*. This algorithm is also attributed to Euclid. Without loss of generality, assume that both* a *and* b *are positive and* $a > b$*. Then the algorithm proceeds as follows:*

$$\begin{aligned} a &= bq_0 + r_0 \ \textit{ with } 0 \leq r_0 < b, & \qquad b &= r_0q_1 + r_1 \ \textit{ with } 0 \leq r_1 < r_0, \\ r_0 &= r_1q_2 + r_2 \ \textit{ with } 0 \leq r_2 < r_1, & \qquad r_1 &= r_2q_3 + r_3 \ \textit{ with } 0 \leq r_3 < r_2, \\ &\vdots = \vdots \\ r_{\ell-1} &= r_\ell q_{\ell+1} + r_{\ell+1} \ \textit{ with } 0 \leq r_{\ell+1} < r_\ell, & \qquad r_\ell &= r_{\ell+1}q_{\ell+2}. \end{aligned}$$

The process will take at most $b-1$ steps as $0 \leq r_0 < b$. Also, note that $\gcd(a,b) = r_{\ell+1}$ and it can be recursively obtained, using backtracking. That is,

$$r_{\ell+1} = r_{\ell-1} - r_\ell q_{\ell+1} = r_{\ell-1} - q_{\ell+1}\left(r_{\ell-2} - r_{\ell-1}q_\ell\right) = r_{\ell-1}\left(1 + q_{\ell+1}q_\ell\right) - q_{\ell+1}r_{\ell-2} = \cdots .$$

Exercise 1.2.13. 1. *Prove that $\gcd(a,b) = \gcd(a,b+a) = \gcd(a+b,b)$, for any two non-zero integers a and b.*

2. *Does there exist $n \in \mathbb{Z}$ such that $\gcd(2n+3, 5n+7) \neq 1$?*

3. *Prove that the system $15x + 12y = b$ has a solution for $x, y \in \mathbb{Z}$ if and only if 3 divides b.*

4. *Prove that $\gcd(a,bc) = 1$ if and only if $\gcd(a,b) = 1$ and $\gcd(a,c) = 1$, for any three non-zero integers a, b and c.*

Definition 1.2.14 (Prime/Composite Numbers)**.** 1. *The positive integer 1 is called the unity or the unit element of $\mathbb{Z}$.*

2. *A positive integer p is said to be a prime, if p has exactly two factors, namely, 1 and p itself.*

3. *An integer that is neither prime and nor is equal to 1, is called composite.*

We are now ready to prove an important result that helps us in proving the fundamental theorem of arithmetic.

Lemma 1.2.15 (Euclid's Lemma)**.** *Let p be a prime and let $a, b \in \mathbb{Z}$. If $p|ab$ then either $p|a$ or $p|b$.*

Proof. If $p|a$, then we are done. So, let us assume that p does not divide a. But p is a prime and hence $\gcd(p,a) = 1$. Thus, by Euclid's algorithm, there exist integers x, y such that $1 = ax + py$. Therefore,

$$b = b \cdot 1 = b \cdot (ax + py) = ab \cdot x + p \cdot by.$$

Now, the condition $p|ab$ implies that p divides $ab \cdot x + p \cdot by = b$. Thus, we have shown that if $p|ab$ then either $p|a$ or $p|b$. ∎

Now, we are ready to prove the fundamental theorem of arithmetics that states that "every positive integer greater than 1 is either a prime or is a product of primes. This product is unique, except for the order in which the prime factors appear".

Theorem 1.2.16 (Fundamental Theorem of Arithmetic)**.** *Let $n \in \mathbb{N}$ with $n \geq 2$. Then there exist prime numbers $p_1 > p_2 > \cdots > p_k$ and positive integers $s_1, s_2, \ldots, s_k$ such that $n = p_1^{s_1} p_2^{s_2} \cdots p_k^{s_k}$, for some $k \geq 1$. Moreover, if n also equals $q_1^{t_1} q_2^{t_2} \cdots q_\ell^{t_\ell}$, for distinct primes $q_1, q_2, \ldots, q_\ell$ and positive integers $t_1, t_2, \ldots, t_\ell$ then $k = \ell$ and for each i, $1 \leq i \leq k$, there exists j, $1 \leq j \leq k$ such that $p_i = q_j$ and $s_i = t_j$.*

Proof. We prove the result using the strong form of the principle of mathematical induction. If n equals a prime, say p then clearly $n = p^1$ and hence the first step of the induction holds true. Hence, let us assume that the result holds for all positive integers that are less than n. We need to prove the result for the positive integer n.

If n itself is a prime then we are done. Else, there exists positive integers a and b such that $n = ab$ and $1 \leq a, b < n$. Thus, by the strong form of the induction hypothesis, there exist primes p_i's, q_j's and positive integers s_i and t_j's such that $a = p_1^{s_1} p_2^{s_2} \cdots p_k^{s_k}$, for some $k \geq 1$ and $b = q_1^{t_1} q_2^{t_2} \cdots q_\ell^{t_\ell}$, for some ℓ. Hence,

$$n = ab = p_1^{s_1} p_2^{s_2} \cdots p_k^{s_k} q_1^{t_1} q_2^{t_2} \cdots q_\ell^{t_\ell}.$$

Now, if some of the p_i's and q_j's are equal, they can be multiplied together to obtain n as a product of distinct prime powers.

Thus, using the string form of the principle of mathematical induction, the result is true for all positive integer n. As per as the uniqueness is concerned, it follows by a repeated application of Lemma 1.2.15.

To see this, observe that p_1 divides $n = q_1^{t_1} q_2^{t_2} \cdots q_\ell^{t_\ell}$ implies that p_1 divides exactly one of them (the primes are distinct), say q_1. Also, it is clear that in this case $s_1 = t_1$. For otherwise, either p_1 will divide $q_2^{t_2} \cdots q_\ell^{t_\ell}$, or $q_1 = p_1$ will divide $p_2^{s_2} \cdots p_k^{s_k}$. This process can be continued a finite number of times to get the required result. ∎

As an application of the fundamental theorem of arithmetic, one has the following well known result. This is the first instance where we have used the contrapositive argument technique to prove the result.

Corollary 1.2.17. *Let $n \in \mathbb{N}$ with $n \geq 2$. Suppose that for any prime $p \leq \sqrt{n}$, p does not divide n then n is prime.*

Proof. Suppose n is not a prime. Then there exists positive integers a and b such that $n = ab$ and $2 \leq a, b \leq n$. Also, note that at least one of them, say $a \leq \sqrt{n}$. For if, both $a, b > \sqrt{n}$ then $n = ab > n$, giving us a contradiction.

Since $a \leq \sqrt{n}$, by Theorem 1.2.16, one of its prime factors, say p will satisfy $p \leq a \leq \sqrt{n}$. Thus, if n has no prime divisor less than or equal to $\sqrt{n}$ then n must be itself be a prime. ∎

Exercise 1.2.18. *1. For every positive integer $n \geq 5$ prove that $2^n > n^2 > n$.*

2. Let $n \in \mathbb{N}$ with $n \geq 2$. Then prove that there exists $s \in \mathbb{N}$, such that $n = 2^s t$, for some odd integer t. : $0 \in \mathbb{N}$.

1.3 Relations and Partitions

We start with the definition of cartesian product of two sets and to define relations.

Definition 1.3.1 (Cartesian Product)**.** *Let A and B be two sets. Then their cartesian product, denoted $A \times B$, is defined as $A \times B = \{(a, b) : a \in A, b \in B\}$.*

Example 1.3.2. *1. Let $A = \{a, b, c\}$ and $B = \{1, 2, 3, 4\}$. Then*

$$\begin{aligned} A \times A &= \{(a,a),(a,b),(a,c),(b,a),(b,b),(b,c),(c,a),(c,b),(c,c)\}. \\ A \times B &= \{(a,1),(a,2),(a,3),(a,4),(b,1),(b,2),(b,3),(b,4),(c,1),(c,2),(c,3),(c,4)\}. \end{aligned}$$

2. The Euclidean plane, denoted $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) : x \in \mathbb{R}\}$.

Definition 1.3.3 (Relation)**.** *A relation R on a non-empty set A, is a subset of $A \times A$.*

Example 1.3.4. *1. Let $A = \{a, b, c, d\}$. Then, some of the relations R on A are:*

(a) $R = A \times A$.

(b) $R = \{(a,a),(b,b),(c,c),(d,d),(a,b),(a,c),(b,c)\}$.

(c) $R = \{(a,a),(b,b),(c,c)\}$.

(d) $R = \{(a,a),(a,b),(b,a),(b,b),(d,d)\}$.

(e) $R = \{(a,a),(a,b),(b,a),(a,c),(c,a),(c,c),(b,b)\}$.

(f) $R = \{(a,b),(b,c),(a,c),(d,d)\}$.

2. Consider the set $\mathbb{Z}$. Some of the relations on $\mathbb{Z}$ are as follows:

(a) $R = \{(a, b) \in \mathbb{Z}^2 : a|b\}$.

(b) Fix a positive integer n and define $R = \{(a, b) \in \mathbb{Z}^2 : n$ divides $a - b\}$.

(c) $R = \{(a, b) \in \mathbb{Z}^2 : a \leq b\}$.

(d) $R = \{(a, b) \in \mathbb{Z}^2 : a > b\}$.

3. Consider the set $\mathbb{R}^2$. Also, let us write $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$. Then some of the relations on $\mathbb{R}^2$ are as follows:

(a) $R = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^2 \times \mathbb{R}^2 : |\mathbf{x}|^2 = x_1^2 + x_2^2 = y_1^2 + y_2^2 = |\mathbf{y}|^2\}$.

(b) $R = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^2 \times \mathbb{R}^2 : \mathbf{x} = \alpha\mathbf{y}$ for some $\alpha \in \mathbb{R}^*\}$.

(c) $R = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^2 \times \mathbb{R}^2 : 4x_1^2 + 9x_2^2 = 4y_1^2 + 9y_2^2\}$.

(d) $R = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^2 \times \mathbb{R}^2 : \mathbf{x} - \mathbf{y} = \alpha(1, 1)$ for some $\alpha \in \mathbb{R}^*\}$.

(e) Fix a $c \in \mathbb{R}$. Now, define $R = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^2 \times \mathbb{R}^2 : y_2 - x_2 = c(y_1 - x_1)\}$.

(f) $R = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^2 \times \mathbb{R}^2 : |\mathbf{x}| = \alpha|\mathbf{y}|\}$, for some positive real number α.

4. Let A be the set of triangles in a plane. Then $R = \{(a, b) \in A^2 : a \sim b\}$, where $\sim$ stands for similarity of triangles.

5. *In $\mathbb{R}$, define a relation $R = \{(a, b) \in \mathbb{R}^2 : |a - b|$ is an integer$\}$.*

6. *Let A be any non-empty set and consider the set $\mathcal{P}(A)$. Then one can define a relation R on $\mathcal{P}(A)$ by $R = \{(S, T) \in \mathcal{P}(A) \times \mathcal{P}(A) : S \subset T\}$.*

Now that we have seen quite a few examples of relations, let us look at some of the properties that are of interest in mathematics.

Definition 1.3.5. *Let R be a relation on a non-empty set A. Then R is said to be*

1. *reflexive if $(a, a) \in R$, for all $a \in A$.*

2. *symmetric if $(b, a) \in R$ whenever $(a, b) \in R$.*

3. *anti-symmetric if, for all $a, b \in A$, the conditions $(a, b), (b, a) \in R$ implies that $a = b$ in A.*

4. *transitive if, for all $a, b, c \in A$, the conditions $(a, b), (b, c) \in R$ implies that $(a, c) \in R$.*

Exercise 1.3.6. *For each of the relations defined in Example 1.3.4, determine which of them are*

1. *reflexive.*

2. *symmetric.*

3. *anti-symmetric.*

4. *transitive.*

We are now ready to define a relation that appears quite frequently in mathematics. Before doing so, let us either use the symbol $\sim$ or $\overset{R}{\sim}$ for relation. That is, if $a, b \in A$ then $a \sim b$ or $a \overset{R}{\sim} b$ will stand for $(a, b) \in R$.

Definition 1.3.7. *Let $\sim$ be a relation on a non-empty set A. Then $\sim$ is said to form an equivalence relation if $\sim$ is reflexive, symmetric and transitive.*

The equivalence class containing $a \in A$, denoted $[a]$, is defined as $[a] := \{b \in A : b \sim a\}$.

Example 1.3.8. 1. *Let $a, b \in \mathbb{Z}$. Then $A \sim b$, if 10 divides $a - b$. Then verify that $\sim$ is an equivalence relation. Moreover, the equivalence classes can be taken as $[0]$, $[1]$, $\ldots$, $[9]$. Observe that, for $0 \leq i \leq 9$, $[i] = \{10n + i : n \in \mathbb{Z}\}$.* This equivalence relation in modular arithmetic is written as $a \equiv b \pmod{10}$.

In general, for any fixed positive integer n, the statement "$a \equiv b \pmod{n}$" (read "a is equivalent to b modulo n" *) is equivalent to saying that $a \sim b$ if n divides $a - b$.*

2. *Determine the equivalence relations that appear in Example 1.3.4. Also, for each equivalence relation, determine a set of equivalence classes.*

Definition 1.3.9 (Partition of a set)**.** *Let A be a non-empty set. Then a partition Π of A, into m-parts, is a collection of non-empty subsets $A_1, A_2, \ldots, A_m$, of A, such that*

1. *$A_i \cap A_j = \emptyset$ (empty set), for $1 \leq i \neq j \leq m$ and*
2. *$\bigcup_{i=1}^{m} A_i = A$.*

Example 1.3.10. *1. The partitions of $A = \{a, b, c, d\}$ into*

   (a) 3-parts are $a|\ b|\ cd$, $a|\ bc|\ d$, $ac|\ b|\ d$, $a|\ bd|\ c$, $ad|\ b|\ c$, $ab|\ c|\ d$, where the expression $a|bc|d$ represents the partition $A_1 = \{a\}$, $A_2 = \{b, c\}$ and $A_3 = \{d\}$.

   (b) 2-parts are

$$a|\ bcd, \quad b|\ acd, \quad c|\ abd, \quad d|\ abc, \quad ab|\ cd, \quad ac|\ bd \quad \text{and} \quad ad|\ bc.$$

2. Let $A = \mathbb{Z}$ and define

   (a) $A_0 = \{2x : x \in \mathbb{Z}\}$ and $A_1 = \{2x + 1 : x \in \mathbb{Z}\}$. Then $\Pi = \{A_0, A_1\}$ forms a partition of Zl into odd and even integers.

   (b) $A_i = \{10n + i : n \in \mathbb{Z}\}$, for $i = 1, 2, \ldots, 10$. Then $\Pi = \{A_1, A_2, \ldots, A_{10}\}$ forms a partition of $\mathbb{Z}$.

3. $A_1 = \{0, 1\}$, $A_2 = \{n \in \mathbb{N} : n$ is a prime$\}$ and $A_3 = \{n \in \mathbb{N} : n \geq 3, n$ is composite$\}$. Then $\Pi = \{A_1, A_2, A_3\}$ is a partition of Nl.

4. Let $A = \{a, b, c, d\}$. Then $\Pi = \{\{a\}, \{b, d\}, \{c\}\}$ is a partition of A.

Observe that the equivalence classes produced in Example 1.3.8.1 indeed correspond to the non-empty sets A_i's, defined in Example 1.3.10.2b. In general, such a statement is always true. That is, suppose that A is a non-empty set with an equivalence relation $\sim$. Then the equivalence classes of $\sim$ in A, gives rise to a partition of A. Conversely, given any partition Π of A, there is an equivalence relation on A whose equivalence classes are the elements of Π. This is proved as the next result.

Theorem 1.3.11. *Let A be a non-empty set.*

1. *Also, let $\sim$ define an equivalence relation on the set A. Then the set of equivalence classes of $\sim$ in A gives a partition of A.*
2. *Let I be a non-empty index set such that $\{A_i : i \in I\}$ gives a partition of A. Then there exists an equivalence relation on A whose equivalence classes are exactly the sets $A_i, i \in I$.*

Proof. Since $\sim$ is reflexive, $a \sim a$, for all $a \in A$. Hence, the equivalence class $[a]$ contains a, for each $a \in A$. Thus, the equivalence classes are non-empty and clearly, their union is the whole set A. We need to show that if $[a]$ and $[b]$ are two equivalence classes of $\sim$ then either $[a] = [b]$ or $[a] \cap [b] = \emptyset$.

Let $x \in [a] \cap [b]$. Then by definition, $x \sim a$ and $x \sim b$. Since $\sim$ is symmetric, one also has $a \sim x$. Therefore, we see that $a \sim x$ and $x \sim b$ and hence, using the transitivity of $\sim$, $a \sim b$. Thus, by definition, $a \in [b]$ and hence $[a] \subseteq [b]$. But $a \sim b$, also implies that $b \sim a$ ($\sim$ is transitive) and hence $[b] \subseteq [a]$. Thus, we see that if $[a] \cap [b] \neq \emptyset$, then $[a] = [b]$. This proves the first part of the theorem.

For the second part, define a relation $\sim$ on A as follows: for any two elements $a, b \in A$, $a \sim b$ if there exists an $i, i \in I$ such that $a, b \in A_i$. It can be easily verified that $\sim$ is indeed reflexive, symmetric and transitive. Also, verify that the equivalence classes of $\sim$ are indeed the sets $A_i, i \in I$. ∎

Exercise 1.3.12. *1. For each of the equivalence relations given in Example 1.3.4, explicitly determine the equivalence classes.*

2. Fix a positive integer n. Suppose, for any two integers a, b, $a \equiv b \pmod{n}$. Then prove that

   (a) for any integer c, $a + c \equiv b + c \pmod{n}$.

   (b) for any integer c, $ac \equiv bc \pmod{n}$.

   (c) $a \equiv b \pmod{m}$, for any positive integer m that divides n.

   (d) $a^s \equiv b^s \pmod{n}$, for any positive integer s.

3. Fix a positive integer n and let a, b, c, d be integers with $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$. Then prove that

   (a) $a + c \equiv b + d \pmod{n}$.

   (b) $ac \equiv bd \pmod{n}$.

4. If m and n are two positive integers and $a \equiv b \pmod{mn}$ then prove that $a \equiv b \pmod{m}$ and $a \equiv b \pmod{n}$. Under what condition on m and n the converse holds true?

5. Let a and n be two positive integers such that $\gcd(a, n) = 1$. Then prove that the system $ax \equiv b \pmod{n}$ has a solution, for every b. Moreover, if x_1 and x_2 are any two solutions then $x_1 \equiv x_2 \pmod{n}$.

6. Let m and n be two positive integers such that $\gcd(m, n) = 1$. Then prove that the system $x \equiv a \pmod{m}$ and $x \equiv b \pmod{n}$ are simultaneously solvable for every choice of a and b.

7. *Let $\mathbb{Z}_m = \{0, 1, 2, \ldots, m-1\}$. In $\mathbb{Z}_m$, we define the binary operation $\oplus_m$ by*

$$a \oplus_m b = \begin{cases} a+b, & \text{if } a+b < m, \\ a+b-m, & \text{if } a+b \geq m. \end{cases}$$

This binary operation is called **addition modulo** m. *In $\mathbb{Z}_m$, we also define the binary operation $\otimes_m$, commonly known as* **multiplication modulo** m, *by*

$$a \otimes_m b = \text{ the remainder when } a \cdot b \text{ is divided by } m.$$

(a) *Prove that*

$$i \pmod n + j \pmod n = i + j \pmod n \text{ and } i \pmod n \cdot j \pmod n = ij \pmod n.$$

(b) *In $\mathbb{Z}_5$, solve the equation $2x \equiv 1 \pmod 5$. What happens when we consider the equation $2x \equiv 1 \pmod 7$ in $\mathbb{Z}_7$? What if we consider $2x \equiv 1 \pmod p$ in $\mathbb{Z}_p$, p a prime?*

1.4 Functions

Definition 1.4.1 (Function)**.** *1. Let A and B be two sets. Then a function $f : A \longrightarrow B$ is a rule that assigns to each element of A exactly one element of B.*

2. *The set A is called the* domain *of the function f.*

3. *The set B is called the* co-domain *of the function f.*

The readers should carefully read the following important remark before proceeding further.

Remark 1.4.2. *1. If $A = \emptyset$, then by convention, one assumes that there is a function, called the empty function, from A to B.*

2. *If $B = \emptyset$, then it can be easily observed that there is no function from A to B.*

3. *Some books use the word "map" in place of "function". So, both the words may be used interchangeably through out the notes.*

4. *Through out these notes, whenever the phrase "let $f : A \longrightarrow B$ be a function" is used, it will be assumed that both A and B are non-empty sets.*

Example 1.4.3. *1. Let $A = \{a, b, c\}$, $B = \{1, 2, 3\}$ and $C = \{3, 4\}$. Then verify that the examples given below are indeed functions.*

(a) *$f : A \longrightarrow B$, defined by $f(a) = 3, f(b) = 3$ and $f(c) = 3$.*

(b) *$f : A \longrightarrow B$, defined by $f(a) = 3, f(b) = 2$ and $f(c) = 2$.*

(c) $f : A \longrightarrow B$, *defined by* $f(a) = 3, f(b) = 1$ *and* $f(c) = 2$.

(d) $f : A \longrightarrow C$, *defined by* $f(a) = 3, f(b) = 3$ *and* $f(c) = 3$.

(e) $f : C \longrightarrow A$, *defined by* $f(3) = a, f(4) = c$.

2. *Verify that the following examples give functions,* $f : \mathbb{Z} \longrightarrow \mathbb{Z}$.

(a) $f(x) = 1$, *if* x *is even and* $f(x) = 5$, *if* x *is odd.*

(b) $f(x) = -1$, *for all* $x \in \mathbb{Z}$.

(c) $f(x) = x \pmod{10}$, *for all* $x \in \mathbb{Z}$.

(d) $f(x) = 1$, *if* $x > 0$, $f(0) = 0$ *and* $f(x) = 1$, *if* $x < 0$.

Exercise 1.4.4. *Do the following give examples of functions? Give reasons for your answer.*

1. *Let* $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ *such that*

(a) $f(x) = 1$, *if* x *is a multiple of* 2 *and* $f(x) = 5$, *if* x *is a multiple of* 3.

(b) $f(x) = 1$, *if* $x = a^2$, *for some* $a \in \mathbb{Z}$ *and* -1, *otherwise.*

(c) $f(x) = x^3$, *for all* $x \in \mathbb{Z}$.

(d) *for a fixed positive integer* n, $f(x) = x^{2n}$, *for all* $x \in \mathbb{Z}$.

(e) *for a fixed positive integer* n, $f(x) = x^{2n+1}$, *for all* $x \in \mathbb{Z}$.

2. *Let* $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$ *such that* $f(x) = \pm\sqrt{x}$, *for all* $x \in \mathbb{R}^+$.

3. *Let* $f : \mathbb{R} \longrightarrow \mathbb{R}$ *such that* $f(x) = \sqrt{x}$, *for all* $x \in \mathbb{R}$.

4. *Let* $f : \mathbb{R} \longrightarrow \mathbb{C}$ *such that* $f(x) = \sqrt{x}$, *for all* $x \in \mathbb{R}$.

5. *Let* $f : \mathbb{R}^* \longrightarrow \mathbb{R}$ *such that* $f(x) = \log_e |x|$, *for all* $x \in \mathbb{R}^*$.

6. *Let* $f : \mathbb{R} \longrightarrow \mathbb{R}$ *such that* $f(x) = \tan x$, *for all* $x \in \mathbb{R}$.

Definition 1.4.5. *Let* $f : A \longrightarrow B$ *be a function. Then,*

1. *for each* $x \in A$, *the element* $f(x) \in B$ *is called the* image *of* x *under* f.

2. *the* range/image *of* A *under* f *equals* $f(A) = \{f(a) : a \in A\}$.

3. *the function* f *is said to be* one-to-one *if "for any two distinct elements* $a_1, a_2 \in A$, $f(a_1) \neq f(a_2)$*".*

4. *the function* f *is said to be* onto *if "for every element* $b \in B$ *there exists an element* $a \in A$, *such that* $f(a) = b$*".*

5. *for any function* $g : B \longrightarrow C$, *the composition* $g \circ f : A \longrightarrow C$ *is a function defined by* $(g \circ f)(a) = g\big(f(a)\big)$, *for every* $a \in A$.

Example 1.4.6. *1. Let* $f : \mathbb{N} \longrightarrow \mathbb{Z}$ *be defined by* $f(x) = \begin{cases} \frac{-x}{2}, & \text{if } x \text{ is even,} \\ \frac{x+1}{2}, & \text{if } x \text{ is odd.} \end{cases}$ *Then prove that* f *is one-one. Is* f *onto?*

Solution: *Let us use the contrapostive argument to prove that* f *is one-one. Let if possible* $f(x) = f(y)$*, for some* $x, y \in \mathbb{N}$*. Using the definition, one sees that* x *and* y *are either both odd or both even. So, let us assume that both* x *and* y *are even. In this case,* $\frac{-x}{2} = \frac{-y}{2}$ *and hence* $x = y$*. A similar argument holds, in case both* x *and* y *are odd.*

Claim: f *is onto.*

Let $x \in \mathbb{Z}$ *with* $x \geq 1$*. Then* $2x - 1 \in \mathbb{N}$ *and* $f(2x-1) = \frac{(2x-1)+1}{2} = x$*. If* $x \in \mathbb{Z}$ *and* $x \leq 0$*, then* $-2x \in \mathbb{N}$ *and* $f(-2x) = \frac{-(-2x)}{2} = x$*. Hence,* f *is indeed onto.*

2. *Let* $f : \mathbb{N} \longrightarrow \mathbb{Z}$ *and* $g : \mathbb{Z} \longrightarrow \mathbb{Z}$ *be defined by,* $f(x) = 2x$ *and* $g(x) = \begin{cases} 0, & \text{if } x \text{ is odd,} \\ x/2, & \text{if } x \text{ is even,} \end{cases}$ *respectively. Then prove that the functions* f *and* $g \circ f$ *are one-one but* g *is not one-one.*

Solution: *By definiton, it is clear that* f *is indeed one-one and* g *is not one-one. But*

$$g \circ f(x) = g(f(x)) = g(2x) = \frac{2x}{2} = x,$$

for all $x \in \mathbb{N}$*. Hence,* $g \circ f : \mathbb{N} \longrightarrow \mathbb{Z}$ *is also one-one.*

Exercise 1.4.7. *For each of the functions given in Example 1.4.3, determine the*

1. *functions that are one-one, onto and/or both.*
2. *range.*

The next theorem gives some result related with composition of functions.

Theorem 1.4.8 (Properties of Functions)**.** *Consider the functions* $f : A \longrightarrow B$*,* $g : B \longrightarrow C$ *and* $h : C \longrightarrow D$*.*

1. *Then* $(h \circ g) \circ f = h \circ (g \circ f)$ *(*associativity holds).
2. If f and g are one-to-one then the function $g \circ f$ is also one-to-one.
3. If f and g are onto then the function $g \circ f$ is also onto.

Proof. First note that $g \circ f : A \longrightarrow C$ and both $(h \circ g) \circ f$, $h \circ (g \circ f)$ are functions from A to D.

Proof of Part 1: The first part is direct, as for each $a \in A$,

$$((h \circ g) \circ f)(a) = (h \circ g)(f(a)) = h(g(f(a))) = h((g \circ f)(a)) = (h \circ (g \circ f))(a).$$

Proof of Part 2: Need to show that "whenever $(g \circ f)(a_1) = (g \circ f)(a_2)$, for some $a_1, a_2 \in A$ then $a_1 = a_2$".

So, let us assume that $g(f(a_1)) = (g \circ f)(a_1) = (g \circ f)(a_2) = g(f(a_2))$, for some $a_1, a_2 \in A$. As g is one-one, the assumption gives $f(a_1) = f(a_2)$. But f is also one-one and hence $a_1 = a_2$.

Proof of Part 3: To show that "given any $c \in C$, there exists $a \in A$ such that $(g \circ f)(a) = c$".

As g is onto, for the given $c \in C$, there exists $b \in B$ such that $g(b) = c$. But f is also given to be onto. Hence, for the b obtained in previous step, there exists $a \in A$ such that $f(a) = b$. Hence, we see that $c = g(b) = g(f(a)) = (g \circ f)(a)$. ■

Definition 1.4.9 (Identity Function)**.** *Fix a set A and let $e_A : A \longrightarrow A$ be defined by $e_A(a) = a$, for all $a \in A$. Then the function e_A is called the identity function or map on A.*

The subscript A in Definition 1.4.9 will be removed, whenever there is no chance of confusion about the domain of the function.

Theorem 1.4.10 (Properties of Identity Function)**.** *Fix two non-empty sets A and B and let $f : A \longrightarrow B$ and $g : B \longrightarrow A$ be any two functions. Also, let $e : A \longrightarrow A$ be the identity map defined above. Then*

1. *e is a one-one and onto map.*
2. *the map $f \circ e = f$.*
3. *the map $e \circ g = g$.*

Proof. Proof of Part 1: Since $e(a) = a$, for all $a \in A$, it is clear that e is one-one and onto.

Proof of Part 2: BY definition, $(f \circ e)(a) = f(e(a)) = f(a)$, for all $a \in A$. Hence, $f \circ e = f$.

Proof of Part 3: The readers are advised to supply the proof. ■

Example 1.4.11. 1. *Let $f, g : \mathbb{N} \longrightarrow \mathbb{N}$ be defined by, $f(x) = 2x$ and $g(x) = \begin{cases} 0, & \text{if } x \text{ is odd,} \\ x/2, & \text{if } x \text{ is even.} \end{cases}$ Then verify that $g \circ f : \mathbb{N} \longrightarrow \mathbb{N}$ is the identity map, whereas $f \circ g$ maps even numbers to itself and maps odd numbers to 0.*

Definition 1.4.12 (Invertible Function)**.** *A function $f : A \longrightarrow B$ is said to be invertible if there exists a function $g : B \longrightarrow A$ such that the map*

1. *$g \circ f : A \longrightarrow A$ is the identity map on A, and*
2. *$f \circ g : B \longrightarrow B$ is the identity map on B.*

Let us now prove that if $f : A \longrightarrow B$ is an invertible map then the map $g : B \longrightarrow A$, defined above is unique.

Theorem 1.4.13. *Let $f : A \longrightarrow B$ be an invertible map. Then the map*

1. *g defined in Definition 1.4.12 is unique. The map g is generally denoted by f^{-1}.*

2. $\left(f^{-1}\right)^{-1} = f$.

Proof. The proof of the second part is left as an exercise for the readers. Let us now proceed with the proof of the first part.

Suppose $g, h : B \longrightarrow A$ are two maps satisfying the conditions in Definition 1.4.12. Therefore, $g \circ f = e_A = h \circ f$ and $f \circ g = e_B = f \circ h$. Hence, using associativity of functions, for each $b \in B$, one has

$$g(b) = g(e_B(b)) = g\left((f \circ h)(b)\right) = (g \circ f)\left(h(b)\right) = e_A(h(b)) = h(b).$$

Hence, the maps h and g are the same and thus the proof of the first part is over. ■

Theorem 1.4.14. *Let $f : A \longrightarrow B$ be a function. Then f is invertible if and only if f is one-one and onto.*

Proof. Let f be invertible. To show, f is one-one and onto.

Since, f is invertible, there exists the map $f^{-1} : B \longrightarrow A$ such that $f \circ f^{-1} = e_B$ and $f^{-1} \circ f = e_A$. So, now suppose that $f(a_1) = f(a_2)$, for some $a_1, a_2 \in A$. Then, using the map f^{-1}, we get

$$a_1 = e_A(a_1) = \left(f^{-1} \circ f\right)(a_1) = f^{-1}\left(f(a_1)\right) = f^{-1}\left(f(a_2)\right) = \left(f^{-1} \circ f\right)(a_2) = e_A(a_2) = a_2.$$

Thus, f is one-one. To prove onto, let $b \in B$. Then, by definition, $f^{-1}(b) \in A$ and $f\left(f^{-1}(b)\right) = \left(f \circ f^{-1}\right)(b) = e_B(b) = b$. Hence, f is onto as well.

Now, let us assume that f is one-one and onto. To show, f is invertible. Consider the map $f^{-1} : B \longrightarrow A$ defined by "$f^{-1}(b) = a$ whenever $f(a) = b$", for each $b \in B$. This map is well-defined as f is onto implies that for each $b \in B$, there exists $a \in A$, such that $f(a) = b$. Also, f is one-one implies that the element a obtained in the previous line is unique.

Now, it can be easily verified that $f \circ f^{-1} = e_B$ and $f^{-1} \circ f = e_A$ and hence f is indeed invertible. ■

We now state the following important theorem whose proof is beyond the scope of this book. The theorem is popularly known as the "Cantor Bernstein Schroeder theorem".

Definition 1.4.15 (Cantor Bernstein Schroeder Theorem)**.** *Let A and B be two sets. Then A and B are said to have the same cardinality if there exists a one-one, onto map $f : A \longrightarrow B$.*

We end this section with the following set of exercises.

Exercise 1.4.16. 1. *Let A and B be two finite sets. Prove that $|A \times B| = |A| \cdot |B|$.*

2. *Prove that the functions f, defined below, are one-one and onto. Also, determine f^{-1}.*

   (a) $f : (-\infty, \infty) \longrightarrow \left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$ *with* $f(x) = \tan x$.

(b) $f : [0, \infty) \longrightarrow [1, \infty)$ *with* $f(x) = x^2 + 1$.

(c) $f : (0, 1) \longrightarrow (1, \infty)$ *with* $f(x) = \dfrac{1}{x}$.

(d) $f : [0, \infty) \longrightarrow [2, \infty)$ *with* $f(x) = |x| + |x - 1| + |x + 1|$.

(e) $f : \mathbb{R} \longrightarrow (0, \infty)$ *with* $f(x) = e^x$, *is the natural exponential function.*

3. *Prove that the pair of sets* A *and* B, *given below, have the same cardinality.*

(a) $A = \mathbb{N}$ *and* $B = \mathbb{Z}$.

(b) $A = (0, 1)$ *and* $B = (1, \infty)$.

(c) $A = (0, 1)$ *and* $B = \mathbb{R}$.

Chapter 2

Counting and Permutations

2.1 Principles of Basic Counting

In the previous chapter, we had seen the following:
Let A and B be two non-empty finite disjoint subsets of a set S. Then

1. $|A \cup B| = |A| + |B|$.
2. $|A \times B| = |A| \cdot |B|$.
3. A and B have the same cardinality if there exists a one-one and onto function $f : A \longrightarrow B$.

2.1.1 Distinguishable Balls

Lemma 2.1.1. *Let M and N be two sets such that $|M| = m$ and $|N| = n$. Then the total number of functions $f : M \longrightarrow N$ equals n^m?*

Proof: Let $M = \{a_1, a_2, \ldots, a_m\}$ and $N = \{b_1, b_2, \ldots, b_n\}$. Since a function is determined as soon as we know the value of $f(a_i)$, for $1 \leq i \leq m$, a function $f : M \longrightarrow N$ has the form

$$f \leftrightarrow \begin{pmatrix} a_1 & a_2 & \cdots & a_m \\ f(a_1) & f(a_2) & \cdots & f(a_m) \end{pmatrix},$$

where $f(a_i) \in \{b_1, b_2, \ldots, b_n\}$, for $1 \leq i \leq m$. As there is no restriction on the function f, $f(a_1)$ has n choices, $b_1, b_2, \ldots, b_n$. Similarly, $f(a_2)$ has n choices, $b_1, b_2, \ldots, b_n$ and so on. Thus, the total number of functions $f : M \longrightarrow N$ is

$$\underbrace{n \cdot n \cdot \cdots \cdot n}_{m \text{ times}} = n^m.$$

■

Remark 2.1.2. *Observe that Lemma 2.1.1 is equivalent to the following question:* IN HOW MANY WAYS CAN m **distinguishable/distinct** BALLS BE PUT INTO n **distinguishable/distinct** BOXES? *Hint: Number the balls as $a_1, a_2, \ldots, a_m$ and the boxes as $b_1, b_2, \ldots, b_n$.*

Exercise 2.1.3. *1. In Hall-III, each student reads exactly one news paper, say "Indian Express" or or . If the number of students reading "Indian Express", "The Times of India" and "Dainik Jagran" are respectively,* 130*,* 155 *and* 235*, then determine the total number of students in Hall-III?*

2. How many distinct ways are there to make a 5 *letter word using the ENGLISH alphabet?*

3. How many distinct ways are there to make a 5 *letter word using the ENGLISH alphabet*

(a) with ONLY consonants?

(b) with ONLY vowels?

(c) with a consonant as the first letter and a vowel as the second letter?

(d) if the vowels appear only at odd positions?

4. Determine the total number of possible outcomes if

(a) two coins are tossed?

(b) a coin and a die are tossed?

(c) two dice are tossed?

(d) three dice are tossed?

(e) five coins are tossed?

5. How many distinct 5*-letter words using only A's, B's, C's, and D's are there that*

(a) contain the word "CAC"?

(b) contain the word "CCC"?

(c) contain the word "CDC"?

(d) does not contain the word "CCD"?

Lemma 2.1.4. *Let M and N be two sets such that $|M| = m$ and $|N| = n$. Then the total number of distinct one-to-one functions $f : M \longrightarrow N$ is $n(n-1)\cdots(n-m+1)$.*

Proof: Observe that "f is one-to-one" means "whenever $x \neq y$ we must have $f(x) \neq f(y)$". Therefore, if $m > n$, then the number of such functions is 0.

So, let us assume that $m \leq n$ with $M = \{a_1, a_2, \ldots, a_m\}$ and $N = \{b_1, b_2, \ldots, b_n\}$. Then by definition, $f(a_1)$ has n choices, $b_1, b_2, \ldots, b_n$. Once $f(a_1)$ is chosen, there are only $n-1$ choices for $f(a_2)$ ($f(a_2)$ has to be chosen from the set $\{b_1, b_2, \ldots, b_n\} \setminus \{f(a_1)\}$). Similarly, there are only $n-2$ choices for $f(a_3)$ ($f(a_3)$ has to be chosen from the set $\{b_1, b_2, \ldots, b_n\} \setminus \{f(a_1), f(a_2)\}$), and so on. Thus, the required number is $n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot (n-m+1)$. ■

Remark 2.1.5. *1. The product $n(n-1)\cdots 3\cdot 2\cdot 1$ is denoted by $n!$, and is commonly called "n factorial".*

2. *By convention, we assume that* $0! = 1$.

3. *Using the factorial notation* $n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot (n-m+1) = \dfrac{n!}{(n-m)!}$. *This expression is generally denoted by* $n_{(m)}$, *and is called the falling factorial of* n. *Thus, if* $m > n$ *then* $n_{(m)} = 0$ *and if* $n = m$ *then* $n_{(m)} = n!$.

4. *The following conventions will be used in these notes:*

$$0! = 0_{(0)} = 1,\ 0^0 = 1,\ n_{(0)} = 1 \text{ for all } n \geq 1,\ 0_{(m)} = 0 \text{ for } m \neq 0.$$

The proof of the next corollary is immediate from Lemma 2.1.4 and hence the proof is omitted.

Corollary 2.1.6. *Let* M *and* N *be two sets such that* $|M| = |N| = n$ *(say). Then the number of one-to-one functions* $f : M \longrightarrow N$ *equals* $n!$, *called "n-factorial".*

Exercise 2.1.7.

1. *How many distinct ways are there to make* 5 *letter words using the ENGLISH alphabet if the letters must be different?*

2. *How many distinct ways are there to arrange the* 5 *letters of the word* $ABCDE$*?*

3. *How many distinct ways are there to arrange the* 5 *couples on* 10 *chairs so that the couples sit together?*

4. *How many distinct ways can* 8 *persons, including Ram and Shyam, sit in a row, with Ram and Shyam sitting next to each other?*

Lemma 2.1.8. *Let* N *be a finite set consisting of* n *elements. Then the number of distinct subsets of* N, *of size* k, $1 \leq k \leq n$, *equals* $\dfrac{n!}{k!\,(n-k)!}$.

Proof: It can be easily verified that the result holds for $k = 1$. Hence, we fix a positive integer k, with $2 \leq k \leq n$. Then observe that any one-to-one function $f : \{1, 2, \ldots, k\} \longrightarrow N$ gives rise to the following:

1. a set $K = \text{Im}(f) = \{f(i) : 1 \leq i \leq k\}$. The set K is a subset of N and $|K| = k$ (as f is one-to-one). Also,

2. given the set $K = \text{Im}(f) = \{f(i) : 1 \leq i \leq k\}$, one gets a one-to-one function $g : \{1, 2, \ldots, k\} \longrightarrow K$, defined by $g(i) = f(i)$, for $1 \leq i \leq k$.

Therefore, we define two sets A and B by

$$A = \{f : \{1, 2, \ldots, k\} \longrightarrow N \mid f \text{ is one-to-one}\}, \text{ and}$$

$$B = \{K \subset N \mid |K| = k\} \times \{f : \{1, 2, \ldots, k\} \longrightarrow K \mid f \text{ is one-to-one}\}.$$

Thus, the above argument implies that there is a bijection between the sets A and B and therefore, using Item 3 on Page 27, it follows that $|A| = |B|$. Also, using Lemma 2.1.4, we know that $|A| = n_{(k)}$ and $|B| = |\{K \subset N \mid |K| = k\}| \times k!$. Hence

$$\text{Number of subsets of } N \text{ of size } k = |\{K \subset N \mid |K| = k\}| = \frac{n_{(k)}}{k!} = \frac{n!}{(n-k)! \cdot k!}.$$

■

Remark 2.1.9. *Let N be a set consisting of n elements.*

1. *Then, for $n \geq k$, the number $\dfrac{n!}{k!\,(n-k)!}$ is generally denoted by $\binom{n}{k}$, and is called "n choose k". Thus, $\binom{n}{k}$ is a positive integer and equals "Number of subsets, of a set consisting of n elements, of size k".*

2. *Let K be a subset of N of size k. Then $N \setminus K$ is again a subset of N of size $n-k$. Thus, there is one-to-one correspondence between subsets of size k and subsets of size $n-k$. Thus, $\binom{n}{k} = \binom{n}{n-k}$.*

3. *The following conventions will be used: $\binom{n}{k} = \begin{cases} 0, & \text{if } n < k, \\ 1, & \text{if } k = 0. \end{cases}$*

Lemma 2.1.10. *Fix a positive integer n. Then for any two symbols x, y*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Proof: The expression $(x+y)^n = \underbrace{(x+y) \cdot (x+y) \cdot \cdots \cdot (x+y)}_{n\text{times}}$. Note that the above multiplication is same as adding all the 2^n products (appearing due to the choice of either choosing x or choosing y, from each of the above n-terms). Since either x or y is chosen from each of the n-terms, the product looks like $x^k y^{n-k}$, for some choice of $k, 0 \leq k \leq n$. Therefore, for a fixed $k, 0 \leq k \leq n$, the term $x^k y^{n-k}$ appears $\binom{n}{k}$ times as we need to choose k places from n places, for x (and thus leaving $n-k$ places for y), giving the expression $\binom{n}{k}$ as a coefficient of $x^k y^{n-k}$.

Hence, the required result follows. ■

Remark 2.1.11. *Fix a positive integer n.*

1. *Then the numbers $\binom{n}{k}$ are called* Binomial Coefficients *as they appear in the expansion of $(x+y)^n$ (see Lemma 2.1.10).*

2. *Substituting $x = y = 1$, one gets $2^n = \sum_{k=0}^{n} \binom{n}{k}$.*

3. *Observe that $(x+y+z)^n = \underbrace{(x+y+z) \cdot (x+y+z) \cdot \cdots \cdot (x+y+z)}_{n\text{times}}$. Note that in this expression, we need to choose, say*

(a) *i places from the n possible places for x ($i \geq 0$),*

(b) *j places from the remaining $n - i$ places for y ($j \geq 0$) and*

thus leaving the $n - i - j$ places for z (with $n - i - j \geq 0$). Hence, one has

$$(x + y + z)^n = \sum_{i,j \geq 0, i+j \leq n} \binom{n}{i} \cdot \binom{n-i}{j} x^i y^j z^{n-i-j}.$$

4. *The expression $\binom{n}{i} \cdot \binom{n-i}{j} = \dfrac{n!}{i!\, j!; (n-i-j)!}$ is also denoted by $\binom{n}{i,j,n-i-j}$.*

5. *Similarly, if $i_1, i_2, \ldots, i_k$ are non-negative integers, such that $i_1 + i_2 + \cdots + i_k = n$, then the coefficient of $x_1^{i_1} x_2^{i_2} \cdots x_k^{i_k}$ in the expansion of $(x_1 + x_2 + \cdots + x_k)^n$ equals*

$$\binom{n}{i_1, i_2, \ldots, i_k} = \frac{n!}{i_1! \cdot i_2! \cdots i_k!}.$$

That is,

$$(x_1 + x_2 + \cdots + x_k)^n = \sum_{\substack{i_1, \ldots, i_k \geq 0 \\ i_1 + i_2 + \cdots + i_k = n}} \binom{n}{i_1, i_2, \ldots, i_k} x_1^{i_1} x_2^{i_2} \cdots x_k^{i_k}.$$

These coefficient and called MULTINOMIAL COEFFICIENTS.

Exercise 2.1.12. *1. In a class there are 17 girls and 20 boys. A committee of 5 students is to be formed to represent the class.*

(a) *Determine the number of ways of forming the committee consisting of 5 students.*

(b) *Suppose the committee also needs to choose two different people from among themselves, who will act as "spokesperson" and "treasurer". In this case, determine the number of ways of forming a committee consisting of 5 students. Note that two committees are different if*

i. *either the members are different, or*

ii. *even if the members are the same, they have different students as spokesperson and/or treasurer.*

(c) *Due to certain restrictions, it was felt that the committee should have at least 3 girls. In this case, determine the number of ways of forming the committee consisting of 5 students (no one is to be designated as spokesperson and/or treasurer).*

2. *Determine the number of ways of arranging the letters of the word*

(a) *ABRACADABARAARCADA.*

(b) *KAGARTHALAMNAGARTHALAM.*

3. *Determine the number of ways of selecting a committee of m people from a group consisting of n_1 women and n_2 men, with $n_1 + n_2 \geq m$.*

Before proceeding further, recall the definition of partition of a non-empty set into m parts given on Page 19.

Definition 2.1.13. *Let $|A| = n$. Then the number of partitions of the set A into m-parts is denoted by $S(n, m)$. The symbol $S(n, m)$ is called the* STIRLING NUMBER OF THE SECOND KIND.

Remark 2.1.14. 1. *The following conventions will be used:*

$$S(n,m) = \begin{cases} 1, & \text{if } n = m \\ 0, & \text{if } n > 0, m = 0 \\ 0, & \text{if } n < m. \end{cases} .$$

2. *If $n > m$ then a recursive method to compute the numbers $S(n, m)$ is given in Lemma 2.1.17. A formula for the numbers $S(n, m)$ is also given in Equation (2.4).*

3. *Consider the problem of* DETERMINING THE NUMBER OF WAYS OF PUTTING m **distinguishable/distinct** BALLS INTO n **indistinguishable** BOXES WITH THE RESTRICTION THAT NO BOX IS EMPTY?

   Let $M = \{a_1, a_2, \ldots, a_m\}$ be the set of m distinct balls. Then, we observe the following:

   (a) *Since the boxes are indistinguishable, we can assume that the number of balls in each of the boxes is in a non-increasing order.*

   (b) *Let A_i, for $1 \leq i \leq n$, denote the balls in the i-th box. Then $|A_1| \geq |A_2| \geq \cdots \geq |A_n|$ and $\bigcup_{i=1}^{n} A_i = M$.*

   (c) *As each box is non-empty, each A_i is non-empty, for $1 \leq i \leq n$.*

   Thus, we see that we have obtained a partition of the set M, consisting of m elements, into n-parts, $A_1, A_2, \ldots, A_n$. Hence, the number of required ways is given by $S(m, n)$, the Stirling number of second kind.

We are now ready to look at the problem of counting the number of onto functions $f : M \longrightarrow N$. But to make the argument clear, we take an example.

Example 2.1.15. *Let $f : \{a, b, c, d, e\} \longrightarrow \{1, 2, 3\}$ be an onto function given by*

$$f(a) = f(b) = f(c) = 1,\ f(d) = 2 \text{ and } f(e) = 3.$$

Then this onto function, gives a partition $B_1 = \{a, b, c\}, B_2 = \{d\}$ and $B_3 = \{e\}$ of the set $\{a, b, c, d, e\}$ into 3-parts. Also, suppose that we are given a partition $A_1 = \{a, d\},\ A_2 = \{b, e\}$

and $A_3 = \{c\}$ *of* $\{a, b, c, d, e\}$ *into* 3*-parts. Then, this partition gives rise to the following* 3! *onto functions from* $\{a, b, c, d, e\}$ *into* $\{1, 2, 3\}$*:*

$$f_1(a) = f_1(d) = 1,\ f_1(b) = f_1(e) = 2, f_1(c) = 3, \quad i.e., \quad f_1(A_1) = 1, f_1(A_2) = 2, f_1(A_3) = 3$$
$$f_2(a) = f_2(d) = 1,\ f_2(b) = f_2(e) = 3, f_2(c) = 2, \quad i.e., \quad f_2(A_1) = 1, f_2(A_2) = 3, f_2(A_3) = 2$$
$$f_3(a) = f_3(d) = 2,\ f_3(b) = f_3(e) = 1, f_3(c) = 3, \quad i.e., \quad f_3(A_1) = 2, f_3(A_2) = 1, f_3(A_3) = 3$$
$$f_4(a) = f_4(d) = 2,\ f_4(b) = f_4(e) = 3, f_4(c) = 1, \quad i.e., \quad f_4(A_1) = 2, f_4(A_2) = 3, f_4(A_3) = 1$$
$$f_5(a) = f_5(d) = 3,\ f_5(b) = f_5(e) = 1, f_5(c) = 2, \quad i.e., \quad f_5(A_1) = 3, f_5(A_2) = 1, f_5(A_3) = 2$$
$$f_6(a) = f_6(d) = 3,\ f_6(b) = f_6(e) = 2, f_6(c) = 1, \quad i.e., \quad f_6(A_1) = 3, f_6(A_2) = 2, f_6(A_3) = 1.$$

Lemma 2.1.16. *Let* M *and* N *be two finite sets with* $|M| = m$ *and* $|N| = n$*. Then the total number of onto functions* $f : M \longrightarrow N$ *is* $n!S(m, n)$*.*

Proof: By definition, "f is onto" implies that "for all $y \in N$ there exists $x \in M$ such that $f(x) = y$. Therefore, the number of onto functions $f : M \longrightarrow N$ is 0, whenever $m < n$. So, let us assume that $m \geq n$ and $N = \{b_1, b_2, \ldots, b_n\}$. Then, we observe the following:

1. Fix i, $1 \leq i \leq n$. Then $f^{-1}(b_i) = \{x \in M | f(x) = b_i\}$ is a non-empty set as f is an onto function.
2. $f^{-1}(b_i) \cap f^{-1}(b_j) = \emptyset$, whenever $1 \leq i \neq j \leq n$ as f is a function.
3. $\bigcup_{i=1}^{n} f^{-1}(b_i) = M$ as the domain of f is M.

Therefore, if we write $A_i = f^{-1}(b_i)$, for $1 \leq i \leq n$, then $A_1, A_2, \ldots, A_n$ gives a partition of M into n-parts. Also, for $1 \leq i \leq n$ and $x \in A_i$, we note that $f(x) = b_i$. That is, for $1 \leq i \leq n$, $|f(A_i)| = |\{b_i\}| = 1$.

Conversely, each onto function $f : M \longrightarrow N$ is completely determined by

- a partition, say $A_1, A_2, \ldots, A_n$, of M into $n = |N|$ parts, and
- a one-to-one function $g : \{A_1, A_2, \ldots, A_n\} \longrightarrow N$, where $f(x) = b_i$, whenever $x \in A_j$ and $g(A_j) = b_i$.

Hence,

$$\begin{aligned} \big|\{f : M \longrightarrow N : f \text{ is onto}\}\big| &= \big|\{g : \{A_1, A_2, \ldots, A_n\} \longrightarrow N : g \text{ is one-to-one}\}\big| \\ &\qquad \times \big|\text{Partition of } M \text{ into } n\text{-parts}\big| \\ &= n!\, S(m, n). \end{aligned} \tag{2.1}$$

Lemma 2.1.17. *Let* m *and* n *be two positive integers and let* $\ell = \min\{m, n\}$*. Then*

$$n^m = \sum_{k=1}^{\ell} \binom{n}{k} k! S(m, k). \tag{2.2}$$

Proof: Let M and N be two sets with $|M| = m$ and $|N| = n$ and let A denote the set of all functions $f: M \longrightarrow N$. We compute $|A|$ using two different methods to get Equation (2.2).

The first method uses Lemma 2.1.1 to give $|A| = n^m$. The second method uses the idea of onto functions. Let $f_0: M \longrightarrow N$ be any function and let $K = f_0(M) = \{f(x) : x \in M\} \subset N$. Then, using f_0, we define a function $g: M \longrightarrow K$, by $g(x) = f_0(x)$, for all $x \in M$. Then clearly g is an onto function with $|K| = k$ for some $k, 1 \leq k \leq \ell = \min\{m, n\}$. Thus, $A = \bigcup_{k=1}^{\ell} A_k$, where $A_k = \{f : M \longrightarrow N \mid |f(M)| = k\}$, for $1 \leq k \leq \ell$. Note that $A_k \cap A_j = \emptyset$, whenever $1 \leq j \neq k \leq \ell$. Now, using Lemma 2.1.8, a subset of N of size k can be selected in $\binom{n}{k}$ ways. Thus, for $1 \leq k \leq \ell$

$$|A_k| = \big|\{K : K \subset N, |K| = k\}\big| \times \big|\{f : M \longrightarrow K \mid f \text{is onto}\}\big| = \binom{n}{k} k! S(m, k).$$

Therefore,

$$|A| = \left|\bigcup_{k=1}^{\ell} A_i\right| = \sum_{k=1}^{\ell} |A_k| = \sum_{k=1}^{\ell} \binom{n}{k} k! S(m, k).$$

■

Remark 2.1.18. *1. The numbers $S(m, k)$ can be recursively calculated using Equation (2.2).*

(a) For example, taking $n \geq 1$ and substituting $m = 1$ in Equation (2.2) gives

$$n = n^1 = \sum_{k=1}^{1} \binom{n}{k} k! S(1, k) = n \cdot 1! \cdot S(1, 1).$$

Thus, $S(1, 1) = 1$. Now, using $n = 1$ and $m \geq 2$ in Equation (2.2) gives

$$1 = 1^m = \sum_{k=1}^{1} \binom{1}{k} k! S(m, k) = 1 \cdot 1! \cdot S(m, 1).$$

Hence, the above two calculations implies that $S(m, 1) = 1$ for all $m \geq 1$.

(b) Use this to verify that $S(5, 2) = 15,\ S(5, 3) = 25, ; S(5, 4) = 10,\ S(5, 5) = 1$.

2. The problem of COUNTING THE TOTAL NUMBER OF ONTO FUNCTIONS $f: M \longrightarrow N$, WITH $|M| = m$ AND $|N| = n$ *is similar to the problem of* DETERMINING THE NUMBER OF WAYS TO PUT m **distinguishable/distinct** BALLS INTO n **distinguishable/distinct** BOXES WITH THE RESTRICTION THAT NO BOX IS EMPTY.

Example 2.1.19. *Determine the number of ways to seat 4 couples in a row if each couple seats together?*

Solution: *A couple can be thought of as one cohesive group (they are to be seated together). So, the 4 cohesive groups can be arranged in 4! ways. But a couple can sit either as "wife and husband" or "husband and wife". So, the total number of arrangements is $2^4\ 4!$.*

We end this section with the following exercises.

Exercise 2.1.20. *1. Fix positive integers n and r with $n \geq r$. Then prove the following results related to Binomial coefficients.*

(a) **Pascal's Identity:** $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$.

(b) $k\binom{n}{k} = n\binom{n-1}{k-1}$.

(c) $\binom{k}{\ell}\binom{n}{k} = \binom{n}{\ell}\binom{n-\ell}{k-\ell}$.

(d) $\sum_{k=\ell}^{n} \binom{k}{\ell}\binom{n}{k} = \binom{n}{\ell}2^{n-\ell}$.

(e) $\binom{m+n}{\ell} = \sum_{k=0}^{\ell} \binom{m}{k}\binom{n}{\ell-k}$.

(f) $\binom{n+r+1}{r} = \sum_{\ell=0}^{r} \binom{n+\ell}{\ell}$.

(g) $\binom{n+1}{r+1} = \sum_{\ell=r}^{n} \binom{\ell}{r}$.

(h) $\binom{n}{\ell} = \sum_{k=0}^{\ell} \binom{t}{k}\binom{n-t}{\ell-k}$, *for any* $t, 0 \leq t \leq n$.

2. chap1:exe5:21 Prove the following results.

(a) Let S be a finite set consisting of n elements, $n \geq 0$. Then S has exactly 2^n subsets.

(b) $\binom{n}{0} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^{n-1}$, *for all* $n \in \mathbb{N}$.

(c) $\sum_{k\geq 0} \binom{n}{2k} = \sum_{k\geq 0} \binom{n}{2k+1}$, *for all* $n \in \mathbb{N}$.

3. Determine the number of n-letter words that are formed using r A's and $n-r$ B's?

4. Determine the number of ways of selecting r distinguishable objects from n distinguishable objects when $n \geq r$?

5. How many ways are there to distribute 20 distinguishable toys among 4 children so that each children gets the same number of toys?

6. In how many ways can m **distinguishable** *balls be put into n* **indistinguishable** *boxes, such that NO box is empty?*

7. In how many ways can m **distinguishable** *balls be put into n* **indistinguishable** *boxes?*

8. Determine a recurrence relation to get the values of $S(n,k)$'s?

9. For a positive integer n, let $b(n)$ denote the number of partitions of the set $\{1, 2, \ldots, n\}$. Then $b(n) = \sum_{m=0}^{n} S(n,m)$ is called the n^{th} Bell number. By definition, $b(0) = 1 = b(1)$. For $2 \leq n \leq 10$, compute the values of $b(n)$.

10. *Let X be a non-empty set. Suppose X_O and X_E, respectively, denote the set of all even and odd subsets of X. Describe a bijection to prove that $|X_O| = |X_E|$. Hence or otherwise, prove that*
$$\sum_{k=0}^{[n/2]} \binom{n}{2k} = \sum_{k=0}^{[n/2]} \binom{n}{2k+1} = 2^{n-1}.$$

11. *Find a closed form expression for $\sum_{k=0}^{n}(2k+1)\binom{n}{2k+1}$ and $\sum_{k=0}^{n}(5k+3)\binom{n}{2k+1}$*

12. *Fix positive integers m and n and let $S = \{f : M \longrightarrow N \mid |M| = m, |N| = n+1\}$. Compute $|S|$ in two ways to obtain*
$$(n+1)^m = \sum_{k=0}^{m} \binom{m}{k} n^k.$$

13. *Suppose 13 people get on the lift at level* **0**. *suppose the lift stops at the levels* **1**, **2**, **3**, **4** *and* **5**. *If all the people get down at some level, calculate the number of ways they can get down so that* **at least one person gets down at each level**.

14. *A function $f : N \longrightarrow N$ is said to be* IDEMPOTENT *if $f\big(f(x)\big) = f(x)$ for all $x \in N$. If $|N| = n$, prove that the number of such functions equals $\sum_{k=1}^{n} k^{n-k}\binom{n}{k}$.*

2.1.2 Indistinguishable Balls and Distinguishable Boxes

Example 2.1.1. *1. Determine the number of distinct strings that can be formed using 3 A's and 6 B's? What if we are interested in the strings that are formed using $+$'s and 1's in place of A'a and B's?*

Solution: *Note that the 3 A's are indistinguishable among themselves and the same holds for 6 B's. Thus, we need to find 3 places, from the $9 = 3 + 6$ places, for the A's. Hence, the answer is $\binom{9}{3}$. The answer will remain the same as we just need to replace A's with $+$'s and B's with 1's in any string of 3 A's and 6 B's.*

2. Determine the number of solutions to the equation $x_1+x_2+x_3+x_4 = 6$, where each $x_i \in \mathbb{Z}$ and $0 \le x_i \le 6$.

Solution: *We will show that this problem is same as Example 2.1.1.1. Take a sequence, say $+111+1+11$ of 3 $+$'s and 6 1's. Then, one can think of this sequence representing a solution $0+3+1+2$ of $x_1+x_2+x_3+x_4 = 6$, where $x_1 = 0, x_2 = 3, x_3 = 1$ and $x_4 = 2$. In the same way, a solution, say $x_1 = 5, x_2 = 0, x_3 = 0$ and $x_4 = 1$ of the given equation gives rise to a sequence $11111+++1$ of 3 $+$'s and 6 1's. Thus the total number of solutions is*
$$\binom{9}{3} = \binom{9}{6} = \binom{6+(4-1)}{6}.$$

We now generalize this example to a general case.

Lemma 2.1.2. *Determine the number of solutions to the equation $x_1+x_2+\cdots+x_n = m$, where each $x_i \in \mathbb{Z}$ and $0 \leq x_i \leq m$.*

Proof: Note that the number m can be replaced with m 1's (or m indistinguishable symbols). Once this is done, using the idea in Example 2.1.1.2, we see that it is enough to find the number of distinct strings formed using $n-1$ +'s and m 1's. As the +'s are indistinguishable among themselves and the same holds for 1's, we get

$$\binom{n-1+m}{m} = \binom{n-1+m}{n-1} = \binom{m+\binom{n-1}{m}}{m}.$$

■

Remark 2.1.3. *Observe that solutions in non-negative integers to the equation $x_1 + x_2 + \cdots + x_n = m$ is same as the following problems.*

1. *Determine the number of ways to put m* **indistinguishable** *balls into n* **distinguishable** *boxes. Hint: Since the balls are indistinguishable, we are interested in finding out the number of balls in each of the distinguishable boxes. So, let the boxes be numbered $1, 2, \ldots, n$ and let x_i, for $1 \leq i \leq n$, denote the number of balls in the i-th box.*

2. *Determine the number of non-decreasing sequences of length m using the numbers $1, 2, \ldots, n$. Hint: Since we are looking at a non-decreasing sequences, we note that the sequence is determined if we know the number of times a particular number has appeared in the sequence. So, let x_i, for $1 \leq i \leq n$, denote the number of times the number i has appeared in the sequence.*

Exercise 2.1.4.

1. *How many 4-letter words (with repetition) are there with the letters in alphabetical order?*

2. *In how many ways can m* **indistinguishable** *balls be put into n* **distinguishable** *boxes with the restriction that no box is empty.*

3. *How many 26-letter permutations of the alphabet have no 2 vowels together?*

4. *How many 26-letter permutations of the alphabet have at least two consonants between any two vowels?*

5. *How many ways can 10 men and 7 women be seated in a row with no 2 women next to each other?*

6. *How many ways can 8 persons, including Ram and Shyam, sit in a row with Ram and Shyam not sitting next to each other?*

7. *How many arrangements of the letters of KAGARTHALAMNAGARTHALAM have the vowels in alphabetical order?*

8. *How many arrangements of the letters of RECURRENCERELATION have no* 2 *vowels adjacent?*

9. *How many nonnegative integer solutions are there to the equation*

$$x_1 + x_2 + \cdots + x_5 = 67?$$

10. *How many positive integer solutions are there to the equation*

$$x_1 + x_2 + \cdots + x_5 = 67?$$

11. *How many nonnegative integer solutions are there to the inequality*

$$x_1 + x_2 + \cdots + x_5 \leq 68?$$

12. *With repetition NOT allowed and order counting, how many ways are there to select* r *things from* n *distinguishable things?*

13. *With repetition NOT allowed and order NOT counting, how many ways are there to select* r *things from* n *distinguishable things?*

14. *With repetition allowed and order counting, how many ways are there to select* r *things from* n *distinguishable things?*

15. *With repetition allowed and order NOT counting, how many ways are there to select* r *things from* n *distinguishable things?*

16. *How many ways are there to arrange the letters in* $ABRACADABRAARCADA$ *so that the first*

    (a) A *precedes the first* B*?*

    (b) B *precedes the first* A *and the first* D *precedes the first* C*?*

    (c) B *precedes the first* A *and the first* A *precedes the first* C*?*

17. *How many ways are there to arrange the letters in* $KAGARTHALAMNAGARTHATAM$ *with the first*

    (a) A *preceding the first* T*?*

    (b) M *preceding the first* G *and the first* H *preceding the first* A*?*

    (c) M *preceding the first* G *and the first* T *preceding the first* G*?*

18. *How many ways are there to distribute* 50 *balls to* 5 *persons if Ram and Shyam together get no more than* 30 *and Mohan gets at least* 10*?*

19. *How many ways can we pick* 20 *letters from* 10 *A's,* 15 *B's and* 15 *C's?*

20. *How many ways are there to select* 12 *integers from the set* $\{1, 2, 3, \ldots, 100\}$ *such that the positive difference between any two of the* 12 *integers is at least* 3*?*

21. *How many* 10*-element subsets of the alphabet have a pair of consecutive letters?*

22. *Determine the number of ways to sit* 5 *men and* 7 *women so that none of the men are sitting next to each other?*

23. *Determine the number of ways to sit* 10 *men and* 7 *women so that none of the women are sitting next to each other?*

24. *Determine the number of ways to sit* 8 *persons, including Ram and Shyam, with Ram and Shyam NOT sitting next to each other?*

2.1.3 Indistinguishable Balls and Indistinguishable Boxes

To study the number of onto functions $f : M \longrightarrow N$, we were lead to the study of "partition of a set consisting of m elements into n parts". In this section, we study the partition of a number m into n parts and look at a few problems in which this idea can be used.

Definition 2.1.5 (Partition of a number)**.** *A partition of a positive integer* m *into* n *parts, is a non-increasing sequence of positive numbers* $x_1, x_2, \ldots, x_n$ *such that* $\sum_{k=1}^{n} x_k = m$*. The number of such partitions is denoted by* $\Pi(m, n)$*.*

Remark 2.1.6.

1. *For example, the distinct partitions of* 7 *into* 4 *parts are given by* $4+1+1+1$, $3+2+1+1$, $2+2+2+1$*. Hence,* $\Pi(7, 4) = 3$*. Verify that* $\Pi(7, 2) = 3$ *and* $\Pi(7, 3) = 4$*.*

2. *For a fixed positive integer* m*,* $\Pi(m)$ *denotes the number of partitions of* m*. Hence,* $\Pi(m) = \sum_{k=1}^{m} \Pi(m, k)$*. Verify that* $\Pi(7) = 15$*.*

3. *By convention, we let* $\Pi(0, 0) = 1$ *and* $\Pi(m, n) = 0$*, whenever* $n > m$*.*

We are now ready to associate the study of partitions of a number with the following problems.

Example 2.1.7.

1. *Determine the number of ways of putting* m **indistinguishable** *balls into* n **indistinguishable** *boxes with the restriction that no box is empty?*

   Solution: *Since the balls are indistinguishable balls, the problem reduces to counting the number of balls in each box with the condition that no box is empty. As the boxes are also indistinguishable, they can be arranged in such a way that the number of balls inside them are in non-increasing order. Hence, we have the answer as* $\Pi(m, n)$*.*

2. *Determine the number of ways to put m* **indistinguishable** *balls into n* **indistinguishable** *boxes?*

   Solution: *The problem can be rephrased as follows: "suppose that each box already has 1 ball, i.e., initially, each of the n boxes are non-empty. Now let us determine the number of ways of putting m indistinguishable balls into the n indistinguishable boxes that are already non-empty." This new problem is same as "in how many ways can $m+n$* **indistinguishable** *balls be put into n* **indistinguishable** *boxes?" Therefore, the answer to our problem reduces to computing $\Pi(m+n, n)$.*

3. *Use the above idea to show that $\Pi(2m, m) = \Pi(m)$ for any positive integer m. Hint: $\Pi(2m, m)$ corresponds to "putting $2m$ indistinguishable balls into m indistinguishable boxes with the condition that no box empty", where as $\Pi(m)$ corresponds to "putting m balls into m indistinguishable boxes".*

Exercise 2.1.8. 1. *Calculate $\Pi(n)$, for $n = 1, 2, 3, \ldots, 8$.*

2. *For a fixed positive integer n, determine a recurrence relation for the numbers $\Pi(n, m)$. Hint: Use the previous exercise to relate $\Pi(m, n)$ with $\Pi(n-1, m-1)$ and/or $\Pi(n-m, m)$?*

2.1.4 Round Table Configurations

Till now, we have been looking at problems that required arranging the objects into a row. That is, we differentiated between the arrangements $ABCD$ and $BCDA$. In this section, we briefly study the problem of arranging the objects into a circular fashion. That is, if we are arranging the four distinct chairs, named A, B, C and D, at a round table then the arrangements $ABCD$ and the arrangement $BCDA$ are the same. That is, the main problem that we come across circular arrangements as compared to problems in the previous sections is "there is no object that can truly be said to be placed at the number 1 position". The problems related with round table configurations will be dealt at length in Chapter 4.

So, to get distinct arrangements at a round table, we need to fix an object and assign it the number 1 position and study the distinct arrangement of the other $n-1$ objects relative to the object which has been assigned position 1. We will look at two examples to understand this idea.

Example 2.1.1. 1. *Determine the number of ways to sit 8 persons at a round table?*

Solution: METHOD 1: *Let us number the chairs as $1, 2, \ldots, 8$. Then, we can pick one of the person and ask him/her to sit on the chair that has been numbered 1. Then relative to this person, the other persons (7 of them) can be arranged in $7!$ ways. So, the total number of arrangements is $7!$.*

METHOD 2: *The total number of arrangements of 8 persons if they are to be seated in a row is $8!$. Since the cyclic arrangement $P_1P_2\cdots P_8$ is same as the arrangement $P_8P_1P_2\cdots P_7$*

and so on, we need to divide the number $8!$ *by* 8 *to get the actual number as* $7!$*.*

2. *Recall Example 2.1.19. Suppose we are now interested in making the* 4 *couples sit in a round table.*
   Solution: *Note that using Example 2.1.1.1, the* 4 *cohesive units can be arranged in* $3!$ *ways. But we can still have the couples to sit either as "wife and husband" or "husband and wife'. Hence, the required answer is* $2^4 3!$*.*

1. Determine the number of ways to sit 5 men and 7 women at a round table with NO 2 men sitting next to each other?

2. Determine the number of ways to sit 10 men and 7 women at a round table so that NO 2 women sitting next to each other?

3. Determine the number of ways to sit 8 persons, including Ram and Shyam, at a round table with Ram and Shyam NOT sitting next to each other?

4. Determine the number of ways to sit 8 persons, including Ram and Shyam, at a round table with Ram and Shyam NOT sitting diametrically opposite to each other?

5. Determine the number of ways to select 6 men from 25 men who are sitting at a round table if NO adjacent men are to be chosen?

2.2 Lattice Paths

Consider a lattice of integer lines in the plane. The set $S = \{(m, n) : \ m, n = 0, 1, 2, \ldots\}$ are said to be the points of the lattice and the lines joining these points are called the edges of the lattice. Now, let us fix two points in this lattice, say (m_1, n_1) and m_2, n_2, with $m_2 \geq m_1$ and $n_2 \geq n_1$. Then we define an INCREASING/LATTICE PATH from (m_1, n_1) to (m_2, n_2) to be a subset $\{e_1, e_2, \ldots, e_k\}$ of S such that

1. either $e_1 = (m_1, n_1 + 1)$ or $e_1 = (m_1 + 1, n_1)$;

2. either $e_k = (m_2, n_2 - 1)$ or $e_k = (m_2 - 1, n_2)$; and

3. if we represent the tuple $e_i = (a_i, b_i)$, for $1 \leq i \leq k$, then for $2 \leq j \leq k$,

   (a) either $a_j = a_{j-1}$ and $b_j = b_{j-1} + 1$

   (b) or $b_j = b_{j-1}$ and $a_j = a_{j-1} + 1$.

That is, the movement on the lattice is either to the RIGHT or UP (see Figure 2.1). Now, let us look at some of the questions related to this topic.

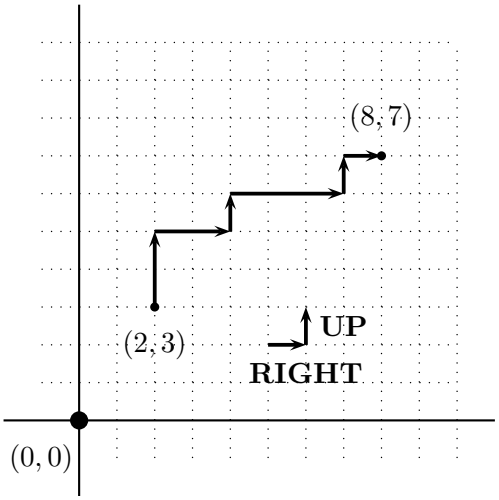


Figure 2.1: A lattice with a lattice path from $(2, 3)$ to $(8, 7)$

Example 2.2.1. *1. Determine the number of lattice paths from $(0, 0)$ to (m, n).*

Solution: *Note that at each stage, the coordinates of the lattice path increases, by exactly one positive value, either in the X-coordinate or in the Y-coordinate. Therefore, to reach (m, n) from $(0, 0)$, the total increase in the X-direction is exactly m and in the Y-direction is exactly n. That is, each lattice path is a sequence of length $m + n$, consisting of m R's (movement along X-axis/RIGHT) and n U's (movement along the Y-axis/UP). So, we need to find m places for the R's among the $m + n$ places (R and U together). Thus, the required answer is $\binom{m+n}{m}$.*

2. Use the method of lattice paths to prove the following result on Binomial Coefficients:

$$\sum_{\ell=0}^{m} \binom{n+\ell}{\ell} = \binom{n+m+1}{m}.$$

Solution: *Observe that the right hand side corresponds to the number of lattice paths from $(0, 0)$ to $(m, n + 1)$, whereas the left hand side corresponds to the number of lattice paths from $(0, 0)$ to (ℓ, n), where $0 \leq \ell \leq m$.*

Now, fix $\ell, 0 \leq \ell \leq m$. Then to each lattice path from $(0, 0)$ to (ℓ, n), say P, we adjoin the path $Q = U \underbrace{RR \cdots R}_{m-\ell \text{ times}}$. Then the path $P \cup Q$, corresponding to a lattice path from $(0, 0)$ to (ℓ, n) and from (ℓ, n) to $(\ell, n + 1)$ and finally from $(\ell, n + 1)$ to $(m, n + 1)$, gives a lattice path from $(0, 0)$ to $(m, n + 1)$. These lattice paths, as we vary ℓ, for $0 \leq \ell \leq m$, are all distinct and hence the result follows.

Exercise 2.2.2. *1. Determine the number of solutions in non-negative integers to the system $x_0 + x_1 + x_2 + \cdots + x_k = n$. Also, find a bijection to prove that the above problem is same as determining the number of lattice paths from $(0, 0)$ to (n, k).*

2. *Construct a proof of* $\binom{n}{0}+\binom{n}{1}+\cdots+\binom{n}{n}=2^n$ *using lattice paths. [Hint:* $\binom{n}{k}$ *represents a lattice path from* $(0,0)$ *to* $(n-k,k)$ *and any subset of* $\{1,2,\ldots,n\}$ *gives positions, where we need to take the step* U*.]*

3. *Construct a proof of* $\binom{n}{0}^2+\binom{n}{1}^2+\cdots+\binom{n}{n}^2=\binom{2n}{n}$ *using lattice paths. [Hint:* $\binom{n}{k}$ *represents a lattice path from* $(0,0)$ *to* $(n-k,k)$ *and* $\binom{n}{k}=\binom{n}{n-k}$ *also represents a lattice path from* $(n-k,k)$ *to* (n,n)*.]*

2.2.1 Catalan Numbers

Determine the number of lattice paths from $(0,0)$ to (n,n) that do not go above the line $Y=X$ (see Figure 2.2).

Solution: The first move from $(0,0)$ is R (corresponding to moving to the point $(1,0)$) as we are not allowed to go above the line $Y=X$. So, in principle, all our lattice paths are from $(1,0)$ to (n,n) with the condition that these paths do not cross the line $Y=X$.

Using Example 2.2.1.1, the total number of lattice paths from $(1,0)$ to (n,n) is $\binom{2n-1}{n}$. So, we need to subtract from $\binom{2n-1}{n}$ a number, say N_0, where N_0 equals the number of lattice paths from $(1,0)$ to (n,n) that **cross** the line $Y=X$.

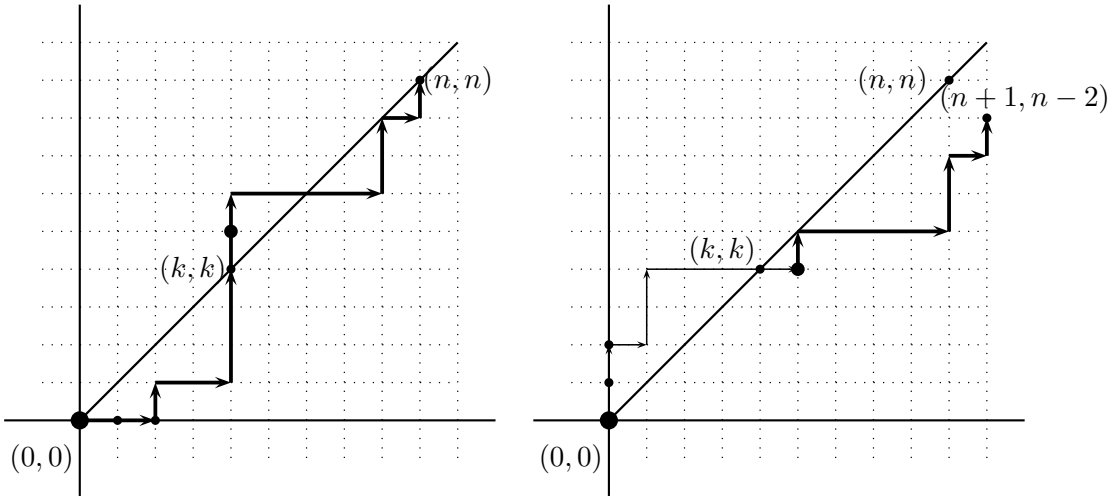


Figure 2.2: Lattice paths giving the mirror symmetry from $(0,0)$ to (k,k)

To compute the value of N_0, we decompose each lattice path that crosses the line $Y=X$ into two sub-paths. Let P be a path from $(1,0)$ to (n,n) that crosses the line $Y=X$. Then this path crosses the line $Y=X$ for the first time at some point, say (k,k), $1\leq k\leq n-1$.

Claim: there exists a 1-1 correspondence between lattice paths from $(1,0)$ to (n,n) that crosses the line $Y=X$ and the lattice paths from $(0,1)$ to $(n+1,n-1)$.

Let $P=P_1P_2\ldots P_{2k-1}P_{2k}P_{2k+1}\ldots P_{2n-1}$ be the path from $(1,0)$ to (n,n) that crosses the line $Y=X$. Then $P_1,P_2,\ldots,P_{2k-2}$ consist of a sequence of $(k-1)$ R's and $(k-1)$ U's. Also, $P_{2k-1}=P_{2k}=U$ and the sub-path $P_{2k+1}P_{2k+2}\ldots P_{2n-1}$ consist of a sequence that has

$(n-k)$ R's and $(n-k-1)$ U's. Also, for any $i, 1 \leq i \leq 2k-2$, in the sub-path $P_1P_2 \dots P_i$,

$$\#\text{ of R's } = |\{j : P_j = R, 1 \leq j \leq i\}| \geq |\{\ell : P_\ell = U, 1 \leq \ell \leq i\}| = \#\text{ of U's.} \tag{2.1}$$

Now, P is mapped to a path Q, such that $Q_i = \begin{cases} P_i, & \text{if } 2k+1 \leq i \leq 2n-1, \\ \{R,U\} \setminus P_i, & \text{if } 1 \leq i \leq 2k. \end{cases}$
Then we see that the path Q consists of exactly $(k-1)+2+(n-k) = n+1$ R's and $(k-1)+(n-k-1) = n-2$ U's. Also, the condition that $Q_i = \{R,U\} \setminus P_i$, for $1 \leq i \leq 2k$, implies that the path Q starts from the point $(0,1)$. Therefore, Q is a path that starts from $(0,1)$ and consists of $(n+1)$ R's and $(n-2)$ U's and hence Q ends at the point $(n+1, n-1)$.

Also, if Q' is a path from $(0,1)$ to $(n+1, n-1)$, then Q' consists of $(n+1)$ R's and $(n-2)$ U's. So, in any such sequence an instant occurs when the number of R's exceeds the number of U's by 2. Suppose this occurrence happens for the first time at the $(2k)^{\text{th}}$ instant, for some $k, 1 \leq k \leq n-1$. Then there are $(k+1)$ R's and $(k-1)$ U's till the first $(2k)^{\text{th}}$ instant and $(n-k)$ R's and $(n-1-k)$ U's in the remaining part of the sequence. So, Q' can be replaced by a path P', such that $(P')_i = \begin{cases} (Q')_i, & \text{if } 2k+1 \leq i \leq 2n-1, \\ \{R,U\} \setminus (Q')_i, & \text{if } 1 \leq i \leq 2k. \end{cases}$

It can be easily verified that P' is a lattice path from $(1,0)$ to (n,n) that crosses the line $Y = X$. Thus, the proof of the claim is complete.

Hence, the number of lattice paths from $(1,0)$ to (n,n) that crosses the line $Y = X$ equals the number of lattice paths from $(0,1)$ to $(n+1, n-1)$. But, using Example 2.2.1.1, the number of lattice paths from $(0,1)$ to $(n+1, n-1)$ equals $\binom{2n-1}{n+1}$. Hence, the number of lattice paths from $(0,0)$ to (n,n) that does not go above the line $Y + X$ is

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

This number is popularly known as the n^{th} Catalan Number, denoted C_n.

Remark 2.2.3. *The book titled "enumerative combinatorics" by Stanley [7] gives a comprehensive list of places in combinatorics where Catalan numbers appear. A few of them are mentioned here for the inquisitive mind. The equivalence between these problems can be better understood after the chapter on recurrence relations.*

1. *Suppose in an election two candidates A and B get exactly n votes. Then C_n equals the number of ways of counting the votes such that candidate A is always ahead of candidate B. For example,*

   $$C_3 = 5 = |\{AAABBB, AABABB, AABBAB, ABAABB, ABABAB\}|.$$

2. *Suppose, we need to multiply $n+1$ given numbers, say $a_1, a_2, \dots, a_{n+1}$. Then the different ways of multiplying these numbers, without changing the order of the elements, equals C_n. For example,*

   $$C_3 = 5 = |\{(((a_1a_2)a_3)a_4), ((a_1a_2)(a_3a_4)), ((a_1(a_2a_3))a_4), (a_1((a_2a_3)a_4)), (a_1(a_2(a_3a_4)))\}|.$$

3. *C_n also equals the number of ways that a convex $(n+2)$-gon cab be sub-divided into triangles by its diagonals so that no two diagonals intersect inside the $(n+2)$-gon? For example, for a pentagon, the different ways are*

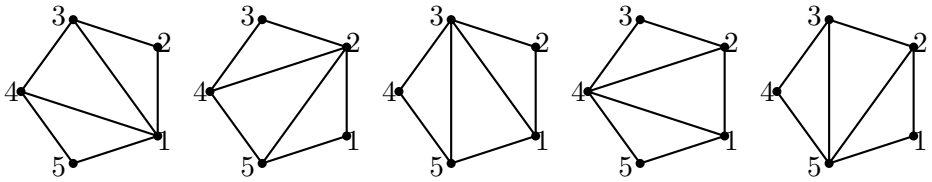


Figure 2.3: Different divisions of pentagon

4. *C_n is also equal to the number of full binary trees on $2n+1$ vertices, where recall that a full binary tree is a rotated binary tree in which every node either has exactly two offsprings or has no offspring.*

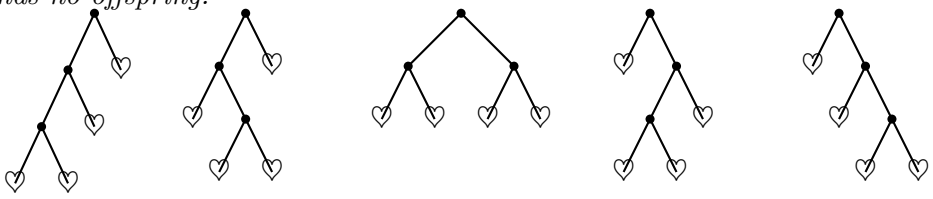


Figure 2.4: Full binary trees on 7 vertices (or 4 leaves)

5. *C_n is also equal to the number of Dyck paths from $(0,0)$ to $(2n,0)$, where recall that a Dyck path is a movement on an integer lattice in which each step is either in the North East or in the South East direction (so the only movement from $(0,0)$ is either to $(1,1)$ or to $(1,-1)$).*

6. *C_n also equals the number of n nonintersecting chords joining $2n$ points on the circumference of a circle.*

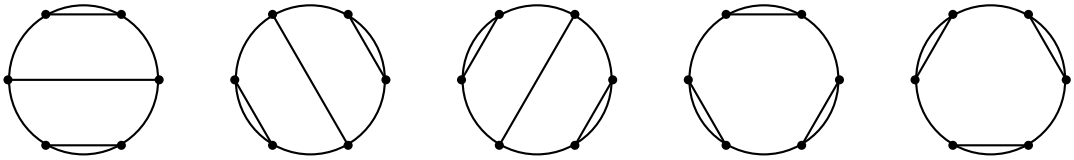


Figure 2.5: Non-intersecting chords using 6 points on the circle

7. *C_n also equals the number of integer sequences that satisfy $1 \leq a_1 \leq a_2 \leq \cdots \leq a_n$ and $a_i \leq i$, for all $i, 1 \leq i \leq n$.*

The following article has been taken from [1].

Let A_n denote the set of all lattice paths from $(0,0)$ to (n,n) and let $B_n \subset A_n$ denote the set of

all lattice paths from $(0,0)$ to (n,n) that does not go above the line $Y = X$. Then, the numerical values of $|A_n|$ and $|B_n|$ imply that $(n+1) \cdot |B_n| = |A_n|$. The question arises:
can we find a partition of the set A_n into $(n+1)$-parts, say $S_0, S_1, S_2, \ldots, S_n$ such that $S_0 = B_n$ and $|S_i| = |S_0|$, for $1 \leq i \leq n$.

The answer is in affirmative. Observe that any path from $(0,0)$ to (n,n) has n right moves. So, the path is specified as soon as we know the successive right moves $R_1, R_2, \ldots, R_n$, where R_i equals ℓ if and only if R_i lies on the line $Y = \ell$. For example, in Figure 2.2, $R_1 = 0,\ R_2 = 0,\ R_3 = 1,\ R_4 = 1, \ldots$. These R_i's satisfy

$$0 \leq R_1 \leq R_2 \leq \cdots \leq R_n \leq n. \tag{2.2}$$

That is, any element of A_n can be represented by an ordered n-tuple $(R_1, R_2, \ldots, R_n)$ satisfying Equation (2.2). Conversely, it can be easily verified that any ordered n-tuple $(R_1, R_2, \ldots, R_n)$ satisfying Equation (2.2) corresponds to a lattice path from $(0,0)$ to (n,n). Note that, using Remark 2.2.3.7, among the above n-tuples, the tuples that satisfy $R_i \leq i-1$, for $1 \leq i \leq n$ are elements of B_n, and vice-versa. Now, we use the tuples that represent the elements of B_n to get $n+1$ maps, $f_0, f_1, \ldots, f_n$, in such a way that $f_j(B_n)$ and $f_k(B_n)$ are disjoint, for $0 \leq j \neq k \leq n$, and $A_n = \bigcup_{k=0}^{n} f_k(B_n)$. In particular, for a fixed $k, 0 \leq k \leq n$, the map $f_k : B_n \longrightarrow A_n$ is defined by

$$f_k\big((R_1, R_2, \ldots, R_n)\big) = (R_{i_1} \oplus_{n+1} k, R_{i_2} \oplus_{n+1} k, \ldots, R_{i_n} \oplus_{n+1} k),$$

where $\oplus_{n+1}$ denotes addition modulo $n+1$ and $i_1, i_2, \ldots, i_n$ is a rearrangement of the numbers $1, 2, \ldots, n$ such that $0 \leq R_{i_1} \oplus_{n+1} k \leq R_{i_2} \oplus_{n+1} k \leq \cdots \leq R_{i_n} \oplus_{n+1} k$. The readers are advised to prove the following exercise as they give the required partition of the set A_n.

Exercise 2.2.4.

1. *Fix $k, 0 \leq k \leq n$. Then prove that $f_k(R) \neq f_k(R')$, whenever $R, R' \in B_n$ and $R \neq R'$.*
2. *For $k \neq 0$, prove that $f_k(B_n) \cap B_n = \emptyset$.*
3. *For each $j, k,\ 0 \leq j \neq k \leq n,\ f_j(B_n) \cap f_k(B_n) = \emptyset$.*
4. *Given any path $P \in A_n$, there exists a positive integer $k,\ 0 \leq k \leq n$ and a path $R \in B_n$ such that $f_k(R) = P$.*

2.3 Some Generalizations

1. Let n, k be non-negative integers with $0 \leq k \leq n$. Then in Lemma 2.1.8, "the number of ways of choosing a subset of size k from a set consisting of n elements" was denoted by the binomial coefficients, $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$. Since, for each $k, 0 \leq k \leq n$, $(n-k)!$ divides $n!$, let us think of $\binom{n}{k} = \dfrac{n \cdot (n-1) \cdots (n-k+1)}{k!}$. With this understanding, the numbers $\binom{n}{k}$ can

be generalized. That is, in the generalized form, for any $n \in \mathbb{C}$ and for any non-negative integer k, one has

$$\binom{n}{k} = \begin{cases} 0, & \text{if } k < 0 \\ 0, & \text{if } k = 0, n \neq k \\ 1, & \text{if } n = k \\ \dfrac{n \cdot (n-1) \cdots (n-k+1)}{k!}, & \text{otherwise.} \end{cases} \tag{2.3}$$

With the notations as above, one has the following theorem and is popularly known as the generalized binomial theorem. We state it without proof.

Theorem 2.3.1. *Let n be any real number. Then*

$$(1+x)^n = 1 + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{r}x^r + \cdots .$$

In particular, if $(1-x)^{-1} = 1 + x + x^2 + x^3 + \cdots$ and if $a, b \in \mathbb{R}$ with $|a| < |b|$, then

$$(a+b)^n = b^n \left(1 + \frac{a}{b}\right)^n = b^n \sum_{r \geq 0} \binom{n}{r} \left(\frac{a}{b}\right)^r = \sum_{r \geq 0} \binom{n}{r} a^r b^{n-r}.$$

Let us now understand Theorem thm:generalized:binomial through the following examples.

(a) Let $n = \dfrac{1}{2}$. In this case, for $k \geq 1$, Equation (2.3) gives

$$\binom{\frac{1}{2}}{k} = \frac{\frac{1}{2} \cdot \left(\frac{1}{2} - 1\right) \cdots \left(\frac{1}{2} - k + 1\right)}{k!} = \frac{1 \cdot (-1) \cdot (-3) \cdots (3-2k)}{2^k k!} = \frac{(-1)^{k-1}(2k-2)!}{2^{2k-1}(k-1)!k!}.$$

Thus,

$$(1+x)^{1/2} = \sum_{k \geq 0} \binom{\frac{1}{2}}{k} x^k = 1 + \frac{1}{2}x + \frac{-1}{2^3}x^2 + \frac{1}{2^4}x^3 + \sum_{k \geq 4} \frac{(-1)^{k-1}(2k-2)!}{2^{2k-1}(k-1)!k!}x^k.$$

This can also be obtained using the Taylor series expansion of $f(x) = (1+x)^{1/2}$ around $x = 0$. Recall that the Taylor series expansion of $f(x)$ around $x = 0$ equals $f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \sum\limits_{k \geq 3} \frac{f^{(k)}(0)}{k!}x^k$, where $f(0) = 1, f'(0) = \frac{1}{2}, f''(0) = \frac{-1}{2^2}$ and in general $f^{(k)}(0) = \frac{1}{2} \cdot (\frac{1}{2} - 1) \cdots (\frac{1}{2} - k + 1)$, for $k \geq 3$.

(b) Let $n = -r$, where r is a positive integer. Then, for $k \geq 1$, Equation (2.3) gives

$$\binom{-r}{k} = \frac{-r \cdot (-r-1) \cdots (-r-k+1)}{k!} = (-1)^k \binom{r+k-1}{k}.$$

Thus,

$$(1+x)^n = \frac{1}{(1+x)^r} = 1 - rx + \binom{r+1}{2}x^2 + \sum_{k \geq 3} \binom{r+k-1}{k}(-x)^k.$$

The readers are advised to get the above expression using the Taylor series expansion of $(1+x)^n$ around $x = 0$.

2. Let $n, m \in \mathbb{N}$. Recall the identity $n^m = \sum_{k=0}^{m} \binom{n}{k} k! S(m,k) = \sum_{k=0}^{n} \binom{n}{k} k! S(m,k)$ that appeared in Lemma 2.1.17 (see Equation (2.2)). We note that for a fixed positive integer m, the above identity is same as the matrix product $X = AY$, where

$$X = \begin{bmatrix} 0^m \\ 1^m \\ 2^m \\ 3^m \\ \vdots \\ n^m \end{bmatrix}, \quad A = \begin{bmatrix} \binom{0}{0} & 0 & 0 & 0 & \cdots & 0 \\ \binom{1}{0} & \binom{1}{1} & 0 & 0 & \cdots & 0 \\ \binom{2}{0} & \binom{2}{1} & \binom{2}{2} & 0 & \cdots & 0 \\ \binom{3}{0} & \binom{3}{1} & \binom{3}{2} & \binom{3}{3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \binom{n}{0} & \binom{n}{1} & \binom{n}{2} & \binom{n}{3} & \cdots & \binom{n}{n} \end{bmatrix} \text{ and } Y = \begin{bmatrix} 0!S(m,0) \\ 1!S(m,1) \\ 2!S(m,2) \\ \vdots \\ n!S(m,n) \end{bmatrix}.$$

Hence, if we know the inverse of the matrix A, we can write $Y = A^{-1}X$. Check that

$$A^{-1} = \begin{bmatrix} \binom{0}{0} & 0 & 0 & 0 & \cdots & 0 \\ -\binom{1}{0} & \binom{1}{1} & 0 & 0 & \cdots & 0 \\ \binom{2}{0} & -\binom{2}{1} & \binom{2}{2} & 0 & \cdots & 0 \\ -\binom{3}{0} & \binom{3}{1} & -\binom{3}{2} & \binom{3}{3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (-1)^n\binom{n}{0} & (-1)^{n-1}\binom{n}{1} & (-1)^{n-2}\binom{n}{2} & (-1)^{n-3}\binom{n}{3} & \cdots & \binom{n}{n} \end{bmatrix}.$$

This gives us a way to calculate the Stirling numbers of second kind as a function of binomial coefficients. That is, verify that

$$S(m,n) = \frac{1}{n!} \sum_{k \geq 0} (-1)^k \binom{n}{k} (n-k)^m. \tag{2.4}$$

The above ideas imply that for all non-negative integers n the identity

$$a(n) = \sum_{k \geq 0} \binom{n}{k} b(k) \text{ holds if and only if } b(n) = \sum_{k \geq 0} (-1)^k \binom{n}{k} a(k) \text{ holds.}$$

3. We end this chapter with an example which has a history (see [3]) of being solved by many mathematicians such as Montmort, N. Bernoulli and de Moivre. We present the idea that was proposed by Euler.

Example 2.3.2. *On a rainy day, n students leave their umbrellas (which are indistinguishable) outside their examination room. Determine the number of ways of collecting the umbrellas so that no student collects the correct umbrella when they finish the examination? This problem is generally known by the* Derangement problem.

Solution: Let the students be numbered $1, 2, \ldots, n$ and suppose that the i^{th} student has the umbrella numbered i, $1 \leq i \leq n$. So, we are interested in the number of permutations

of the set $\{1, 2, \ldots, n\}$ such that the number i is not at the i^{th} position, for $1 \leq i \leq n$. Let D_n represent the number of derangements. Then it can be checked that $D_2 = 1$ and $D_3 = 2$. They correspond the permutations 21 for $n = 2$ and 231, 312 for $n = 3$. We will try to find a relationship of D_n with D_i, for $1 \leq i \leq n-1$.

Let us have a close look at the required permutations. We note that n should not be placed at the n^{th} position. So, n has to appear some where between 1 and $n-1$. That is, for some i, $1 \leq i \leq n-1$

(a) n appears at the i^{th} position and i appears at the n^{th} position, or

(b) n appears at the i^{th} position and i does not appear at the n^{th} position.

Case (a): Fix $i, 1 \leq i \leq n-1$. Then the position of n and i is fixed and the remaining numbers j, for $j \neq i, n$ should not appear at the j^{th} place. As $j \neq i, n$, the problem reduces to the number of derangements of $n-2$ numbers which by our notation equals D_{n-2}. As i can be any one of the integers $1, 2, \ldots, n-1$, the number of derangements corresponding to the first case equals $(n-1)D_{n-2}$.

Case (b): Fix $i, 1 \leq i \leq n-1$. Then the position of n is at the i^{th} place but i is not placed at the n^{th} position. So, in this case, the problem reduces to placing the numbers $1, 2, \ldots, n-1$ at the places $1, 2, \ldots, i-1, i+1, \ldots, n$ such that the number i is not to be placed at the n^{th} position and for $j \neq i$, j is not placed at the j^{th} position. Let us rename the positions as $a_1, a_2, \ldots, a_{n-1}$, where $a_i = n$ and $a_j = j$ for $j \neq i$.

Then, with this renaming, the problem reduces to placing the numbers $1, 2, \ldots, n-1$ at places $a_1, a_2, \ldots, a_{n-1}$ such that the number j, for $1 \leq j \leq n-1$, is not placed at a_j'th position. This corresponds to the derangement of $n-1$ numbers that this by our notation equals D_{n-1}. Thus, in this case the number of derangements equals $(n-1)D_{n-1}$.

Hence, $D_n = (n-1)D_{n-1} + (n-1)D_{n-2}$. Or equivalently, $D_2 = 1$ and $D_1 = 0$ imply $D_n - nD_{n-1} = -\big(D_{n-1} - (n-1)D_{n-2}\big) = (-1)^2\big(D_{n-2} - (n-2)D_{n-3}\big) = \cdots = (-1)^n$. Therefore,

$$\begin{aligned} D_n &= nD_{n-1} + (-1)^n = n\big((n-1)D_{n-2} + (-1)^{n-1}\big) + (-1)^n \\ &= n(n-1)D_{n-2} + n(-1)^{n-1} + (-1)^n \\ &\vdots \\ &= n(n-1)\cdots 4 \cdot 3\, D_2 + n(n-1)\cdots 4(-1)^3 + \cdots + n(-1)^{n-1} + (-1)^n \\ &= n!\left(1 + \frac{-1}{1!} + \frac{(-1)^2}{2!} + \cdots + \frac{(-1)^{n-1}}{(n-1)!} + \frac{(-1)^n}{n!}\right). \end{aligned}$$

Or, in other words $\lim\limits_{n \longrightarrow \infty} \dfrac{D_n}{n!} = \dfrac{1}{e}$.

We end this chapter with another set of exercises.

2.3.1 Miscellaneous Exercises

1. Prove that there exists a bijection between any two of the following sets.

   (a) The set of words of length n on an alphabet consisting of m letters.

   (b) The set of maps of an n-set into a m-set.

   (c) The set of distributions of n distinct objects into m distinct boxes.

   (d) The set of n-tuples on m letters.

2. Prove that there exists a bijection between any two of the following sets.

   (a) The set of n letter words with distinct letters out of an alphabet consisting of m letters.

   (b) The set of one-one functions from an n-set into a m-set.

   (c) The set of distributions of n distinct objects into m distinct boxes, subject to "if an object is put in a box, no other object can be put in the same box".

   (d) The set of n-tuples on m letters, without repetition.

   (e) The set of permutations of m symbols taken n at a time.

3. Prove that there exists a bijection between any two of the following sets.

   (a) The set of increasing words of length n on m ordered letters.

   (b) The set of distributions on n non-distinct objects into m distinct boxes.

   (c) The set of combinations of m symbols taken n at a time with repetitions permitted.

4. Determine the number of ways of distributing n distinct objects into m distinct boxes so that objects in each box are arranged in a definite order.

5. Let X be a finite set and let $\mathbb{W}$ be the set of non-negative integers. A function $f : X \longrightarrow \mathbb{W}$ is said to give rise to a MULTISET if $f(x)$ is finite, for all $x \in X$. The number $f(x)$ is called the SIZE of $x \in X$ and $\sum_{x \in X} f(x)$ is called the size of the multiset. For example, if $X = \{a, b, c\}$ and $f(a) = 2, f(b) = 3$ and $f(c) = 1$ then the corresponding multiset is $\{a, a, b, b, b, c\}$ and the size of this multiset is 6.

   If $|X| = n$, determine the number of multiset of size k. *[Hint: Let $X = \{a_1, a_2, \ldots, a_n\}$ and let f_i denote the image of a_i, $1 \leq i \leq n$. Then, is the problem equivalent to finding the number of solutions in non-negative integers to the equation $f_1 + f_2 + \cdots + f_n = k$?]*

6. Fix a positive integer n. A COMPOSITION of n is an expression of n as a sum of positive integers. For example, if $n = 4$, then the distinct compositions are

$$4,\ \ 3+1,\ \ 1+3,\ \ 2+2,\ \ 2+1+1,\ \ 1+1+2,\ \ 1+2+1,\ \ 1+1+1+1.$$

Let $S_k(n)$ denote the number of compositions of n into k parts. Then $S_1(4) = 1,\ S_2(4) = 3,\ S_3(4) = 3$ and $S_4(4) = 1$. Determine the numbers $S_k(n)$, for $1 \leq k \leq n$. Also, determine $\sum_{k\geq 1} S_k(n)$. *[Hint: Is the problem equivalent to finding the number of solutions in positive integers to the equation $f_1 + f_2 + \cdots + f_k = n$?]*

7. Let $S = \{1, 2, \ldots, n\}$. Let $s(n, k)$ denote the number of permutations of S that have k disjoint cycles. For example, $s(4, 2) = 11$ and it corresponds to the permutations $(12)(34), (13)(24), (14)(23), (1)(234), (1)(243), (134)(2), (143)(2), (124)(3), (142)(3), (123)(4)$ and $(132)(4)$. By convention, we take $s(0, 0) = 1$ and $s(n, 0) = 0 = s(0, n)$, whenever $n \geq 1$.

   (a) Determine a recurrence relation for the numbers $s(n, k)$. *[Hint: Either the number n appears as a sinlge/separate cycle or it appears with some other number(s). In the first case, we are looking at $s(n-1, k-1)$ and in the second case, we are looking at $(n-1)s(n-1, k)$, as the number n can be placed at any of the $n-1$ positions to form a permutation. That is, $s(n, k) = (n-1)s(n-1, k) + s(n-1, k-1)$.]*

   (b) Recall that $x^{(n)} = x(x+1)(x+2)\cdots(x+n-1)$. Then prove that if $x^{(n)} = \sum_{k\geq 0} a(n, k)x^k$ then the numbers $a(n, k)$'s satisfy the same recurrence relation as $s(n, k)$'s with the same initial conditions.

   Hence, they give rise to the same sequence of numbers and are commonly known as the Stirling numbers of first kind.

Notes: Most of the ideas for this chapter have come from the books [2], [4] and [5].

Chapter 3

Advanced Counting

3.1 Pigeonhole Principle

The pigeonhole principle states that *if there are $n+1$ pigeons and n holes (boxes), then there is at least one hole (box) that contains two or more pigeons.* It can be easily verified that the pigeonhole principle is equivalent to the following statements:

1. If m pigeons are put into m pigeonholes, there is an empty hole if and only if there's a hole with more than one pigeon.

2. If n pigeons are put into m pigeonholes, with $n > m$, then there is a hole with more than one pigeon.

3. For two finite sets A and B, there exists a one to one and onto function $f : A \longrightarrow B$ if and only if $|A| = |B|$.

Remark 3.1.1. *Recall that the expression $\lceil x \rceil$, called the* CEILING FUNCTION*, is the smallest integer ℓ, such that $\ell \geq x$ and the expression $\lfloor x \rfloor$, called the* FLOOR FUNCTION*, is the largest integer k, such that $k \leq x$.*

1. **[Generalized Pigeonhole Principle]** if there are n pigeons and m holes with $n > m$, then there is at least one hole that contains $\lceil \frac{n}{m} \rceil$ pigeons.

2. *Dirichlet was the one who popularized this principle.*

Example 3.1.2. *1. Let a be an irrational number. Then prove that there exist infinitely many rational numbers $s = \frac{p}{q}$, such that $|a - s| < \frac{1}{q^2}$.*

Proof. Let $N \in \mathbb{N}$. Without loss of generality, we assume that $a > 0$. By $\{\alpha\}$, we will denote the fractional part of α. That is, $\{\alpha\} = \alpha - \lfloor \alpha \rfloor$.

Now, consider the fractional parts $\{0\}, \{a\}, \{2a\}, \ldots, \{Na\}$ of the first $(N+1)$ multiples of a and the N subintervals $[0, \frac{1}{N}), [\frac{1}{N}, \frac{2}{N}), \ldots, [\frac{N-1}{N}, 1)$ of $[0, 1)$. Clearly $\{ka\}$, for k a positive integer, cannot be an integer as a is an irrational number. Thus, by the pigeonhole principle, two of the above fractional parts must fall into the same subinterval. That is, there exist integers u, v and w such that $u > v$ but

$$\{ua\} \in [\frac{w}{N}, \frac{w+1}{N}) \quad \text{and} \quad \{va\} \in [\frac{w}{N}, \frac{w+1}{N}).$$

Thus, $|\{ua\} - \{va\}| < \dfrac{1}{N}$ and $|\{ua\} - \{va\}| = |(u-v)a - (\lfloor ua \rfloor - \lfloor va \rfloor)|$. Now, let $q = u - v$ and $p = \lfloor ua \rfloor - \lfloor va \rfloor$. Then $p, q \in \mathbb{Z}, q \neq 0$ and $|qa - p| < \frac{1}{N}$. Dividing by q, we get

$$|a - \frac{p}{q}| < \frac{1}{Nq} \leq \frac{1}{q^2} \quad \text{as } 0 < q \leq N.$$

Therefore, we have found a rational number $\dfrac{p}{q}$ such that $|a - \dfrac{p}{q}| < \dfrac{1}{q^2}$. We will now show that the number of such pairs (p, q) is infinite.

On the contrary, assume that there are only a finite number of rational numbers, say $r_1, r_2, \ldots, r_M$ such that

$$r_i = \frac{p_i}{q_i}, \quad \text{for } i = 1, \ldots, M, \quad \text{and} \quad |a - r_i| < \frac{1}{q_i^2}.$$

Since a is an irrational number, none of the differences $|a - r_i|$, for $i = 1, 2, \ldots, M$, will be exactly 0. Therefore, there exists an integer Q such that

$$|a - r_i| > \frac{1}{Q}, \quad \text{for all} \quad i = 1, 2, \ldots, M.$$

We now, apply our earlier argument to this Q. The argument gives the existence of a fraction $r = \dfrac{p}{q}$ such that $|a - r| < \frac{1}{Q} < |a - r_i|$, for $1 \leq iM$. Hence $r \neq r_i$ for all $i = 1, 2, \ldots, M$. On the other hand, we also have, $|a - r| < \dfrac{1}{q^2}$ contradicting the assumption that the fractions r_i, for $i = 1, 2, \ldots, M$, were all the fractions with this property. ■

2. *Let $\{a_1, a_2, \ldots, a_{mn+1}\}$ be a sequence of distinct $mn + 1$ real numbers. Then prove that this sequence has a subsequence of either $(m + 1)$ numbers that is strictly increasing or $(n + 1)$ numbers that is strictly decreasing.*

Observation: *The statement is NOT TRUE if there are exactly mn numbers. For example, consider the sequence $4, 3, 2, 1, 8, 7, 6, 5, 12, 11, 10, 9$ of $12 = 3 \times 4$ distinct numbers. This sequence neither has an increasing subsequence of 4 numbers nor a decreasing subsequence of 5 numbers.*

Proof. Let T be the given sequence. That is, $T = \{a_k\}_{k=1}^{mn+1}$ and define

$$\ell_i = \max_s \{s : \text{ an increasing subsequence of length } s \text{ exists starting with } a_i\}.$$

Then there are $mn+1$ positive integers $\ell_1, \ell_2, \ldots, \ell_{mn+1}$. If there exists a $j, 1 \leq j \leq mn+1$, such that $\ell_j \geq m+1$, then by definition of ℓ_j, there exists an increasing sequence of length $m+1$ starting with a_j and thus the result follows. So, on the contrary assume that $\ell_i \leq m$, for $1 \leq i \leq mn+1$.

That is, we have $mn+1$ numbers $(\ell_1, \ldots, \ell_{mn+1})$ and all of them have to be put in the boxes numbered $1, 2, \ldots, m$. So, by the generalized pigeonhole principle, there are at least $\left\lceil \frac{mn+1}{m} \right\rceil = n+1$ numbers (ℓ_i's) that lies in the same box. Therefore, let us assume that there exist numbers $1 \leq i_1 < i_2 < \cdots < i_{n+1} \leq mn+1$, such that

$$\ell_{i_1} = \ell_{i_2} = \cdots = \ell_{i_{n+1}}. \tag{3.1}$$

That is, the length of the largest increasing subsequences of T starting with the numbers $a_{i_1}, a_{i_2}, \ldots, a_{i_{n+1}}$ are all equal. We now claim that $a_{i_1} > a_{i_2} > \cdots > a_{i_{n+1}}$.

We will show that $a_{i_1} > a_{i_2}$. A similar argument will give the other inequalities and hence the proof of the claim. On the contrary, let if possible $a_{i_1} < a_{i_2}$ (recall that a_i's are distinct) and consider a largest increasing subsequence $a_{i_2} = \alpha_1 < \alpha_2 < \cdots < \alpha_{i_2}$ of T, starting with a_{i_2}, that has length ℓ_{i_2}. This subsequence with the assumption that $a_{i_1} < a_{i_2}$ gives an increasing subsequence

$$a_{i_1} < a_{i_2} = \alpha_1 < \alpha_2 < \cdots < \alpha_{\ell_{i_2}}$$

of T, starting with a_{i_1}, of length $\ell_{i_2}+1$. So, by definition of ℓ_i's, $\ell_{i_1} \geq \ell_{i_2}+1$. This gives a contradiction to the equality, $\ell_{i_1} = \ell_{i_2}$, in Equation (3.1). Hence the proof of the example is complete. ■

3. *Prove that there exist two powers of* 3 *whose difference is divisible by* 2011.

Proof. Consider the set $S = \{1 = 3^0, 3, 3^2, 3^3, \ldots, 3^{2011}\}$. Then $|S| = 2012$. Also, we know that when we divide positive integers by 2011 then the possible remainders are $0, 1, 2, \ldots, 2010$ (corresponding to exactly 2011 boxes). So, if we divide the numbers in S with 2011, then by pigeonhole principle there will exist at least two numbers $0 \leq i < j \leq 2011$, such that the remainders of 3^j and 3^i, when divided by 2011, are equal. That is, 2011 divides $3^j - 3^i$. Hence, this completes the proof.

Observe that this argument also implies that "there exists a positive integer ℓ such that 2011 *divides* $3^\ell - 1$ *as* $\gcd(3, 2011) = 1$*".* ■

4. *Prove that there exists a power of three that ends with* 0001.

Proof. Consider the set $S = \{1 = 3^0, 3, 3^2, 3^3, \ldots\}$. Now, let us divide each element of S by 10^4. As $|S| > 10^4$, there exist $i > j$ such that the remainders of 3^i and 3^j, when

are divided by 10^4, are equal. But $\gcd(10^4, 3) = 1$ and thus, 10^4 divides $3^\ell - 1$. That is, $3^\ell - 1 = s \cdot 10^4$ for some positive integer s. That is $3^\ell = s \cdot 10^4 + 1$ and hence the result follows. ■

Exercise 3.1.3. *1. At a party of n people, some pair of people are friends with the same number of people at the party. We assume that each person is friendly to at least one person at the party.*

2. *Let n be an odd integer. Then prove for any permutation σ of the set $\{1, 2, ..., n\}$ the product $P(\sigma) = (1 - \sigma(1))(2 - \sigma(2))...(n - \sigma(n))$ is necessarily even.*

3. *Prove that among any five points selected inside an equilateral triangle with side equal to 1 unit, there always exists a pair at the distance not greater than .5 units.*

4. *Let S be a set consisting of five lattice points. Prove that there exist two points in S, say P and Q, such that the mid-point of P and Q is also a lattice point?*

5. *Suppose $f(x)$ is a polynomial with integral coefficients. If $f(x) = 14$ for three distinct integers, say a, b and c, then prove that for no integer $f(x)$ can be equal to 15.*

6. *Suppose $f(x)$ is a polynomial with integral coefficients. If $f(x) = 11$ for five distinct integers, say $a_1, a_2, \ldots, a_5$ then prove that for no integer $f(x)$ can be equal to 9.*

7. *Let n be an odd positive number. Then prove that there exists a positive integer ℓ such that n divides $2^\ell - 1$.*

8. *Does there exist a multiple of 2013 that has all its digits 2? Explain your answer.*

9. *Does there exist a multiple of 2013 that ends with 23? Explain your answer.*

10. *Does there exist a multiple of 2013 that starts with 23? Explain your answer.*

11. *Consider a chess board with two of the diagonally opposite corners removed. Is it possible to cover the board with pieces of rectangular dominos whose size is exactly two board squares?*

12. *Let $x_1, x_2, \ldots, x_{2012}$ be a sequence of 2012 integers. Prove that there exist $1 \leq i < j \leq n$ such that $x_i + x_{i+1} + \cdots + x_{j-1} + x_j$ is a multiple of 2012.*

13. *Let $x_1, x_2, \ldots, x_{1008}$ be a sequence of 1008 integers. Prove that there exist $1 \leq i < j \leq n$ such that $x_j + x_i$ or $x_j - x_i$ is a multiple of 2013.*

14. *During the year 2000, a book store, which was open 7 days a week, sold at least one book each day, and a total of 600 books over the entire year. Show that there must have been a period of consecutive days when exactly 125 books were sold.*

15. *Let $S = \{1, 2, \ldots, 10\}$ and T be any subset of S consisting of 6 elements. Then prove that T has at least two elements whose sum is odd.*

16. *Suppose you are given a set A of ten different integers from the set $T = \{1, 2, ..., 116\}$. Prove that you can always find two disjoint non-empty subsets, S and T of A, such that the sum of elements in S equals the sum of elements in T.*

17. *Does there exist a number of the form $777\cdots7$ which is a multiple of 2007.*

18. *Show that if we select a subset of $n+1$ numbers from the set $\{1, 2, \ldots, 2n\}$ then some pair of numbers in the subset are relatively prime.*

19. *Show that the pigeonhole principle is the same as saying that at least one of the numbers $a_1, a_2, \ldots, a_n$ is greater than or equal to their average $\dfrac{a_1 + a_2 + \cdots + a_n}{n}$.*

20. *Consider two discs A and B, each having $2n$ equal sectors. Suppose each sector is painted either yellow or green. On disc A exactly n sectors are colored yellow and exactly n are colored green. For disc B there are no constrains. Show that there is a way of putting the two discs, one above the other, so that at least n corresponding regions have the same colors.*

21. *Prove that however one selects 55 integers $1 \leq x_1 < x_2 < x_3 < \cdots < x_{55} \leq 100$, there will be some two that differ by 9, some two that differ by 10, a pair that differ by 12, and a pair that differ by 13. Surprisingly, there need not be a pair of numbers that differ by 11.*

22. *There are 7 distinct real numbers. Is it possible to select two of them, say x and y such that $0 < \dfrac{x-y}{1+xy} < \dfrac{1}{\sqrt{3}}$?*

23. *Given any sequence of n integers, positive or negative, not necessarily all different, prove that there exists a consecutive subsequence that has the property that the sum of the members of this subsequence is a multiple of n.*

24. *Color the plane with two colors, say yellow and green. Then prove the following:*

    (a) *there exist two points at a distance of 1 unit which have been colored with the same color.*

    (b) *there is an equilateral triangle all of whose vertices have the same color.*

    (c) *there is a rectangle all of whose vertices have the same color.*

25. *Show that in any group of six people there are either three mutual friends or three mutual strangers.*

3.2 Principle of Inclusion and Exclusion

The following result is well known and hence we omit the proof.

Theorem 3.2.1. *Let U be a finite set. Suppose A and B are two subsets of U. Then the number of elements of U that are neither in A nor in B are*

$$|U| - \big(|A| + |B| - |A \cap B|\big).$$

Or equivalently, $|A \cup B| = |A| + |B| - |A \cap B|$.

A generalization of this to three subsets A, B and C is also well known. To get a result that generalizes Theorem 3.2.1 for n subsets $A_1, A_2, \ldots, A_n$, we need the following notations:

$$S_1 = \sum_{i=1}^{n} |A_i|, \quad S_2 = \sum_{1 \le i < j \le n} |A_i \cap A_j|, \quad S_3 = \sum_{1 \le i < j < k \le n} |A_i \cap A_j \cap A_k|, \cdots, S_n = |A_1 \cap A_2 \cap \cdots \cap A_n|.$$

With the notations as defined above, we have the following theorem, called the inclusion-exclusion principle. This theorem can be easily proven using the principle of mathematical induction. But we give a separate proof for better understanding.

Theorem 3.2.2. *[Inclusion-Exclusion Principle] Let $A_1, A_2, \ldots, A_n$ be n subsets of a finite set U. Then the number of elements of U that are in none of $A_1, A_2, \ldots, A_n$ is given by*

$$|U| - S_1 + S_2 - S_3 + \cdots + (-1)^n S_n. \tag{3.1}$$

Or equivalently,

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = S_1 - S_2 + \cdots + (-1)^{n-1} S_n. \tag{3.2}$$

Proof. We show that if an element $x \in U$ belongs to exactly k of the subsets $A_1, A_2, \ldots, A_n$, for some $k \ge 1$, then its contribution in (3.1) is zero. Suppose x belongs to exactly k subsets $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$. Then we observe the following:

1. The contribution of x in $|U|$ is 1.
2. The contribution of x in S_1 is k as $x \in A_{i_j},\ 1 \le j \le k$.
3. The contribution of x in S_2 is $\binom{k}{2}$ as $x \in A_{i_j} \cap A_{i_l},\ 1 \le j < l \le k$.
4. The contribution of x in S_3 is $\binom{k}{3}$ as $x \in A_{i_j} \cap A_{i_l} \cap A_{i_m},\ 1 \le j < l < m \le k$.

   Proceeding this way, we have

5. The contribution of x in S_k is $\binom{k}{k} = 1$, and
6. The contribution of x in S_ℓ for $\ell \ge k+1$ is 0.

So, the contribution of x in (3.1) is

$$1-k+\binom{k}{2}-\binom{k}{3}+\cdots+(-1)^{k-1}\binom{k}{k-1}+(-1)^k\binom{k}{k}+0\cdots+0=(1-1)^k=0.$$

This completes the proof of the theorem. The readers are advised to prove the equivalent condition. ■

Example 3.2.3. *1. Determine the number of* 10*-letter words on ENGLISH alphabet that does not contain all the vowels.*

Solution: *Let U be the set consisting of all the* 10*-letters words on ENGLISH alphabet and let A_α be a subset of U that does not contain the letter α. Then we need to compute*

$$|A_a\cup A_e\cup A_i\cup A_o\cup A_u|=S_1-S_2+S_3-S_4+S_5,$$

where $S_1=\sum\limits_{\alpha\in\{a,e,i,o,u\}}|A_\alpha|=\binom{5}{1}25^{10}$, $S_2=\binom{5}{2}24^{10}$, $S_3=\binom{5}{3}23^{10}$, $S_4=\binom{5}{4}22^{10}$ and $S_5=21^{10}$. So, the required answer is $\sum\limits_{k=1}^{5}(-1)^{k-1}\binom{5}{k}(26-k)^{10}$.

2. Determine the number of integers between 1 *and* 1000 *that are coprime to* $2,3,11$ *and* 13*.*

Solution: *Let $U=\{1,2,3,\ldots,1000\}$ and let $A_i=\{n\in U: i \text{ divides } n\}$, for $i=2,3,11,13$. Then note that we need the value of $|U|-|A_2\cup A_3\cup A_{11}\cup A_{13}|$. Observe that*

$$|A_2|=\lfloor\frac{1000}{2}\rfloor=500,\ |A_3|=\lfloor\frac{1000}{3}\rfloor=333,\ |A_{11}|=\lfloor\frac{1000}{11}\rfloor=90,\ |A_{13}|=\lfloor\frac{1000}{13}\rfloor=76,$$
$$|A_2\cap A_3|=\lfloor\frac{1000}{6}\rfloor=166,\ |A_2\cap A_{11}|=\lfloor\frac{1000}{22}\rfloor=45,\ |A_2\cap A_{13}|=\lfloor\frac{1000}{26}\rfloor=38,$$
$$|A_3\cap A_{11}|=\lfloor\frac{1000}{33}\rfloor=30,\ |A_3\cap A_{13}|=\lfloor\frac{1000}{39}\rfloor=25,\ |A_{11}\cap A_{13}|=\lfloor\frac{1000}{143}\rfloor=6,$$
$$|A_2\cap A_3\cap A_{11}|=15,\ |A_2\cap A_3\cap A_{13}|=12,\ |A_2\cap A_{11}\cap A_{13}|=3,$$
$$|A_3\cap A_{11}\cap A_{13}|=2,\ |A_2\cap A_3\cap A_{11}\cap A_{13}|=1.$$

Thus, the required number is

$$1000-\big((500+333+90+76)-(166+45+38+30+25+6)-(15+12+3+2)-1\big)=1000-720=280.$$

3. (Euler's ϕ-function Or Euler's totient function) Let n denote a positive integer.Then the Euler ϕ-function is defined by

$$\phi(n)=\big|\{k\ :1\le k\le n,\ \gcd(n,k)=1\}\big|. \tag{3.3}$$

Determine a formula for $\phi(n)$ in terms of its prime factors.

Solution: *Let $n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_k^{\alpha_k}$ be the unique decomposition of n as product of distinct*

primes $p_1, p_2, \ldots, p_k$, $U = \{1, 2, \ldots, n\}$ *and let* $A_{p_i} = \{m \in U : p_i \text{ divides } m\}$, *for* $1 \leq i \leq k$. *Then, by definition*

$$\begin{aligned}
\phi(n) &= |U| - S_1 + S_2 - S_3 + \cdots + (-1)^k S_k \\
&= n - \sum_{i=1}^{k} \frac{n}{p_i} + \sum_{1 \leq i < j \leq k} \frac{n}{p_i p_j} - \cdots + (-1)^k \frac{n}{p_1 p_2 \cdots p_k} \\
&= n \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right). \qquad (3.4)
\end{aligned}$$

Exercise 3.2.4. *1. Prove Theorem 3.2.2 using the principle of mathematical induction.*

2. *Recall the Derangement problem (see Page 48). On a rainy day, n students leave their umbrellas (which are indistinguishable) outside their examination room. Find the number of ways in which no student collects the correct umbrella when they finish the examination.*

3. *Determine the number of ways to put* 30 *indistinguishable red balls into* 4 *distinguishable boxes with at most* 10 *balls in each box.*

4. *Determine the number of ways to put* 30 *distinguishable balls into* 10 *distinguishable boxes such that at least* 1 *box is empty.*

5. *Determine the number of ways to put r distinguishable balls into n distinguishable boxes such that at least* 1 *box is empty.*

6. *Determine the number of onto functions* $f : M \longrightarrow N$, *where* $|M| = m$, $|N| = n$ *and* $n \leq m$ *(Recall (2.1) for another expression).*

7. *Determine the number of ways to put r distinguishable balls into n distinguishable boxes so that no box is empty.*

8. *Determine the number of ways to distribute* 40 *distinguishable books to* 25 *boys so that each boy gets at least one book.*

9. *Determine the number of ways to arrange* 10 *integers, say* $1, 2, 3, \ldots, 10$, *so that the number i is never followed immediately by* $i + 1$.

10. *Determine the number of strings of length* 15 *consisting of the* 10 *digits,* $0, 1, \ldots, 9$, *so that no string contains all the* 10 *digits.*

11. *In how many ways can n pairs of socks be hung on a line so that adjacent socks are from different pairs, if socks within a pair are indistinguishable and each pair is different.*

12. *Suppose* 15 *people get on a lift that stops at* 5 *floors, say* a, b, c, d *and* e. *Determine the number of ways for people to get out of the lift if at least one person gets out at each floor.*

13. *Determine the number of ways of permuting the* 26 *letters of the ENGLISH alphabet so that none of the patterns lazy, run, show and pet occurs.*

14. *Let* x *be a positive integer less than or equal to* 9999999.

    (a) *Find the number of* x*'s for which the sum of the digits in* x *equals* 30.

    (b) *How many of the solutions obtained in the first part consist of* 7 *digits.*

Chapter 4

Polya Theory

4.1 Groups

Our aim in this chapter is to look at groups and use it to the study of questions of the type:

1. How many different necklace configurations are possible if we use 6 beads of 3 different colors? Or for that matter what if we use n beads of m different colors?

2. How many different necklace configurations are possible if we use 12 beads among which 3 are *red*, 5 are *blue* and 4 are *green*? And a generalization of this problem.

3. Counting the number of chemical compounds which can be derived by the substitution of a given set of radicals in a given molecular structure.

It can be easily observed that if we want to look at different color configurations of a necklace formed using 6 beads, we need to understand the symmetries of a hexagon. Such a study is achieved through what in literature is called *groups*. Once we have learnt a bit about groups, we study *group action*. This study helps us in defining an equivalence relation on the set of color configurations for a given necklace. And it turns out that the number of distinct color configurations is same as the number of equivalence classes.

Before coming to the definition and its properties, let us look at the properties of the sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$. We know that the sets $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ satisfy the following:

Binary Operation: for every $a, b \in \mathbb{Z}\ (\mathbb{Q}, \mathbb{R}, \mathbb{C})$, $a + b$, called the addition of a and b, is an element of $\mathbb{Z}\ (\mathbb{Q}, \mathbb{R}, \mathbb{C})$;

Addition is Associative: for every $a, b, c \in \mathbb{Z}\ (\mathbb{Q}, \mathbb{R}, \mathbb{C})$, $(a + b) + c = a + (b + c)$;

Additive Identity: the element zero, denoted $\mathbf{0}$, is an element of $\mathbb{Z}\ (\mathbb{Q}, \mathbb{R}, \mathbb{C})$ and has the property that for every $a \in \mathbb{Z}\ (\mathbb{Q}, \mathbb{R}, \mathbb{C})$, $a + \mathbf{0} = a = \mathbf{0} + a$;

Additive Inverse: For every element $a \in \mathbb{Z}\ (\mathbb{Q}, \mathbb{R}, \mathbb{C})$, there exists an element $-a \in \mathbb{Z}\ (\mathbb{Q}, \mathbb{R}, \mathbb{C})$ such that $a + (-a) = \mathbf{0} = -a + a$;

Addition is Commutative: We also have $a + b = b + a$ for every $a, b \in \mathbb{Z}\ (\mathbb{Q}, \mathbb{R}, \mathbb{C})$.

Now, let us look at the sets $\mathbb{Z}^* = \mathbb{Z} - \{\mathbf{0}\}, \mathbb{Q}^* = \mathbb{Q} - \{\mathbf{0}\}, \mathbb{R}^* = \mathbb{R} - \{\mathbf{0}\}$ and $\mathbb{C}^* = \mathbb{C} - \{\mathbf{0}\}$. As in the previous case, we see that similar statements hold true for the sets $\mathbb{Z}^*, \mathbb{Q}^*, \mathbb{R}^*$ and $\mathbb{C}^*$. Namely,

Binary Operation: for every $a, b \in \mathbb{Z}^*\ (\mathbb{Q}^*, \mathbb{R}^*, \mathbb{C}^*)$, $a \cdot b$, called the multiplication of a and b, is an element of $\mathbb{Z}^*\ (\mathbb{Q}^*, \mathbb{R}^*, \mathbb{C}^*)$;

Multiplication is Associative: for every $a, b, c \in \mathbb{Z}^*\ (\mathbb{Q}^*, \mathbb{R}^*, \mathbb{C}^*)$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$;

Multiplicative Identity: the element one, denoted $\mathbf{1}$, is an element of $\mathbb{Z}^*\ (\mathbb{Q}^*, \mathbb{R}^*, \mathbb{C}^*)$ and for all $a \in \mathbb{Z}^*\ (\mathbb{Q}^*, \mathbb{R}^*, \mathbb{C}^*)$, $a \cdot \mathbf{1} = a = \mathbf{1} \cdot a$;

Multiplication is Commutative: One also has $a \cdot b = b \cdot a$ for every $a, b \in \mathbb{Z}^*\ (\mathbb{Q}^*, \mathbb{R}^*, \mathbb{C}^*)$.

Observe that if we choose $a \in \mathbb{Z}^*$ with $a \neq 1, -1$ then there does not exist an element $b \in \mathbb{Z}^*$ such that $a \cdot b = 1 = b \cdot a$. Whereas, for the sets $\mathbb{Q}^*, \mathbb{R}^*$ and $\mathbb{C}^*$ one can always find a b such that $a \cdot b = 1 = b \cdot a$.

Based on the above examples, an abstract notion called *groups* is defined. Formally, one defines a group as follows.

Definition 4.1.1 (Group)**.** *A group G, usually denoted $(G, *)$, is a non-empty set together with a binary operation, say $*$, such that the elements of G satisfy the following:*

1. *for every $a, b, c \in G$, $(a * b) * c = a * (b * c)$ (*Associativity Property*) holds in G;*
2. *there is an element $\mathbf{e} \in G$ such that $a * \mathbf{e} = a = \mathbf{e} * a$, for all $a \in G$ (*Existence of Identity*);*
3. *for every element $a \in G$, there exists an element $b \in G$ such that $a * b = \mathbf{e} = b * a$ (*Existence of Inverse*).*

In addition, if the set G satisfies $a * b = b * a$, for every $a, b \in G$, then G is said to be **an abelian (commutative)** *group.*

Before proceeding with examples of groups that concerns us, we state a few basic results in group theory without proof. The readers are advised to prove it for themselves.

Remark 4.1.2. *Let $(G, *)$ be a group. Then the following hold:*

1. *The identity element of G is unique. Hence, the identity element is denoted by e.*
2. *For each fixed $a \in G$, the element $b \in G$ such that $a * b = e = b * a$ is also unique. Therefore, for each $a \in G$, the element b that satisfies $a * b = e = b * a$ is denoted by a^{-1}.*

3. *Also, for each* $a \in G$, $(a^{-1})^{-1} = a$.

4. *If* $a*b = a*c$, *for some* $a, b, c \in G$ *then* $b = c$. *Similarly, if* $b*d = c*d$, *for some* $b, c, d \in G$ *then* $b = c$. *That is, the cancelation laws in* G.

5. *For each* $a, b \in G$, $(ab)^{-1} = b^{-1}a^{-1}$.

6. *By convention, we assume* $a^0 = e$, *for all* $a \in G$.

7. *For each* $a \in G$, $(a^n)^{-1} = (a^{-1})^n$, *for all* $n \in \mathbb{Z}$.

In the remaining part of this chapter, the binary operation may not be explicitly mentioned as it will be clear from the context. Now, let us now look at a few examples that will be used later in this chapter.

Example 4.1.3. Symmetric group on n **letters:** *Let* N *denote the set* $\{1, 2, \ldots, n\}$. *A function* $f : N \longrightarrow N$ *is called a permutation on* n *elements if* f *is both one-to-one and onto. Let* $\mathcal{S}_n = \{f : N \longrightarrow N \mid f \text{ is one to one and onto}\}$. *That is,* $\mathcal{S}_n$ *is the set of all permutations of the set* $\{1, 2, \ldots, n\}$. *Then the following can be verified:*

1. *Suppose* $f, g \in \mathcal{S}_n$. *Then* $f : N \longrightarrow N$ *and* $g : N \longrightarrow N$ *are two one-to-one and onto functions. Therefore, one uses the composition of functions to define the composite function* $f \circ g : N \longrightarrow N$ *by* $(f \circ g)(x) = f\big(g(x)\big)$. *Then it can be easily verified that* $f \circ g$ *is also one-to-one and onto. Hence* $f \circ g \in \mathcal{S}_n$. *That is, "composition of function", denoted* $\circ$, *defines a binary operation in* $\mathcal{S}_n$.

2. *It is well known that the composition of functions is an associative operation and thus* $(f \circ g) \circ h = f \circ (g \circ h)$.

3. *The function* $\mathbf{e} : N \longrightarrow N$ *defined by* $\mathbf{e}(i) = i$, *for all* $i = 1, 2, \ldots, n$ *is the identity function. That is, check that* $f \circ \mathbf{e} = f = \mathbf{e} \circ f$, *for all* $f \in \mathcal{S}_n$.

4. *Now let* $f \in \mathcal{S}_n$. *As* $f : N \longrightarrow N$ *is a one-to-one and onto function,* $f^{-1} : N \longrightarrow N$ *defined by* $f^{-1}(i) = j$, *whenever* $f(j) = i$, *for all* $i = 1, 2, \ldots, n$, *is a well defined function and is also one-to-one and onto. That is, for each* $f \in \mathcal{S}_n$, $f^{-1} \in \mathcal{S}_n$ *and* $f \circ f^{-1} = \mathbf{e} = f^{-1} \circ f$.

Thus $(\mathcal{S}_n, \circ)$ *is a group. This group is called the* Symmetric/Permutation *group on* n *letters. If* $\sigma \in \mathcal{S}_n$ *then one represents this by writing* $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$. *This representation of an element of* $\mathcal{S}_n$ *is called a* TWO ROW NOTATION. *Observe that as* σ *is one-to-one and onto function from* N *to* N, *it can be checked that* $N = \{\sigma(1), \sigma(2), \ldots, \sigma(n)\}$. *Hence, there are* n *choices for* $\sigma(1)$, $n - 1$ *choices for* $\sigma(2)$ *(all elements of* N *except* $\sigma(1)$*) and so on. Thus, the total number of elements in* $\mathcal{S}_n$ *is* $n! = n(n-1)\cdots 2 \cdot 1$.

Before discussing other examples, let us try to understand the group $\mathcal{S}_n$. As seen above, any element $\sigma \in \mathcal{S}_n$ can be represented using a two-row notation. There is another notation for permutations that is often very useful. This notation is called the *cycle notation*. Let us try to understand this notation.

Definition 4.1.4. *Let $\sigma \in \mathcal{S}_n$ and let $S = \{i_1, i_2, \ldots, i_k\} \subseteq \{1, 2, \ldots, n\}$ be distinct. If σ satisfies*

$$\sigma(i_\ell) = i_{\ell+1}, \text{ for all } \ell = 1, 2, \ldots, k-1, \ \sigma(i_k) = i_1 \quad \text{and } \sigma(r) = r \text{ for } r \notin S$$

*then σ is called a k-*cycle *and is denoted by $\sigma = (i_1, i_2, \ldots i_k)$ or $(i_2, i_3, \ldots, i_k, i_1)$ and so on.*

Example 4.1.5.

1. *The permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 1 & 5 \end{pmatrix}$ in cycle notation can be written as (1234), (2341), (3412), or (4123) as $\sigma(1) = 2, \sigma(2) = 3, \sigma(3) = 4, \sigma(4) = 1$ and $\sigma(5) = 5$.*

2. *The permutation $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 4 & 6 & 5 \end{pmatrix}$ in cycle notation equals $(123)(65)$ as $\sigma(1) = 2, \sigma(2) = 3, \sigma(3) = 1$, $\sigma(4) = 4$, $\sigma(5) = 6$ and $\sigma(6) = 5$. That is, this element is formed with the help of two cycles (123) and (56).*

3. *Consider two permutations $\sigma = (143)(27)$ and $\tau = (1357)(246)$. Then, there composition, denoted $\sigma \circ \tau$, is obtained as follows:*
   $(\sigma \circ \tau)(1) = \sigma\big(\tau(1)\big) = \sigma(3) = 1$, $(\sigma \circ \tau)(2) = \sigma\big(\tau(2)\big) = \sigma(4) = 3$, $(\sigma \circ \tau)(3) = \sigma\big(\tau(3)\big) = \sigma(5) = 5, (\sigma \circ \tau)(4) = \sigma\big(\tau(4)\big) = \sigma(6) = 6$, $(\sigma \circ \tau)(5) = 2$, $(\sigma \circ \tau)(6) = 7$ *and* $(\sigma \circ \tau)(7) = 4$. *Hence*

$$\sigma \circ \tau = (143)(27)(1357)(246) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 3 & 5 & 6 & 2 & 7 & 4 \end{pmatrix} = (235)(467).$$

4. *Similarly, verify that $(1456)(152) = (16)(245)$.*

Definition 4.1.6. *Two cycles $\sigma = (i_1, i_2, \ldots, i_t)$ and $\tau = (j_1, j_2, \ldots, j_s)$ are said to be* disjoint *if*

$$\{i_1, i_2, \ldots, i_t\} \cap \{j_1, j_2, \ldots, j_s\} = \emptyset.$$

The proof of the following theorem can be obtained from any standard book on abstract algebra.

Theorem 4.1.7. *Let $\sigma \in \mathcal{S}_n$. Then σ can be written as a product of disjoint cycles.*

Remark 4.1.8. *Observe that the representation of a permutation as a product of disjoint cycles, none of which is the identity, is unique up to the order of the disjoint cycles. The representation of an element $\sigma \in \mathcal{S}_n$ as product of disjoint cycles is called the* cyclic decomposition *of σ.*

Example 4.1.9. *1. Symmetries of regular n-gons in plane.*

(a) Suppose a unit square is placed at the coordinates $(0,0,0),(1,0,0),(0,1,0)$ *and* $(1,1,0)$. *Our aim is to move the square in space such that the position of the vertices may change but they are still placed at the above mentioned coordinates. The question arises, what are the possible ways can this be done? It can be easily verified that the possible configurations are as follows (see Figure 4.1):*

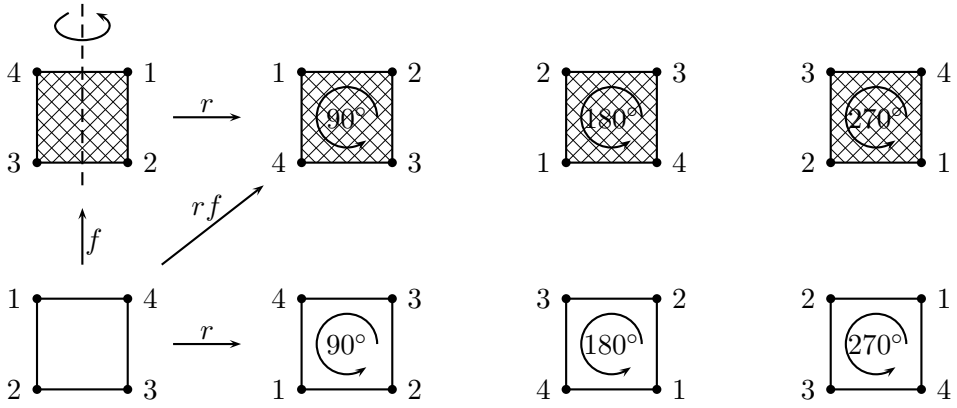


Figure 4.1: Symmetries of a square.

Let r *denote the counter-clockwise rotation of the square by* 90° *and* f *denote the flipping of the square along the vertical axis passing through the midpoint of opposite horizontal edges (see Figure 4.1). Then note that the possible configurations correspond to the set*

$$G = \{e, r, r^2, r^3, f, rf, r^2f, r^3f\} \text{ with relations } r^4 = e = f^2 \text{ and } fr^3 = rf. \quad (4.1)$$

Using (4.1), observe that $(rf)^2 = (rf)(rf) = r(fr)f = r(r^3f)f = r^4f^2 = e$. *Similarly, it can be checked that* $(r^2f)^2 = (r^3f)^2 = e$. *That is, all the terms* f, rf, r^2f *and* r^3f *are flips. The group* G *is generally denoted by* D_4 *and is called the* **Dihedral group** *or the* **symmetries of a square**. *This group can also be represented as follows:*

$$\{e, (1234), (13)(24), (1432), (14)(23), (24), (12)(34), (13)\}.$$

Exercise: *Relate the two representations of the group* D_4.

(b) In the same way, one can define the symmetries of an equilateral triangle (see Figure 4.2). This group is denoted by D_3 *and is represented as*

$$D_3 = \{e, r, r^2, f, rf, r^2f\} \text{ with relations } r^3 = e = f^2 \text{ and } fr^2 = rf, \quad (4.2)$$

where r *is a counter-clockwise rotation by* $120^\circ = \dfrac{2\pi}{3}$ *and* f *is a flip. Using Figure 4.2, one can check that the group* D_3 *can also be represented by*

$$D_3 = \{e, (ABC), (ACB), (BC), (CA), (AB)\}.$$

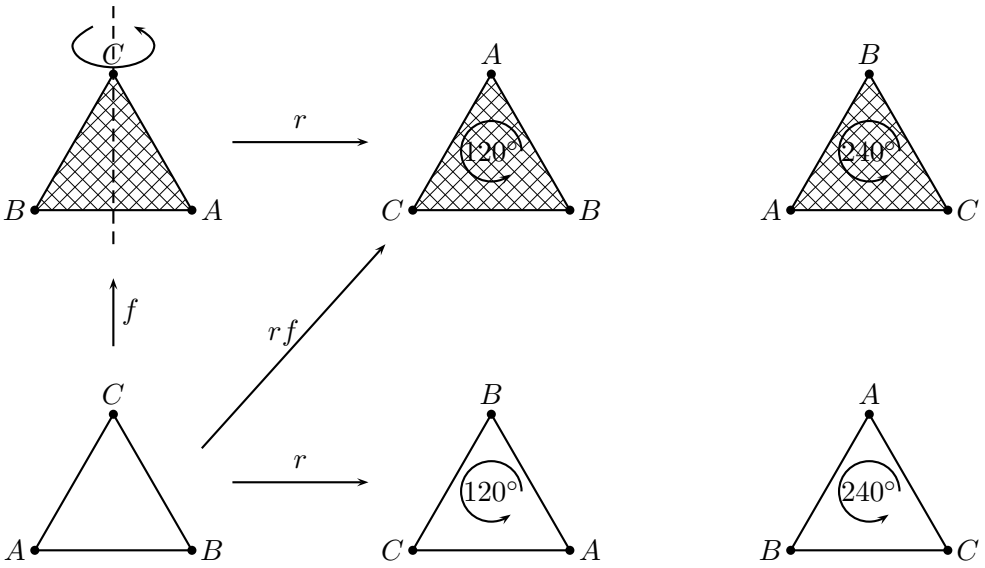


Figure 4.2: Symmetries of an Equilateral Triangle.

(c) *For a regular pentagon, it can be verified that the group of symmetries of a regular pentagon is given by* $G = \{e, r, r^2, r^3, r^4, f, rf, r^2f, r^3f, r^4f\}$ *with* $r^5 = 2 = f^2$ *and* $rf = fr^4$, *where* r *denotes a counter-clockwise rotation through an angle of* $72^\circ = \frac{2\pi}{5}$ *and* f *is a flip along a line that passes through a vertex and the midpoint of the opposite edge. Or equivalently, if we label the vertices of a regular pentagon, counter-clockwise, with the numbers* $1, 2, 3, 4$ *and* 5 *then*

$$\begin{aligned} G \;=\; & \{e, (1,2,3,4,5), (1,3,5,2,4), (1,4,2,5,3), (1,5,4,3,2), (2,5)(3,4), \\ & (1,3)(4,5), (1,5)(2,4), (1,2)(3,5), (1,4)(2,3)\}. \end{aligned}$$

(d) *In general, one can define symmetries of a regular* n*-gon. This group is denoted by* D_n,*has* $2n$ *elements and is represented as*

$$\{e, r, r^2, \ldots, r^{n-1}, f, rf, \ldots, r^{n-1}f\} \;\; \textit{with} \;\; r^n = e = f^2 \;\; \textit{and} \;\; fr^{n-1} = rf. \qquad (4.3)$$

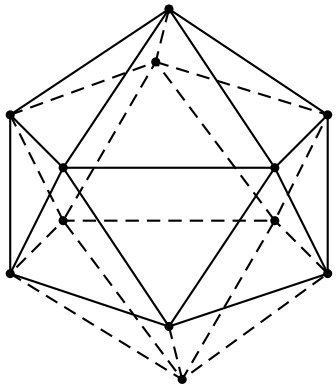
Here the symbol r *stands for a counter-clockwise rotation through an angle of* $\frac{2\pi}{n}$ *and* f *stands for a vertical flip.*

2. *Symmetries of regular platonic solids.*

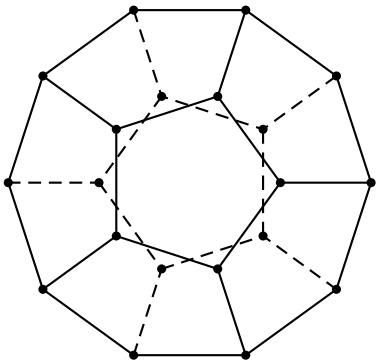
(a) *Recall from geometry that a tetrahedron is a* 3*-dimensional regular object that is composed of* 4*-equilateral triangles such that any three triangles meet at a vertex (see Figure 4.1). Observe that a tetrahedron has* 6 *edges,* 4 *vertices and* 4 *faces. If we denote the vertices of the tetrahedron with numbers* $1, 2, 3$ *and* 4*, then the symmetries of the tetrahedron can be represented with the help of the group,*

$$G = \{e, (234), (243), (124), (142), (123), (132), (134), (143), (12)(34), (13)(24), (14)(23)\},$$

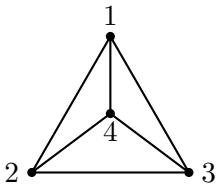
where, for distinct numbers i, j, k and ℓ, the element (ijk) is formed by a rotation of 120° *along the line that passes through the vertex ℓ and the centroid of the equilateral triangle with vertices i, j and k. Similarly, the group element $(ij)(k\ell)$ is formed by a rotation of* 180° *along the line that passes through mid-points of the edges (ij) and $(k\ell)$.*



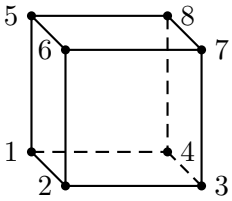
An Icosahedron



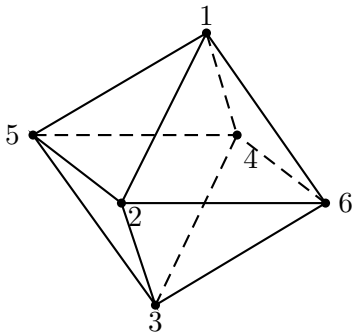
A Dodecahedron



A Tetrahedron



A Cube



An Octahedron

Figure 4.3: Regular Platonic solids.

(b) Consider the Cube and the Octahedron given in Figure 4.3. It can be checked that the group of symmetries of the two figures has 24 *elements. We give the group elements for the symmetries of the cube, when the vertices of the cube are labeled. The readers are required to compute the group elements for the symmetries of the octahedron. For the cube (see Figure 4.3), the group elements are*

i. e, the identity element;

ii. $3 \times 3 = 9$ *elements that are obtained by rotations along lines that pass through the center of opposite faces (*3 *pairs of opposite faces and each face is a square: corresponds to a rotation of* 90°*). In terms of the vertices of the cube, the group elements are*

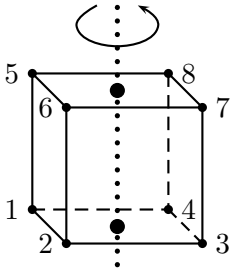
$$(1234)(5678), (13)(24)(57)(68), (1432)(5876), (1265)(3784), (16)(25)(38)(47),$$
$$(1562)(3487), (1485)(2376), (18)(45)(27)(36), (1584)(2673),$$

iii. $2 \times 4 = 8$ *elements that are obtained by rotations along lines that pass through opposite vertices (*4 *pairs of opposite vertices and each vertex is incident with* 3 *edges: correspond to a rotation of* $120°$*). The group elements in terms of the vertices of the cube are*

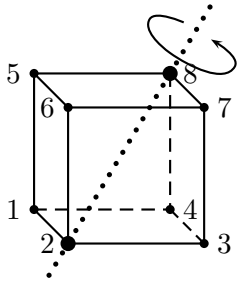
$$(254)(368), (245)(386), (163)(457), (136)(475), (275)(138),$$
$$(257)(183), (168)(274), (186)(247),$$

iv. $1 \times 6 = 6$ *elements that are obtained by rotations along lines that pass through the midpoint of opposite edges (*6 *pairs of opposite edges: corresponds to a rotation of* $180°$*). The corresponding elements in terms of the vertices of the cube are*

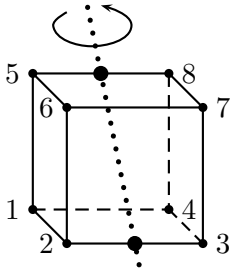
$$(14)(67)(28)(35), (23)(58)(17)(46), (15)(37)(28)(64), (26)(48)(17)(35),$$
$$(12)(78)(36)(45), (34)(56)(17)(28).$$



$90°$ rotations along lines that pass through opposite faces.



$120°$ rotations along lines that pass through opposite vertices.



$180°$ rotations along lines that pass through opposite edges.

Figure 4.4: Understanding the group of symmetries of a cube.

(c) Consider now the icosahedron and the dodecahedron (see Figure 4.3). Note that the icosahedron has 12 *vertices,* 20 *faces and* 30 *edges and the dodecahedron has* 20 *vertices,* 12 *faces and* 30 *edges. It can be checked that the group of symmetries of the two figures has* 60 *elements. We give the idea of the group elements for the symmetries of the icosahedron. The readers are required to compute the group elements for the symmetries of the dodecahedron. For the icosahedron, one has*

i. e, the identity element;

ii. $2 \times 10 = 20$ *elements that are obtained by rotations along lines that pass through the center of opposite faces (*10 *pairs of opposite faces and each face is an equilateral triangle: corresponds to a rotation of* $120°$*);*

iii. $6 \times 4 = 24$ *elements that are obtained by rotations along lines that pass through opposite vertices (*6 *pairs of opposite vertices and each vertex is incident with* 5 *edges: corresponds to a rotation of* $72°$*);*

iv. $1 \times 15 = 15$ *elements that are obtained by rotations along lines that pass through the midpoint of opposite edges (*15 *pairs of opposite edges: corresponds to a rotation of* 180°*).*

Exercise 4.1.10. *Determine the group of symmetries of a parallelogram, a rectangle, a rhombus and an octahedron?*

By now, we have already come across lots of examples of groups that arise as symmetries of different objects. To proceed further, we study the notion of subgroup of a given group. That is, if $(G, *)$ is a group and H is a non-empty subset of G then under what condition is $(H, *)$ a group in its own right (it is important to note that the binary operation is the same as that in G). Formally, we have the following definition.

Definition 4.1.11 (Subgroup)**.** *Let $(G, *)$ be a group. Then a non-empty subset H of G is said to be a subgroup of G, if H itself forms a group with respect to the binary operation $\star$.*

Example 4.1.12.

1. *Let G be a group with identity element e. Then G and $\{e\}$ are themselves groups and hence they are subgroups of G. These two subgroups are called* **trivial subgroups**.

2. *$\mathbb{Z}$, the set of integers, and $\mathbb{Q}$, the set of rational numbers, are subgroups of $(\mathbb{R}, +)$, the set of real numbers with respect to addition.*

3. *The set $\{e, r^2, f, r^2 f\}$ forms a subgroup of D_4.*

4. *Let $\sigma \in \mathcal{S}_4$. Then, using Theorem 4.1.7, we know that σ has a cycle representation. With this understanding it can be easily verified that the group D_4 is a subgroup of $\mathcal{S}_4$.*

5. *Consider $H = \{e, r, r^2, \ldots, r^{n-1}\}$ as a subset of D_n. Then it can be easily verified that H is a subgroup of D_4.*

Before proceeding further, let us look at the following two results which help us in proving "whether or not a given non-empty set H of a group G is a subgroup of not"?

Theorem 4.1.13 (Subgroup Test)**.** *Let G be a group and let H be a non-empty subset of G. Then H is a subgroup of G if for each $a, b \in H$, $ab^{-1} \in H$.*

Proof. As H is non-empty, we can find an $x \in H$. Therefore, for $a = x$ and $b = x$, the condition $ab^{-1} \in H$ implies that $e = aa^{-1} \in H$. Thus, H has the identity element of G. Hence, for each $h \in H \subset G$, $eh = h = he$.

We now need to prove that for each $h \in H$, $h^{-1} \in H$. To do so, note that for $a = e$ and $b = h$ the condition $ab^{-1} \in H$ reduces to $h^{-1} = eh^{-1} \in H$.

As a third step, we show that H is closed with respect to the binary operation of G. So, let us assume that $x, y \in H$. Then by the previous paragraph, $y^{-1} \in H$. Therefore, for $a = x$ and

$b = y^{-1}$ the condition $ab^{-1} \in H$ implies that $xy = x(y^{-1})^{-1} \in H$. Hence, H is also closed with respect to the binary operation of G.

Finally, we see that since the binary operation of H is same as that of G and since associativity holds G it holds in H as well. ■

We now give another result without proof that helps us in deciding whether a non-empty subset of a group is a subgroup or not.

Theorem 4.1.14. *[Two-Step Subgroup Test] Let H be a non-empty subset of a group G. Then H is a subgroup if the two conditions given below hold.*

1. *For each $a, b \in H$, $ab \in H$ (i.e., H is closed with respect to the binary operation of G).*
2. *For each $a \in H$, $a^{-1} \in H$.*

We now give a few examples to understand the above theorems.

Example 4.1.15. *1. Consider the group $(\mathbb{Z}, +)$. Then in the following cases, the given subsets do not for a subgroup.*

   (a) Let $H = \{0, 1, 2, 3, \ldots\} \subset \mathbb{Z}$. Note that, for each $a, b \in H$, $a + b \in H$ and the identity element $0 \in H$. But H is not a subgroup of $\mathbb{Z}$, as for all $n \neq 0$, $-n \notin H$.

   (b) Let $H = \mathbb{Z} \setminus \{0\} = \{\ldots, -3, -2, -1, 1, 2, 3, \ldots\} \subset \mathbb{Z}$. Note that, the identity element $0 \notin H$ and hence H is not a subgroup of $\mathbb{Z}$.

   (c) Let $H = \{-1, 0, 1\} \subset \mathbb{Z}$. Then H contains the identity element 0 of $\mathbb{Z}$ and for each $h \in H$, $h^{-1} = -h \in H$. But H is not a subgroup of $\mathbb{Z}$ as $1 + 1 = 2 \notin H$.

2. *Let G be an abelian group with identity e. Consider the sets $H = \{x \in G : x^2 = e\}$ and $K = \{x^2 : x \in G\}$. Then prove that both H and K are subgroups of G.*

   Proof. *Clearly $e \in H$ and $\mathbf{e} \in K$. Hence, both H and K are non-empty subsets of G. We first show that H is a subgroup of G.*

   As H is non-empty, pick $x, y \in H$. Thus, $x^2 = e = y^2$. We will now use Theorem 4.1.13, to show that $xy^{-1} \in H$. But this is equivalent to showing that $\left(xy^{-1}\right)^2 = e$. But this is clearly true as G is abelian implies that

$$\left(xy^{-1}\right)^2 = x^2(y^{-1})^2 = e(y^2)^{-1} = e^{-1} = e.$$

   Thus, H is indeed a subgroup of G by Theorem 4.1.13.

   Now, let us prove that K is a subgroup of G. We have already seen that K is non-empty. Thus, we just need to show that for each $x, y \in K$, $xy^{-1} \in K$.

   Note that $x, y \in K$ implies that there exists $a, b \in G$ such that $x = a^2$ and $y = b^2$. As $b \in G, b^{-1} \in G$. Also, $xy^{-1} = a^2(b^2)^{-1} = a^2(b^{-1})^2 = (ab^{-1})^2 \in K$ as G is abelian and $ab^{-1} \in G$. Thus, K is also a subgroup of G.

As a last result of this section, we prove that the condition of the above theorems can be weakened if we assume that H is a finite, non-empty subset of a group G.

Theorem 4.1.16. *[Finite Subgroup Test] Let G be a group and let H be a non-empty finite subset of G. If H is closed with respect to the binary operation of G then H is a subgroup of G.*

Proof. By Theorem 4.1.14, we just need to show that for each $a \in H$, $a^{-1} \in H$. If $a = e \in H$ then $a^{-1} = e^{-1} = e \in H$. So, let us assume that $a \neq e$ and $a \in H$. Now consider the set $S = \{a, a^2, a^3, \ldots, a^n, \ldots\}$. As H is closed with respect to the binary operation of G, $S \subset H$. But H has only finite number of elements. Hence, all these elements of S cannot be distinct. That is, there exist positive integers, say m, n with $m > n$, such that $a^m = a^n$. Thus, using Remark 4.1.2, one has $a^{m-n} = e$. Hence, $a^{-1} = a^{m-n-1}$ and by definition $a^{m-n-1} \in H$. ■

Exercise 4.1.17.

1. *Consider the group D_3. Does the subset $\{e, rf\}$ form a subgroup of D_3?*
2. *Determine all the subgroups of D_4.*
3. *Fix a positive integer n and consider the group D_n. Now, for each integer i, $0 \leq i \leq n-1$, does the set $\{e, r^i f\}$ form a subgroup of D_n? Justify your answer.*
4. *Determine all the subgroups of the group of symmetries of a tetrahedron.*
5. *Determine all the subgroups of the group of symmetries of a cube.*

4.2 Lagrange's Theorem

In this section, we prove the first fundamental theorem for groups that have finite number of elements. To do so, we start with the following example to motivate our definition and the ideas that they lead to.

Example 4.2.1. *Consider the set $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$. Then $\mathbb{R}^2$ is an abelian group with respect to component wise addition. That is, for each $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$, the binary operation is defined by $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$. Check that if H is a subgroup of $\mathbb{R}^2$ then H represents a line passing through $(0, 0)$.*

Hence, $H_1 = \{(x, y) \in \mathbb{R}^2 : y = 0\}$, $H_2 = \{(x, y) \in \mathbb{R}^2 : x = 0\}$ and $H_3 = \{(x, y) \in \mathbb{R}^2 : y = 3x\}$ are subgroups of $\mathbb{R}^2$. Note that H_1 represents the X-axis, H_2 represents the Y-axis and H_3 represents a line passes through the origin and has slope 3.

Fix the element $(2, 3) \in \mathbb{R}^2$. Then

1. *$(2, 3) + H_1 = \{(2, 3) + (x, y) : y = 0\} = \{(2 + x, 3) : x \in \mathbb{R}\}$. This is the equation of a line that passes through the point $(2, 3)$ and is parallel to the X-axis.*
2. *verify that $(2, 3) + H_2$ represents a line that passes through the point $(2, 3)$ and is parallel to the Y-axis.*

3. $(2,3) + H_3 = \{(2+x, 3+3x) : x \in \mathbb{R}\} = \{(x,y) \in \mathbb{R}^2 : y = 3x - 3\}$. *So, this represents a line that has slope* 3 *and passes through the point* $(2,3)$.

So, we see that if we fix a subgroup H *of* $\mathbb{R}^2$ *and take any point* $(x_0, y_0) \in \mathbb{R}^2$, *then the set* $(x_0, y_0) + H$ *gives a line that is a parallel shift of the line represented by* H *and* $(x_0, y_0) + H$ *contains the point* (x_0, y_0). *Hence, it can be easily observed that*

1. (x_1, y_1) *lies on the line* $(x_0, y_0) + H$ *if and only if* $(x_0, y_0) + H = (x_1, y_1) + H$.

2. *for any two* $(x_0, y_0), (x_1, y_1) \in \mathbb{R}^2$, *either* $(x_0, y_0) + H = (x_1, y_1) + H$ *or they represent two parallel lines which themselves are parallel to the line represented by* H.

3. $\bigcup_{x \in \mathbb{R}} \bigcup_{y \in \mathbb{R}} (x, y) + H = \mathbb{R}^2$.

That is, if we define a relation, denoted $\sim$, *in* $\mathbb{R}^2$ *by* $(x_1, y_1) \sim (x_2, y_2)$, *whenever* $(x_1 - x_2, y_1 - y_2) \in H$, *then the above observations imply that this relation is an equivalence relation. Hence, as* (x, y) *vary over all the points of* $\mathbb{R}^2$, *we get a partition of* $\mathbb{R}^2$. *Moreover, the equivalence class containing the point* (x_0, y_0) *is the set* $(x_0, y_0) + H$.

Therefore, we see that given a subgroup H of a group G, it may be possible to partition the group G into subsets that are in some sense similar to H itself. Example 4.2.1 also implies that for each $g \in G$, we need to consider the set $g + H$, if G is an additive group or either the set gH or the set Hg, if G is a multiplicative group. So, we are led to the following definition.

Definition 4.2.2 (Left and Right Coset)**.** *Let* G *be a group and let* H *be a subgroup of* G. *Then for each* $g \in G$ *the set*

1. $gH = \{gh : h \in H\}$ *is called the left coset of* H *in* G.

2. $Hg = \{hg : h \in H\}$ *is called the right coset of* H *in* G.

Remark 4.2.3. *Since the identity element* $e \in H$, *for each fixed* $g \in G$, $g = ge \in gH$. *Hence, we often say that* gH *is the left coset of* H *containing* g. *Similarly,* $g \in Hg$ *and hence* Hg *is said to be the right coset of* H *containing* g.

Example 4.2.4. *Consider the group* D_4 *and let* $H = \{e, f\}$ *and* $K = \{e, r^2\}$ *be two subgroups of* D_4. *Then observe the following:*

$$H = \{e, f\} = Hf, \quad Hr = \{r, fr\} = H\ fr,$$
$$H\ r^2 = \{r^2, fr^2\} = H\ fr^2 \quad \text{and} \quad H\ r^3 = \{r^3, fr^3\} = H\ fr^3. \tag{4.1}$$
$$H = \{e, f\} = fH, \quad rH = \{r, rf\} = rf\ H,$$
$$r^2\ H = \{r^2, r^2 f\} = r^2 f\ H \quad \text{and} \quad r^3\ H = \{r^3, r^3 f\} = r^3 f\ H. \tag{4.2}$$
$$K = \{e, r^2\} = Kr^2 = r^2 K, \quad Kr = \{r, r^3\} = rK = Kr^3 = r^3 K$$
$$Kf = \{f, r^2 f\} = fK = K\ r^2 f = r^2 f\ K \quad \text{and}$$
$$K\ fr = \{fr, fr^3\} = fr\ K = K\ fr^3 = fr^3\ K. \tag{4.3}$$

From (4.1) and (4.2), we note that in general $Hg \neq gH$, for each $g \in D_4$, whereas from (4.3), we see that $Kg = gK$, for each $g \in D_4$. So, there should be a way to distinguish between these two subgroups of D_4. This leads to to study of normal subgroups and beyond. The interested reader can look at any standard book in abstract algebra to go further in this direction.

Now, let us come back to the partition of a group using its subgroup. The proof of the theorem is left as an exercise for the readers.

Theorem 4.2.5. *Let H be a subgroup of a group G. Suppose $a, b \in G$. Then the following results hold for left cosets of H in G:*

1. *$aH = H$ if and only if $a \in H$,*
2. *aH is a subgroup of G if and only if $a \in H$,*
3. *either $aH = bH$ or $aH \cap bH = \emptyset$,*
4. *$aH = bH$ if and only if $a^{-1}b \in H$.*

Similarly one obtains the following results for right cosets of H in G.

1. *$Ha = H$ if and only if $a \in H$,*
2. *Ha is a subgroup of G if and only if $a \in H$,*
3. *either $Ha = Hb$ or $Ha \cap Hb = \emptyset$,*
4. *$Ha = Hb$ if and only if $ab^{-1} \in H$.*

Furthermore, $aH = Ha$ if and only if $H = aHa^{-1} = \{aha^{-1} : h \in H\}$.

To proceed further, we need the following definition.

Definition 4.2.6 (Order of a Group)**.** *The number of elements in G, denoted $|G|$, is called the order of G. If $|G| < \infty$, then G is called a group of finite order.*

We are now ready to prove the main result of this section, namely the Lagrange's Theorem.

Theorem 4.2.7. *Let H be a subgroup of a finite group G. Then $|H|$ divides $|G|$. Moreover, the number of distinct left (right) cosets of H in G equals $\dfrac{|G|}{|H|}$.*

Proof. We give the proof for left cosets. A similar proof holds for right cosets. Since G is a finite group, the number of left cosets of H in G is finite. Let $g_1H, g_2H, \ldots, g_mH$ be the collection of all left cosets of H in G. Then by Theorem 4.2.5, two cosets are either equal or they are disjoint. That is, G is a disjoint union of the sets $g_1H, g_2H, \ldots, g_mH$.

Also, it can be easily verified that $|aH| = |bH|$, for each $a, b \in G$. Hence, $|g_iH| = |H|$, for all $i = 1, 2, \ldots, m$. Thus, $|G| = \left|\bigcup_{i=1}^{m} g_iH\right| = \sum_{i=1}^{m} |g_iH| = m|H|$ (the disjoint union gives the second equality). Thus, $|H|$ divides $|G|$ and the number of left cosets equals $m = \dfrac{|G|}{|H|}$. ∎

Remark 4.2.8. *The number m in Theorem 4.2.7 is called* the index *of H in G, and is denoted by $[G : H]$ or $i_G(H)$.*

Theorem 4.2.7 is a statement about any subgroup of a finite group. It may so happen that the group G and its subgroup H may have infinite number of elements but the number of left (right) cosets of H in G may be finite. One still talks of index of H in G in such cases. For example, consider $H = 10\mathbb{Z}$ as a subgroup of the additive group $\mathbb{Z}$. Then $[\mathbb{Z} : H] = 10$. In general, for a fixed positive integer m, consider the subgroup $m\mathbb{Z}$ of the additive group $\mathbb{Z}$. Then it can be easily verified that $[\mathbb{Z} : m\mathbb{Z}] = m$.

Before coming to our next remark, we need the following definition and example.

Definition 4.2.9 (Order of an Element)**.** *Let G be a group and let $g \in G$. Then* **the smallest positive integer** *m such that $g^m = e$ is called the order of g. If there is no such positive integer then g is said to have* **infinite order**. *The order of an element is denoted by $\mathbf{O}(g)$.*

Example 4.2.10. *1. The only element of order 1 in a group G is the identity element of G.*

2. In D_4, the elements r^2, f, rf, r^2f, r^f have order 2, whereas the elements r and r^3 have order 4.

Exercise 4.2.11. *1. Prove that for each $a \in G$, $\mathbf{O}(a) = \mathbf{O}(a^{-1})$.*

2. Determine the order of each subgroup that were obtained in Exercise 4.1.17.

With the definition of the order of an element, we now prove that in general, the converse of Lagrange's Theorem is not true. To see this consider the group G discussed in Example 4.1.9.2a. This group has 12 elements and 6 divides 12. Whereas it can be shown that G doesn't have a subgroup of order 6. We give a proof for better understanding of cosets.

Proof. Let if possible, H be a subgroup of order 6 in G, where

$$G = \{e, (234), (243), (124), (142), (123), (132), (134), (143), (12)(34), (13)(24), (14)(23)\}.$$

Observe that G has exactly 8 elements of the form (ijk), for distinct numbers i, j and k, and each has order 3. Hence, G has exactly 8 elements of order 3. Let $a \in G$ with $\mathbf{O}(a) = 3$. Then using Theorem 4.2.5, we see that cosets of H in G will be exactly 2 and at the same time, the possible cosets could be H, aH and a^2H (as $a^3 = e$, no other coset exists). Hence, at most two of the cosets H, aH and a^2H are distinct. But, using Theorem 4.2.5, it can be easily verified that the equality of any two of them gives $a \in H$. Therefore, all the 8 elements of order 3 must be elements of H. That is, H must have at least 9 elements (8 elements of order 3 and one identity). This is absurd as $|H| = 6$. ∎

we now derive some important corollaries of Lagrange's Theorem. We omit the proof as it can be found in any standard textbook in abstract algebra. The first corollary is about the order of an element of a finite group. The observation that for each $g \in G$, the set $H = \{e, g, g^2, g^3, \ldots\}$ forms a subgroup of any finite group G gives the proof of the next result.

Corollary 4.2.12. *Let G be a finite group and let $g \in G$. Then $\mathbf{O}(g)$ divides $|G|$.*

Remark 4.2.13. *Corollary 4.2.12 implies that if G is a finite group of order n then the possible orders of its elements are the divisors of n. For example, if $|G| = 30$ then for each $g \in G$, $\mathbf{O}(g) \in \{1, 2, 3, 5, 6, 10, 15, 30\}$.*

Let G be a finite group. Then in the first corollary, we have shown that for any $g \in G$, $\mathbf{O}(g)$ divides $|G|$. Therefore, $|G| = m \cdot \mathbf{O}(g)$, for some positive integer m. Hence

$$g^{|G|} = g^{m \cdot \mathbf{O}(g)} = (g^{\mathbf{O}(g)})^m = e^m = e.$$

This observation gives our next result.

Corollary 4.2.14. *Let G be a finite group. Then, for each $g \in G, g^{|G|} = e$.*

Let P be an odd prime and consider the set $\mathbb{Z}_p^* = \{1, 2, \ldots, p-1\}$. Then, check that $\mathbb{Z}_p^*$ forms a group with respect to the binary operation

$$a \odot b = \text{the remainder, when } ab \text{ is divided by } p.$$

Applying Corollary 4.2.14 to $\mathbb{Z}_p^*$ gives the famous result called the *Fermat's Little Theorem*. To state this, recall that for $a, b \in \mathbb{Z}$, the notation "$a \equiv b \pmod{p}$" indicates that p divides $a - b$.

Corollary 4.2.15. *Let a be any positive integer and let p be a prime. Then $a^{p-1} \equiv 1 \pmod{p}$, if p does not divide a. In general, $a^p \equiv a \pmod{p}$.*

We now state without proof a generalization of the Fermat's Little Theorem, popularly known as the Euler's Theorem. to do so, let $U_n = \{k : 1 \leq k \leq n, \gcd(k, n) = 1\}$, for each positive integer n. Then U_n, with binary operation

$$a \odot b = \text{the remainder, when } ab \text{ is divided by } n$$

forms a group. Also, recall that the symbol $\varphi(n)$ gives the number of integers between 1 and n that are coprime to n. That is, $|U_n| = \varphi(n)$, for each positive integer n. Now applying Corollary 4.2.14 to U_n, gives the next result.

Corollary 4.2.16. *Let $a, n \in \mathbb{Z}$ with $n > 0$. If $\gcd(a, n) = 1$ then $a^{\varphi(n)} \equiv 1 \pmod{n}$.*

Example 4.2.17. *1. Find the unit place in the expansion of 13^{1001}.*

Solution : *Observe that $13 \equiv 3 \pmod{10}$. So, $13^{1001} \equiv 3^{1001} \pmod{10}$. Also, $3 \in U_{10}$ and therefore by Corollary 4.2.14, $3^{|U_{10}|} = 1 \pmod{10}$. But $|U_{10}| = 4$ and $1001 = 4 \cdot 250 + 1$. Thus,*

$$13^{1001} \equiv 3^{1001} \equiv 3^{4 \cdot 250 + 1} \equiv (3^4)^{250} \cdot 3^1 \equiv 1 \cdot 3 \equiv 3 \pmod{10}.$$

Hence, the unit place in the expansion of 13^{1001} is 3.

2. *Find the unit and tens place in the expansion of* 23^{1002}.

   Solution : *Observe that* $23 \in U_{100}$ *and* $23^{|U_{100}|} = 1 \pmod{100}$. *But* $|U_{100}| = 40$ *and* $1002 = 40 \cdot 25 + 2$. *Hence*

$$23^{1002} \equiv 23^{40 \cdot 25 + 2} \equiv (23^{40})^{25} \cdot 23^2 \equiv 1 \cdot 23^2 \equiv 529 \equiv 29 \pmod{100}.$$

   Hence, the unit place is 9 *and the tens place is* 2 *in the expansion of* 23^{1002}.

4.3 Group Action

Definition 4.3.1. *Let* $(G, \cdot)$ *be a group with identity* e. *Then* G *is said to act on a set* X, *if there exists an operator* $\star : G \times X \longrightarrow X$ *satisfying the following conditions:*

1. $e \star x = x$, *for all* $x \in X$, *and*
2. $g \star (h \star x) = (g \cdot h) \star x$, *for all* $x \in X$ *and* $g, h \in G$.

Remark 4.3.2.

1. *Let us assume that* X *consists of a set of points and let us suppose that the group* G *acts on* X *by moving the points. Then, Definition 4.3.1 can be interpreted as follows:*
   (a) *the first condition implies that the identity element of the group does not move any element of* X. *That is, the points in* X *remain fixed when they are acted upon by the identity element of* G.
   (b) *the second condition implies that if a point, say* $x_0 \in X$, *is first acted upon by an element* $h \in G$ *and then by an element* $g \in G$ *then the final position of* x_0 *is same as the position it would have reached if it was acted exactly once by the element* $g \star h \in G$.
2. *Fix an element* $g \in G$. *Then the set* $\{g \star x : x \in X\} = X$.
   For otherwise, there exist $x, y \in X$ *such that* $g \star x = g \star y$. *Then, by definition,*

$$x = e \star x = (g^{-1} \cdot g) \star x = g^{-1} \star (g \star x) = g^{-1}(g \star y) = (g^{-1} \cdot g) \star y = e \star y = y.$$

   That is, g *just permutes the elements of* X. *Or equivalently, each* $g \in G$ *gives rise to a one-one, onto function from* X *into itself.*
3. *There may exist* $g, h \in G$, *with* $g \neq h$ *such that* $g \star x = h \star x$, *for all* $x \in X$,.

Before proceeding further with definitions and results related with group action, let us look at a few examples.

Example 4.3.3. *1. Consider the dihedral group* $D_6 = \{e, r, \ldots, r^5, f, rf, \ldots, r^5 f\}$, *with* $r^6 = e = f^2$ *and* $rf = fr^5$. *Here,* f *stands for the vertical flip and* r *stands for counter clockwise rotation by an angle of* $\frac{\pi}{3}$. *Then* D_6 *acts on the labeled edges/vertices of a regular hexagon by permuting the labeling of the edges/vertices (see Figure 4.5).*

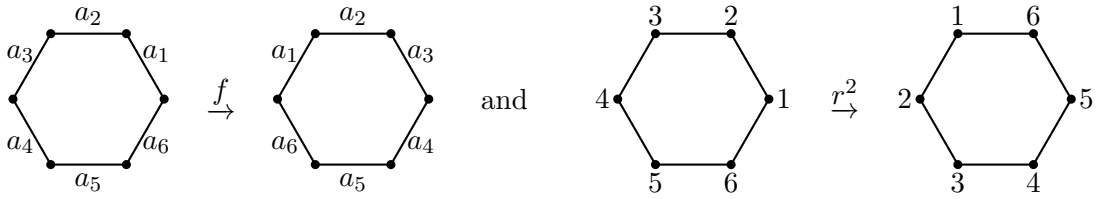


Figure 4.5: Action of f on labeled edges and of r^2 on labeled vertices of a regular hexagon.

2. *Let X denote the set of ways of coloring the vertices of a square with two colors, say, Red and Blue. Then X equals the set of all functions $h : \{1, 2, 3, 4\} \longrightarrow \{Red, Blue\}$, where the vertices south-west, south-east, north-east and north-west are respectively, labeled as $1, 2, 3$ and 4. Then, using Lemma 2.1.1, $|X| = 16$. The distinct colorings have been depicted in Figure 4.6, where R stands for the vertex colored "Red" and B stands for the vertex colored "Blue". For example, the figure labeled x_9 in Figure 4.6 corresponds to $h(1) = R = h(4)$ and $h(2) = B = h(3)$. Now, let us denote the permutation (1234) by r and the permutation $(12)(34)$ by f. Then the dihedral group $D_4 = \{e, r, r^2, r^3, f, rf, r^2f, r^3f\}$ acts on the set X. For example,*

   (a) *x_1 and x_{16} are mapped to itself under the action of every element of D_4. That is, $g \star x_1 = x_1$ and $g \star x_{16} = x_{16}$, for all $g \in G$.*

   (b) *$r \star x_2 = x_5$ and $f \star x_2 = x_3$.*

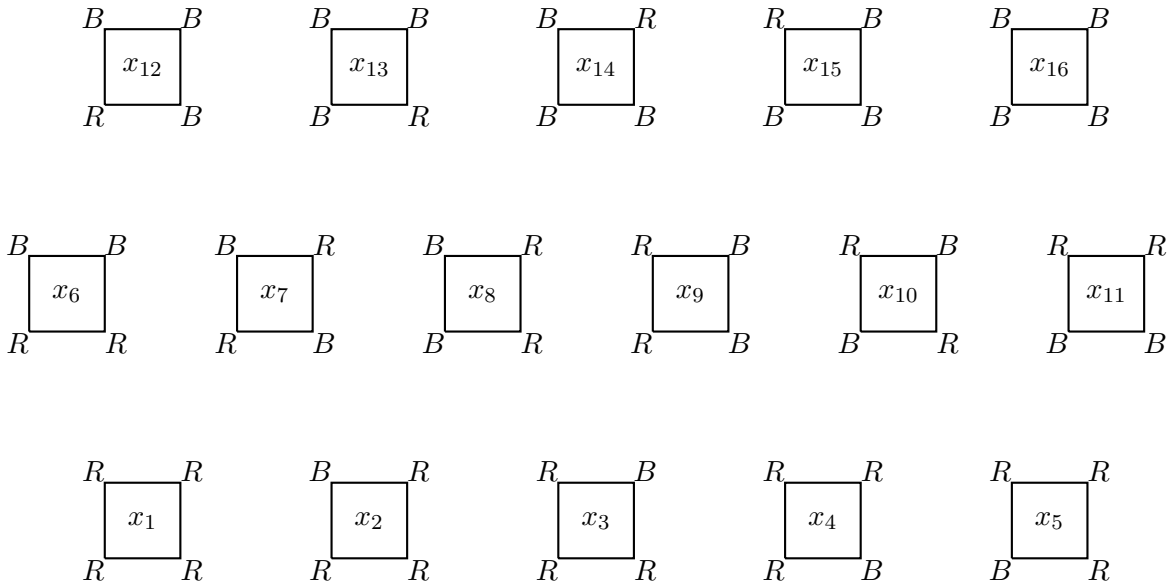


Figure 4.6: Coloring the vertices of a square.

There are three important sets associated with a group action. We first define them and then try to understand them using an example.

Definition 4.3.4. *Let G act on a set X. Then*

1. *for each fixed $x \in X$, $\mathcal{O}(x) = \{g \star x : \ g \in G\} \subset X$ is called the* Orbit *of x.*

2. *for each fixed $x \in X$, $G_x = \{g \in G : g \star x = x\} \subset G$ is called the* Stabilizer *of x in G.*

3. *for each fixed $g \in G$, $F_g = \{x \in X : g \cdot x = x\} \subset X$ is called the* Fix *of g.*

Let us now understand the above definitions using the following example.

Example 4.3.5. *Consider the set X given in Example 4.3.3.2. Then using the depiction of the set X in Figure 4.6, we have*

$$\mathcal{O}(x_2) = \{x_2, x_3, x_4, x_5\},\ G_{x_2} = \{e, rf\}, \quad \text{and} \ \ F_{rf} = \{x_1, x_2, x_4, x_7, x_{10}, x_{13}, x_{15}, x_{16}\}.$$

The readers should compute the different sets by taking other examples to understand the above defined sets.

We now state a few results associated with the above definitions. The proofs are omitted as they can be easily verified.

Proposition 4.3.6. *Let G act on a set X.*

1. *Then for each fixed $x \in X$, the set G_x is a subgroup of G.*

2. *Define a relation, denoted $\sim$, on the set X, by $x \sim y$ if there exists $g \in G$, such that $g \star x = y$. Then prove that $\sim$ defines an equivalence relation on the set X. Furthermore, the equivalence class containing $x \in X$ equals $\mathcal{O}(x) = \{g \star x : \ g \in G\} \subset X$.*

3. *Fix $x \in X$ and let $t \in \mathcal{O}(x)$. Then $\mathcal{O}(x) = \mathcal{O}(t)$. Moreover, if $g \star x = t$ then $G_x = g^{-1} G_t g$.*

Let G act on a set X. Then Proposition 4.3.6 helps us to relate the distinct orbits of X under the action of G with the cosets of G. This is stated and proved as the next result.

Theorem 4.3.7. *Let a group G act on a set X. Then for each fixed $x \in X$, there is a one-to-one correspondence between the elements of $\mathcal{O}(x)$ and the set of all left cosets of G_x in G. In particular,*

$$|\mathcal{O}(x)| = [G \ : \ G_x], \quad \text{the number of left cosets of } \ G_x \text{ in } G.$$

Moreover, if G is a finite group then $|G| = |\mathcal{O}(x)| \cdot |G_x|$, for all $x \in X$.

Proof. Let S be the set of distinct left cosets of G_x in G. Then $S = \{gG_x : \ g \in G\}$ and $|S| = [G : \ G_x]$. Consider the map $\tau : \ S \longrightarrow \mathcal{O}(x)$ by $\tau(gG_x) = g \star x$. Let us first check that this map is well-defined.

So, suppose that the left cosets gG_x and hG_x are equal. That is, $gG_x = hG_x$. Then, using Theorem 4.2.5 and the definition of group action, one obtains the following sequence of assertions:

$$gG_x = hG_x \iff h^{-1}g \in G_x \iff (h^{-1}g) \star x = x \iff h^{-1} \star (g \star x) = x \iff g \star x = h \star x.$$

Thus, by definition of the map τ, one has $gG_x = hG_x \iff \tau(gG_x) = \tau(hG_x)$. Hence, τ is not only well-defined but also one-one.

To show τ is onto, note that for each $y \in \mathcal{O}(x)$, there exists an $h \in G$, such that $h \star x = y$. Also, for this choice of $h \in G$, the coset $hG_x \in S$. Therefore, for this choice of $h \in G$, $\tau(hG_x) = h \star x = y$ holds. Hence, τ is onto.

Therefore, we have shown that τ gives a one-to-one correspondence between $\mathcal{O}(x)$ and the set S. This completes the proof of the first part. The other part follows by observing that by definition ; $[G : G_x] = \dfrac{|G|}{|G_x|}$, for each subgroup G_x of G whenever $|G|$ is finite. ■

The following lemmas are an immediate consequence of Proposition 4.3.6 and Theorem 4.3.7. We give the proof for the sake of completeness.

Lemma 4.3.8. *Let G be a finite group acting on a set X. Then, for each $y \in X$,*

$$\sum_{x \in \mathcal{O}(y)} |G_x| = |G|.$$

Proof. Recall that, for each $x \in \mathcal{O}(y)$, $|\mathcal{O}(x)| = |\mathcal{O}(y)|$. Hence, using Theorem 4.3.7, one has $|G| = |G_x| \cdot |\mathcal{O}(x)|$, for all $x \in X$. Therefore,

$$\sum_{x \in \mathcal{O}(y)} |G_x| = \sum_{x \in \mathcal{O}(y)} \frac{|G|}{|\mathcal{O}(x)|} = \sum_{x \in \mathcal{O}(y)} \frac{|G|}{|\mathcal{O}(y)|} = \frac{|G|}{|\mathcal{O}(y)|} \sum_{x \in \mathcal{O}(y)} 1 = \frac{|G|}{|\mathcal{O}(y)|} |\mathcal{O}(y)| = |G|.$$ ■

Theorem 4.3.9. *Let G be a finite group acting on a set X. Let N denote the number of distinct orbits of X under the action of G. Then*

$$N = \frac{1}{|G|} \sum_{x \in X} |G_x|.$$

Proof. By Lemma 4.3.8, note that $\sum_{x \in \mathcal{O}(y)} |G_x| = |G|$, for all $y \in X$. Let $x_1, x_2, \ldots, x_N$ be the representative of the distinct orbits of X under the action of G. Then

$$\frac{1}{|G|} \sum_{x \in X} |G_x| = \frac{1}{|G|} \sum_{i=1}^{N} \sum_{y \in \mathcal{O}(x_i)} |G_{x_i}| = \frac{1}{|G|} \sum_{i=1}^{N} |G| = \frac{1}{|G|} N \cdot |G| = N.$$ ■

Example 4.3.10. *Let us come back to Example 4.3.3.2. Check that the number of distinct colorings are*

$$\frac{1}{|G|} \sum_{i=1}^{16} |G_{x_i}| = \frac{1}{8} (8 + 2 + 2 + 2 + 2 + 2 + 4 + 2 + 2 + 4 + 2 + 2 + 2 + 2 + 2 + 8) = 6.$$

Observation: As the above example illustrates, we are able to find the number of distinct configurations using this method. But it is important to observe that this method requires us to list all elements of X. That is, if we need to list all the elements of X then we can already pick the ones that are distinct. So, the question arises what is the need of Theorem 4.3.9. Also,

if we color the vertices of the square with 3 colors, then $|X| = 3^4 = 81$, whereas the number of elements of D_4 (the group that acts as the group of symmetries of a square) remains 8. So, one feels that the calculation may become easy if one has to look at the elements of the group D_4 as one just needs to look at 8 elements of D_4. So, the question arises, can we get a formula that relates the number of distinct orbits with the elements of the group, in place of the elements of the set X? This query has an affirmative answer and is given as our next result.

Lemma 4.3.11 (Cauchy-Frobenius-Burnside's Lemma)**.** *Let G be a finite group acting on a set X. Let N denote the number of distinct orbits of X under the action of G. Then*

$$N = \frac{1}{|G|} \sum_{g \in G} |F_g|.$$

Proof. Consider the set $S = \{(g, x) \in G \times X \ : \ g \star x = x\}$. We calculate $|S|$ by two methods. As the first method, let us fix $x \in X$. Then, for each fixed $x \in X$, G_x gives the collection of elements of G that satisfy $g \star x = x$. So, $|S| = \sum\limits_{x \in X} |G_x|$.

As the second method, let us fix $g \in G$. Then, for each fixed $g \in G$, F_g gives the collection of elements of X that satisfy $g \star x = x$. So, $|S| = \sum\limits_{g \in G} |F_g|$. Thus, using two separate methods, one has $\sum_{x \in X} |G_x| = |S| = \sum_{g \in G} |F_g|$. Hence, using Theorem 4.3.9, we have

$$N = \frac{1}{|G|} \sum_{g \in G} |F_g|. \qquad \blacksquare$$

Example 4.3.12. *Let us come back to Example 4.3.3.2. Check that $|F_e| = 16$, $|F_r| = 2$, $|F_{r^2}| = 4, |F_{r^3}| = 2$, $|F_f| = 4$, $|F_{rf}| = 8, |F_{r^2 f}| = 4$ and $|F_{r^3 f}| = 8$. Hence, the number of distinct configurations are*

$$\frac{1}{|G|} \sum_{g \in G} |F_g| = \frac{1}{8} (16 + 2 + 4 + 2 + 4 + 8 + 4 + 8) = 6.$$

It seems that we may still need to know all the elements of X to compute the above terms. But it will be shown in the next section that to compute $|F_g$, for any $g \in G$, we just need to know the decomposition of g as product of disjoint cycles.

4.4 The Cycle Index Polynomial

Let G be a group acting on a set X. Then as mentioned at the end of the previous section, we need to understand the cycle decomposition of each $g \in G$ as product of disjoint cycles. Redfield and Polya observed that elements of G with the same cyclic decomposition made the same contribution to the sets of *fixed points*. They defined the notion of cycle index polynomial to keep track of the cycle decomposition of the elements of G. Let us start with a few definitions and examples to better understand the use of cycle decomposition of an element of a permutation group.

Definition 4.4.1. *A permutation $\sigma \in \mathcal{S}_n$ is said to have the cycle structure $1^{k_1}\ 2^{k_2} \cdots n^{k_n}$, if the cycle representation of σ has k_i cycles of length i, for $1 \leq i \leq n$. Observe that $\sum_{i=1}^{t} i \cdot k_i = n$.*

Example 4.4.2. *1. Let e be the identity element of $\mathcal{S}_n$. Then $e = (1)\ (2)\ \cdots (n)$ and hence the cycle structure of e, as an element of $\mathcal{S}_n$ equals 1^n.*

2. Let $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 3 & 6 & 7 & 10 & 14 & 1 & 2 & 13 & 15 & 4 & 11 & 5 & 8 & 12 & 9 \end{pmatrix}$. Then it can be easily verified that in the cycle notation, $\sigma = (1\ 3\ 7\ 2\ 6)\ (4\ 10)\ (5\ 14\ 12)\ (8\ 13)\ (9\ 15)\ (11)$. Thus, the cycle structure of σ is $1^1 2^3 3^1 5^1$.

3. Consider the group G of symmetries of the tetrahedron (see Example 4.1.9.2a). Then the elements of G have the following cycle structure:

1^4 *for exactly* 1 *element corresponding to the identity element*;
$1^1 3^1$ *for exactly* 8 *elements corresponding to* 3 *cycles*;
2^2 *for exactly* 3 *elements corresponding to* $(12)(34), (13)(24)\&(14)(23)$.

Exercise 4.4.3. *Determine the cycle structure of the following groups:*

1. The group in Example 4.1.9.2a.

2. The group in Example 4.1.9.1d. Note that this will depend on the factorization of n as product of distinct primes.

Definition 4.4.4. *Let G be a permutation group on n symbols. For a fixed $g \in G$, let $\ell_k(g)$ denote the number of cycles of length $k, 1 \leq k \leq n$, in the cycle representation of g. Then the cycle index polynomial of G, as a permutation group on n symbols, is a polynomial in n variables $z_1, z_2, \ldots, z_n$ given by*

$$P_G(z_1, z_2, \ldots, z_n) = \frac{1}{|G|} \left(\sum_{g \in G} z_1^{\ell_1(g)}\ z_2^{\ell_2(g)}\ \cdots z_n^{\ell_n(g)} \right).$$

Before, we look at a few examples, note that for each fixed $g \in G$, the condition that g has exactly $\ell_k(g)$ cycles of length $k, 1 \leq k \leq n$, implies that each term $z_1^{\ell_1(g)}\ z_2^{\ell_2(g)}\ \cdots z_n^{\ell_n(g)}$ in the summation satisfies $1 \cdot \ell_1(g) + 2 \cdot \ell_2(g) + \cdots + n \cdot \ell_n(g) = n$.

Example 4.4.5. *1. Let G be the dihedral group D_4 (see Example 4.1.9.2). Then*

$e = (1)(2)(3)(4) \longrightarrow z_1^4,\ r = (1234) \longrightarrow z_4,\ r^3 = (1432) \longrightarrow z_4,\ r^2 = (13)(24) \longrightarrow z_2^2,$
$f = (14)(23) \longrightarrow z_2^2,\ rf = (1)(3)(24) \longrightarrow z_1^2 z_2,\ r^2 f = (12)(34) \longrightarrow z_2^2,\ r^3 f = (13)(2)(4) \longrightarrow z_1^2 z_2.$

Thus, $P_G(z_1, z_2, z_3, z_4) = \frac{1}{8}\left(z_1^4 + 2z_4 + 3z_2^2 + 2z_1^2 z_2\right)$.

2. *Let G be the dihedral group D_5 (see Example 4.1.9.1c). Then*
$P_G(z_1, z_2, z_3, z_4, z_5) = \frac{1}{10}\left(z_1^5 + 4z_5 + 5z_1z_2^2\right)$.

3. *Verify that the cycle index polynomial of the symmetries of a cube induced on the set of vertices equals*
$$P_G(z_1, z_2, \ldots, z_8) = \frac{1}{24}\left(z_1^8 + 6z_4^2 + 9z_2^4 + 8z_1^2z_3^2\right).$$

4.4.1 Applications

Let S be an object (a geometrical figure) and let X be the finite set of points of S. Also, let C be a finite set (say, of colors). Consider the set Ω that denotes the set of all functions from X to C. Observe that an element of Ω gives a color pattern on the object S. Let G be a subgroup of the group of permutations of the object S. Hence, G acts on the elements of X. Let us denote this action by $\star$. So, $g \star x \in X$, for all $x \in X$.

One can also obtain an action of G on Ω, denoted $\circledast$, by the following rule:
Fix an element $x \in X$. Then, for each $\phi \in \Omega$ and $g \in G$, $g \circledast \phi$ is an element of Ω and hence it gives a function from X to C. Hence, one defines

$$(g \circledast \phi)(x) = \phi(g^{-1} \star x), \text{ for all } \phi \in \Omega.$$

We claim that $\circledast$ indeed defines a group action on the set Ω. To do so, note that for each $h, g \in G$ and $\phi \in \Omega$, the definition of the action on X and Ω gives

$$\begin{aligned}(h \circledast (g \circledast \phi))(x) &= (g \circledast \phi)(h^{-1} \star x) = \phi\left(g^{-1} \star (h^{-1} \star x)\right) = \phi\left(g^{-1}h^{-1} \star x\right)\\ &= \phi\left((hg)^{-1} \star x\right) = (hg \circledast \phi)(x).\end{aligned}$$

Since, $(h \circledast (g \circledast \phi))(x) = (hg \circledast \phi)(x)$, for all $x \in X$, one has $h \circledast (g \circledast \phi) = hg \circledast \phi$, for each $h, g \in G$ and $\phi \in \Omega$. Hence, the proof of the claim is complete. Now, using the above notations, we have the following theorem.

Theorem 4.4.6. *Let C, S, X and Ω be as defined above. Also, let G be a subgroup of the group of permutations of the object S. Then the number of distinct color patterns (distinct elements of Ω), distinct up to the action of G, is given by*

$$P_G(|C|, |C|, \ldots, |C|).$$

Proof. Let $|X| = n$. Then observe that G is a subgroup of $\mathcal{S}_n$. So, each $g \in G$ can be written as a product of disjoint cycles. Also, by Burnside's Lemma 4.3.11, N, the number of distinct color patterns (distinct orbits under the action of G), equals $\frac{1}{|G|}\sum_{g \in G} |F_g|$, where

$$F_g = \{\, \phi \in \Omega \;:\; (g \circledast \phi)(x) = \phi(x), \text{ for all } x \in X\}.$$

We claim that "$g \in G$ fixes a color pattern (or an element of Ω) if and only if ϕ colors the elements in a given cycle of g with the same color".

Suppose that $g \circledast \phi = \phi$. That is, $(g \circledast \phi)(x) = \phi(x)$, for all $x \in X$. So, using the definition, one has $\phi(g^{-1} \star x) = \phi(x)$, for all $x \in X$. In particular, for a fixed $x_0 \in X$, one also has

$$\phi(x_0) = \phi(g \star x_0) = \phi\big(g^2 \star x_0\big) = \cdots .$$

Note that, for each fixed $x_0 \in X$ and $g \in G$, the permutation $(x_0,\ g\star x_0,\ g^2\star x_0, \ldots)$ corresponds to a cycle of g. Therefore, if g fixes a color pattern ϕ, *i.e.*, $g \circledast \phi = \phi$, then ϕ assigns the same color to each element of any cycle of g.

Conversely, fix an element $g \in G$ and let ϕ be a color pattern (a function) that has the property that every point in a given cycle of g is colored with the same color. That is, $\phi(x) = \phi(g \star x)$, for each $x \in X$. Or equivalently, $\phi(x) = \phi\big(g^{-1} \star x\big) = (g \circledast \phi)(x)$, for all $x \in X$. Hence, by definition, $g \circledast \phi = \phi$. Thus, g fixes the color pattern ϕ. Hence, the proof of the claim is complete.

Therefore, we observe that for a fixed $g \in G$, a cycle of g can be given a color independent of another cycle of g. Also, the number of distinct colors equals $|C|$. Hence, for a fixed $g \in G$, $|F_g| = |C|^{\ell_1(g)} \cdot |C|^{\ell_2(g)} \ \cdots |C|^{\ell_n(g)}$, where for each $k,\ 1 \leq k \leq n$, $\ell_k(g)$ denotes the number of cycles of g of length k. Thus,

$$N = \frac{1}{|G|} \sum_{g \in G} |F_g| = \frac{1}{|G|} \sum_{g \in G} |C|^{\ell_1(g)} \cdot |C|^{\ell_2(g)} \ \cdots |C|^{\ell_n(g)} = P_G(|C|, |C|, \ldots, |C|). \qquad \blacksquare$$

We now give a few examples to indicate the importance of Theorem 4.4.6.

Example 4.4.7. *1. Determine the number of distinct color patterns, when the vertices of a pentagon is colored with* 3 *colors.*

Solution: *We know that the group D_5, is the group of symmetries of a pentagon. Hence, D_5 acts on the color patterns. Verify that*

$$P_{D_5}(z_1, z_2, \ldots, z_5) = \frac{1}{|D_5|}(z_1^5 + 4z_5 + 5z_1z_2^2) = \frac{z_1^5 + 4z_5 + 5z_1z_2^2}{10}.$$

Thus, by Theorem 4.4.6, the required number equals $N = \dfrac{1}{10}(3^5 + 4 \cdot 3 + 5 \cdot 3 \cdot 3^2) = 39$.

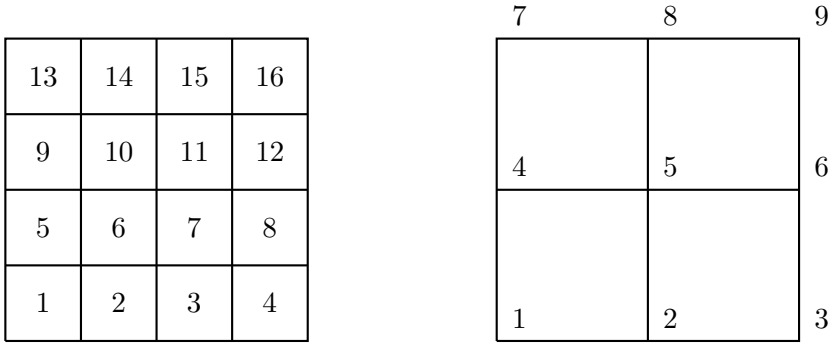
2. Suppose we are given beads of 3 *different colors and that there are at least* 6 *beads of each color. Determine the distinct necklace patterns that are possible using the* 6 *beads.*

Solution: *Since we are forming a necklace using* 6 *beads, the group D_6 acts on the* 6 *beads of the necklace. Also, the cycle index polynomial of D_6 equals $P_{D_6}(z_1, z_2, \ldots, z_5, z_6) = \dfrac{1}{|D_6|}(z_1^6 + 2z_6 + 2z_3^2 + z_2^3 + 3z_2^3 + 3z_1^2z_2^2)$. Hence, by Theorem 4.4.6, the number of distinct necklace patterns equals $\dfrac{1}{12}(3^6 + 2 \cdot 3 + 2 \cdot 3^2 + 4 \cdot 3^3 + 3 \cdot 3^2 \cdot 3^2) = 92$.*

3. Consider the 2×2 square given in Figure 4.7. Determine the number of distinct color patterns, when the vertices of the given figure are colored with two colors.

Solution: *Observe that D_4 is the group of symmetries of the 2×2 square and it needs to*

act on 9 *vertices. So, we need to write the elements of* D_4 *as a subgroup of* $\mathcal{S}_9$*. Hence, the cycle index polynomial is given by* $P_{D_4}(z_1, \ldots, z_9) = \dfrac{z_1^9 + 2z_1z_4^2 + z_1z_2^4 + 4z_1^3z_2^3}{8}$ *and the number of distinct color patterns equals* 102.



The 4×4 Square The 2×2 Square

Figure 4.7: Faces and Vertices of Squares

4. *Determine the number of distinct color patterns when the faces of a cube are colored with* 2 *colors.*

   Solution: *Using the group of symmetries of the cube given on Page 69, the cycle index polynomial corresponding to the faces equals* $P_G(z_1, \ldots, z_{12}) = \dfrac{z_1^{12} + 6z_4^3 + 3z_2^6 + 8z_3^4 + 6z_1^2z_2^5}{24}$. *Thus, the required number is* 218.

Exercise 4.4.8. *Determine the number of distinct color patterns when*

1. *the faces of the* 4×4 *square given in Figure 4.7 are colored with* 2 *colors.*

2. *the edges of a cube are colored with* 2 *colors. Hint: The cycle index polynomial equals* $P_G(z_1, z_2, \ldots, z_6) = \dfrac{1}{24}\left(z_1^6 + 6z_1^2z_4 + 3z_1^2z_2^2 + 6z_2^3 + 8z_3^2\right)$.

4.4.2 Polya's Inventory Polynomial

In this section, the ideas of the previous subsection are generalized. This generalization allows us to count the distinct number of necklaces even if there are not sufficient number of beads of each color. To do this, each element of C is assigned a *weight*, that in turn gives weight to each color pattern. This weight may be a number, a variable or in general, an element of a commutative ring with identity. The setup for our study remains the same. To start with, we have the following definitions.

Definition 4.4.9 (Weight of a color pattern)**.** *Let* A *be a commutative ring with identity (the elements of* A *are called weights). Let* $w : C \longrightarrow A$ *be a map that assigns weights to each color. Then the weight of a color pattern* $\phi : X \longrightarrow C$*, with respect to the weight function* w *is given by* $w(\phi) = \prod_{x \in X} w\big(\phi(x)\big)$.

Fix $g \in G$. Then we have seen that g fixes a color pattern $\phi \in \Omega$ if and only if ϕ colors the elements in a given cycle of g with the same color. Similarly, for each fixed $g \in G$ and $\phi \in \Omega$, one has

$$w(g \circledast \phi) = \prod_{x \in X} w\big(g \circledast \phi(x)\big) = \prod_{x \in X} w\big(\phi(g^{-1} \star x)\big) = \prod_{y \in X} w\big(\phi(y)\big) = w(\phi), \tag{4.1}$$

as $\{g \star x : x \in X\} = X$ (see Remark 4.3.2). That is, for a fixed $\phi \in \Omega$, the weight of each element of $\mathcal{O}(\phi) = \{g \circledast \phi : g \in G\}$ is the same and it equals $w(\phi)$. That is, $w(\phi) = w(\psi)$, whenever $\psi = g \circledast \phi$, for some $g \in G$.

Example 4.4.10. *Let X consist of the set of faces of a cube, G be the group of symmetries of the cube and let C consist of two colors 'Red' and 'Blue'. Thus, if the weights R and B are assigned to the two elements of C then the weight*

1. *B^6 corresponds "all faces being colored Blue";*
2. *R^2B^4 corresponds to "any two faces being colored 'Red' and the remaining four faces being colored 'Blue';*
3. *R^3B^3 corresponds to "any three faces being colored 'Red' and the remaining three faces being colored 'Blue' and so on.*

The above examples indicate that different color patterns need not have different weights. We also need the following definition to state and prove results in this area.

Definition 4.4.11 (Pattern Inventory)**.** *Let G be a group acting on the set Ω, the set of color patterns and let $w : C \longrightarrow A$ be a weight function. The pattern inventory, denoted I, under the action of G on Ω, with respect to w, is the sum of the weights of the orbits. That is, $I = \sum_{\Delta} w(\Delta)$, where the sum runs over all the distinct orbits Δ obtained by the action of G on Ω.*

With the above definitions, we are ready to prove the Polya's Enumeration Theorem. To do so, we first need to prove the weighted Burnside's Lemma. This Lemma is the weighted version of the Burnside's Lemma 4.3.11.

Lemma 4.4.12. *With the definitions and notations as above,*

$$I = \sum_{\Delta} w(\Delta) = \frac{1}{|G|} \sum_{g \in G} \sum_{\substack{\phi \in \Omega \\ g \circledast \phi = \phi}} w(\phi),$$

where the sum runs over all the distinct orbits Δ obtained by the action of G on Ω.

Proof. As G acts on Ω, for each $\alpha \in \Omega$, the application of Lemma 4.3.7 gives $|G_\alpha| \cdot |\mathcal{O}(\alpha)| = |G|$. Since Δ is an orbit under the action of G, for each $\phi \in \Delta$, $|G_\phi| \cdot |\Delta| = |G|$. Also, by definition, $w(\Delta) = w(\phi)$, for all $\phi \in \Delta$. Thus,

$$w(\Delta) = w(\phi) = \frac{1}{|\Delta|} \sum_{\phi \in \Delta} w(\phi) = \sum_{\phi \in \Delta} \frac{1}{|\Delta|} w(\phi) = \sum_{\phi \in \Delta} \frac{|G_\phi|}{|G|} w(\phi) = \frac{1}{|G|} \sum_{\phi \in \Delta} |G_\phi| \cdot w(\phi).$$

Let $F_g = \{\phi \in \Omega : g \circledast \phi = \phi\}$. Then $\sum_{\phi\in\Omega}\sum_{g\in G_\phi} w(\phi) = \sum_{g\in G}\sum_{\phi\in F_g} w(\phi)$ and hence

$$\begin{aligned} I &= \sum_{\Delta} w(\Delta) = \sum_{\Delta}\frac{1}{|G|}\sum_{\phi\in\Delta}|G_\phi|\cdot w(\phi) = \frac{1}{|G|}\sum_{\Delta}\sum_{\phi\in\Delta}|G_\phi|\cdot w(\phi) = \frac{1}{|G|}\sum_{\phi\in\Omega}|G_\phi|\cdot w(\phi) \\ &= \frac{1}{|G|}\sum_{\phi\in\Omega}\sum_{g\in G_\phi} w(\phi) = \frac{1}{|G|}\sum_{g\in G}\sum_{\phi\in F_g} w(\phi). \end{aligned}$$ ■

We are now in a position to prove the Polya's Enumeration Theorem. Before doing so, recall that F_g consists precisely of those color schemes which color each cycle of g with just one color (see the argument used in the second paragraph in the proof of Theorem 4.4.6).

Theorem 4.4.13 (Polya's Enumeration Theorem)**.** *With the definitions and notations as above,*

$$I = \sum_{\Delta} w(\Delta) = P_G(x_1, x_2, \ldots, x_n),$$

where the sum runs over all the distinct orbits Δ obtained by the action of G on Ω and $x_i = \sum_{c\in C} w(c)^i$, is the i^{th} power sum of the weights of the colors. In particular, if weight of each color is 1, $I = P_G(|C|, |C|, \ldots, |C|)$.

Proof. Using the weighted Burnside Lemma 4.4.12, we need to prove that

$$\sum_{g\in G}\sum_{\phi\in F_g} w(\phi) = \sum_{g\in G} x_1^{\ell_1(g)} x_2^{\ell_2(g)} \cdots x_n^{\ell_n(g)},$$

where $\ell_i(g)$ is the number of cycles of length i in the cycle representation of g.

Now, fix a $g \in G$. Suppose g has exactly t disjoint cycles, say $g_1, g_2, \ldots, g_t$. As F_g consists precisely of those color schemes which color each cycle of g with just one color, we just need to determine the weight of such a color pattern. To do so, for $1 \le i \le t$, define X_i to be that subset of X whose elements form the cycle g_i. Then, it is easy to see that $X_1, X_2, \ldots, X_t$ defines a partition of X. Also, the condition that x and $g \star x$ belong to the same cycle of g, one has $w(\phi(s_i)) = w(\phi(g \star s_i))$, for each $s_i \in X_i, 1 \le i \le t$. Thus, for each $\phi \in F_g$,

$$w(\phi) = \prod_{x\in X} w(\phi(x)) = \prod_{i=1}^{t}\prod_{x\in X_i} w(\phi(x)) = \prod_{i=1}^{t} w(\phi(s_i))^{|X_i|}.$$

Note that if we pick a term from each factor in $\prod_{i=1}^{t}\left(\sum_{c\in C} w(c)^{|X_i|}\right)$ and take the product of these terms, we obtain all the terms of $\prod_{i=1}^{t}\left(\sum_{c\in C} w(c)^{|X_i|}\right)$. All these terms also appear in $\sum_{\phi\in F_g}\prod_{i=1}^{t} w(\phi(s_i))^{|X_i|}$ because as ϕ is allowed to vary over all elements of F_g, the images $\phi(s_i)$, for $1 \le i \le t$, take all values in C. The argument can also be reversed and hence it follows that

$$\sum_{\phi\in F_g} w(\phi) = \sum_{\phi\in F_g}\prod_{i=1}^{t} w(\phi(s_i))^{|X_i|} = \prod_{i=1}^{t}\left(\sum_{c\in C} w(c)^{|X_i|}\right).$$

Now, assume that g has $\ell_k(g)$ cycles of length k, $1 \leq k \leq n$. This means that in the collection $|X_1|, |X_2|, \ldots, |X_t|$, the number 1 appears $\ell_1(g)$ times, the number 2 appears $\ell_2(g)$ times and so on till the number n appears $\ell_n(g)$ times (note that some of the $\ell_i(g)$'s may be zero). Consequently, $\prod_{i=1}^{t}\left(\sum_{c\in C} w(c)^{|X_i|}\right)$ equals $\prod_{k=1}^{n} x_k^{\ell_k(g)}$, as $x_1 = \sum_{c\in C} w(c)$, $x_2 = \sum_{c\in C} w(c)^2$ and so on till $x_n = \sum_{c\in C} w(c)^n$. Hence, $\sum_{\phi\in F_g} w(\phi) = \prod_{k=1}^{n} x_k^{\ell_k(g)}$ and thus, the required result follows. ■

Example 4.4.14. *1. Consider a necklace consisting of* 6 *beads. If there are* 3 *color choices, say* R, B *and* G*, then determine*

(a) the number of necklaces that have at least one R *bead.*

(b) the number of necklaces that have three R*, two* B *and one* G *bead.*

Solution: *Recall that* D_6 *acts on a regular hexagon and its cycle index polynomial equals*

$$P_{D_6}(z_1, z_2, \ldots, z_6) = \frac{1}{12}(z_1^6 + 4z_2^3 + 2z_3^2 + 2z_6 + 3z_1^2 z_2^2).$$

So, for the first part, at least one R *needs to be used and the remaining can be any number of* B *and/or* G*. So, we define the weight of the color* R *as* x *and that of* B *and* G *as* 1*. Therefore, by Polya's Enumeration Theorem 4.4.13,*

$$\begin{aligned} I &= \frac{1}{12}\big((x+1+1)^6 + 4(x^2+1+1)^3 + 2(x^3+1+1)^2 \\ &\qquad +2(x^6+1+1) + 3(x+1+1)^2(x^2+1+1)^2\big) \\ &= x^6 + 2x^5 + 9x^4 + 16x^3 + 29x^2 + 20x + 15. \end{aligned}$$

So, the required answer is $1 + 2 + 9 + 16 + 29 + 20 = 77$.

For the second part, define the weights as R, B *and* G *itself. Then*

$$\begin{aligned} I &= \frac{1}{12}\big((R+B+G)^6 + 4(R^2+B^2+G^2)^3 + 2(R^3+B^3+G^3)^2 \\ &\qquad +2(R^6+B^6+G^6) + 3(R+B+G)^2(R^2+B^2+G^2)^2\big). \end{aligned}$$

The required answer equals the coefficient of R^3B^2G *in* I*, which equals*

$$\frac{1}{12}\left(\binom{6}{3,2,1} + 3\cdot 2\cdot 2\right) = \frac{1}{12}\left(\frac{6!}{3!2!} + 6\right) = 6.$$

We end this chapter with a few Exercises. But before doing so, we give the following example with which Polya started his classic paper on this subject.

Example 4.4.15. *Suppose we are given* 6 *similar spheres in three different colors, say, three Red, two Blue and one Yellow (spheres of the same color being indistinguishable). In how many ways can we distribute the six spheres on the* 6 *vertices of an octahedron freely movable in space?*
Solution: *Here* $X = \{1, 2, 3, 4, 5, 6\}$ *and* $C = \{R, B, Y\}$*. Using Example 4.1.9.2b on Page 69*

the cycle index polynomial corresponding to the symmetric group of the octahedron that acts on the vertices of the octahedron is given by

$$\frac{1}{24}\left(z_1^6 + 6z_1^2z_4 + 3z_1^2z_2^2 + 8z_3^2 + 6z_2^3\right).$$

Hence, the number of patterns of the required type is the coefficient of the term R^3B^2Y in

$$\begin{aligned} I &= \frac{1}{24}\big((R+B+Y)^6 + 6(R+B+Y)^2(R^4+B^4+Y^4) + 3(R+B+Y)^2(R^2+B^2+Y^2)^2 \\ &\qquad +8(R^3+B^3+Y^3)^2 + 6(R^2+B^2+Y^2)^3\big). \end{aligned}$$

Verify that this number equals 3.

Exercise 4.4.16. *1. Three black and three white beads are strung together to form a necklace. If the beads of the same color are indistinguishable, determine the number of distinct necklace patterns, if the necklace can be only be rotated. What is the number if the necklace can be rotated and turned over?*

2. *Suppose the edges of a regular tetrahedron are being colored with white and black. Then determine the number of patterns that have exactly four black edges and two white edges.*

3. *Consider the molecules C_2H_4, C_6H_6 and CH_4 given in Figure 4.8. In each case, determine the number of possible molecules that can be formed, if the hydrogen atoms can be replaced by either Fluorine, Chlorine or Bromine.*

4. *In essentially how many different ways can we color the vertices of a cube if n colors are available?*

5. *Three ear-rings are shown in Figure 4.8. In each case, the ear-ring can be rotated along the horizontal axis passing through the central vertex (highlighted with dark circle). Then determine the following:*

   (a) *The group that acts on the ear-rings.*

   (b) *Write the elements of the group as a subgroup of $\mathcal{S}_n$, for a proper choice of n.*

   (c) *Determine the number of distinct color patterns when there are sufficient number of beads of both the colors "RED" and "BLUE".*

   (d) *Determine the possible distinct color patterns, when there are exactly* 6 *beads of color "RED".*

6. *Let p be a prime suppose that we want to make a necklace consisting of p beads. If for each bead, one has n choices of colors, then determine the number of distinct necklace patterns. Use this number to prove the Fermat's little theorem.*

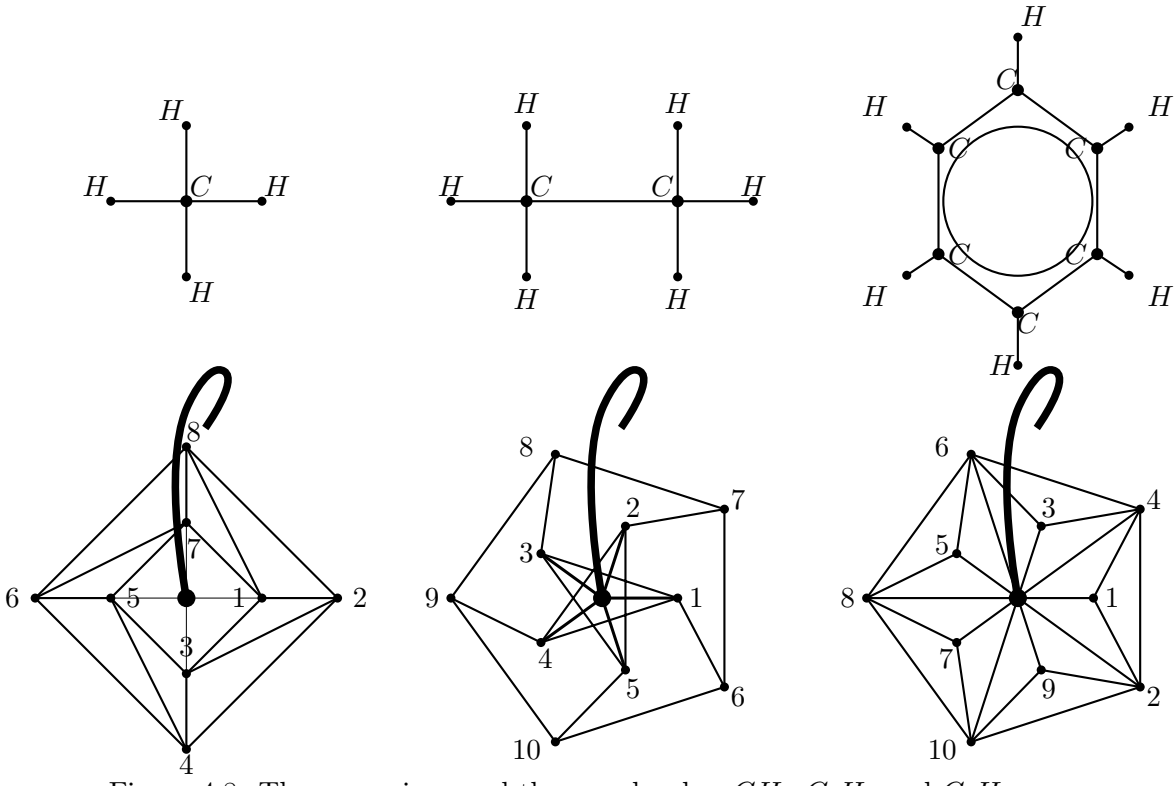


Figure 4.8: Three ear-rings and three molecules, CH_4, C_2H_6 and C_6H_6.

7. *Prove that the cycle index polynomial for the vertices, edges and faces of the octahedron is respectively, equal to*

$$\begin{aligned} P_G(z_1, z_2, \ldots, z_6) &= \frac{1}{24}\left(z_1^6 + 6z_1^2 z_4 + 3z_1^2 z_2^2 + 8z_3^2 + 6z_2^3\right), \\ P_G(z_1, z_2, \ldots, z_{12}) &= \frac{1}{24}\left(z_1^{12} + 6z_4^3 + 3z_2^6 + 8z_3^4 + 6z_1^2 z_2^5\right), \\ P_G(z_1, z_2, \ldots, z_8) &= \frac{1}{24}\left(z_1^8 + 6z_4^2 + 9z_2^4 + 8z_1^2 z_3^2\right). \end{aligned}$$

Chapter 5

Generating Functions and Its Applications

5.1 Formal Power Series

In this chapter, we will try to first develop the theory of generating functions by getting closed form expressions for some known recurrence relations. These ideas will be used to get some binomial identities.

To do so, we first recall from Page 47 that for all $n \in \mathbb{Q}$ and $k \in \mathbb{Z},\ k \geq 0$, the binomial coefficients, $\binom{n}{k}$, are well defined, using the idea that $\binom{n}{k} = \dfrac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$. We now start with the definition of "formal power series" and then its properties are studied in some detail.

Definition 5.1.1 (Formal power series)**.** *An algebraic expression of the form $f(x) = \sum\limits_{n\geq 0} a_n x^n$ is called a formal power series in the indeterminate x.*

Remark 5.1.2.

1. *Given a sequence of numbers $\{a_n : n = 0, 1, 2, \ldots\}$, one associates two formal power series, namely, $\sum\limits_{n\geq 0} a_n x^n$ and $\sum\limits_{n\geq 0} a_n \dfrac{x^n}{n!}$. The expression $\sum\limits_{n\geq 0} a_n x^n$ is called the* generating function *and the expression $\sum\limits_{n\geq 0} a_n \dfrac{x^n}{n!}$ is called the* exponential generating function*, for the numbers $\{a(n) : n \geq 0\}$.*

2. *Let $f(x) = \sum\limits_{n\geq 0} a_n x^n$ be a formal power series. Then the coefficient of x^n, for $n \geq 0$, in $f(x)$ is denoted by $[x^n]f(x)$. That is, $a_0 = [x^0]f(x)$ and for $n \geq 1$, $a_n = [x^n]f(x)$.*

3. *One just thinks of them as algebraic expressions. In general, one is not interested in evaluating them for any value of x. But if at all there is a need to evaluate, then one does find out the "radius of convergence" of the series and evaluates within that radius.*

In this chapter, our main aim is to study the series by means of algebraic rules. Let $\mathcal{P}(x)$ denote the set of all formal power series in the indeterminate x, with coefficients from $\mathbb{R}$. We need the following definition to proceed further.

Definition 5.1.3 (Equality of two formal power series)**.** *Two elements $f(x) = \sum_{n\geq 0} a_n x^n$ and $g(x) = \sum_{n\geq 0} b_n x^n$ of $\mathcal{P}(x)$ are said to be equal if $a_n = b_n$, for all $n \geq 0$.*

We are now ready to define the algebraic rules:

Definition 5.1.4. *Let $f(x) = \sum_{n\geq 0} a_n x^n, g(x) = \sum_{n\geq 0} b_n x^n \in \mathcal{P}(x)$. Then their*

1. *sum/addition is defined by*

$$f(x) + g(x) = \sum_{n\geq 0} a_n x^n + \sum_{n\geq 0} b_n x^n = \sum_{n\geq 0} (a_n + b_n) x^n.$$

2. *product is defined by*

$$f(x)\cdot g(x) = \left(\sum_{n\geq 0} a_n x^n\right)\cdot\left(\sum_{n\geq 0} b_n x^n\right) = \sum_{n\geq 0} c_n x^n, \quad \text{where} \quad c_n = \sum_{k=0}^{n} a_k b_{n-k}, \quad \text{for } n \geq 0.$$

This product is also called the Cauchy *product.*

Remark 5.1.5. 1. *In case of exponential power series, i.e., the product of $f(x) = \sum_{n\geq 0} a_n \frac{x^n}{n!}$ and $g(x) = \sum_{n\geq 0} b_n \frac{x^n}{n!}$ equals $\sum_{n\geq 0} d_n x^n$, where $d_n = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}$, for $n \geq 0$.*

2. *Note that the expression e^{e^x-1} is a well defined formal power series as the definition $e^y = \sum_{n\geq 0} \frac{y^n}{n!}$ implies that $e^{e^x-1} = \sum_{n\geq 0} \frac{(e^x-1)^n}{n!}$ and hence*

$$[x^m]e^{e^x-1} = [x^m]\sum_{n\geq 0} \frac{(e^x-1)^n}{n!} = \sum_{n=0}^{m} [x^m]\ \frac{(e^x-1)^n}{n!}. \tag{5.1}$$

That is, for each $m \geq 0$, $[x^m]e^{e^x-1}$ is a sum of a finite number of real numbers. Where as the expression e^{e^x} is not a formal power series as the computation of $[x^m]e^{e^x}$, for all $m \geq 0$, will indeed require an infinite sum.

Thus, under the algebraic operations defined above, it can be checked that the set $\mathcal{P}(x)$ forms a *Commutative Ring* with identity, where the identity element is given by the formal power series $f(x) = 1$. In this ring, the element $f(x) = \sum_{n\geq 0} a_n x^n$ is said to have a *reciprocal* if there exists another element $g(x) = \sum_{n\geq 0} b_n x^n \in \mathcal{P}(x)$ such that $f(x)\cdot g(x) = 1$. So, the questions arises, under what conditions on the coefficients of $f(x)$, can we find $g(x) \in \mathcal{P}(x)$ such that $f(x)g(x) = 1$. The answer to this question is given in the following proposition.

Proposition 5.1.6. *Let $f(x) = \sum_{n\geq 0} a_n x^n \in \mathcal{P}(x)$. Then there exists $g(x) \in \mathcal{P}(x)$ satisfying $f(x)\cdot g(x) = 1$ if and only if $a_0 \neq 0$.*

Proof. Let $g(x) = \sum_{n\geq 0} b_n x^n \in \mathcal{P}(x)$. Then, by the definition of Cauchy product, $f(x)g(x) = \sum_{n\geq 0} c_n x^n$, where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$, for all $n \geq 0$. Therefore, using the definition of equality of two power series, we see that $f(x)g(x) = 1$ if and only if $c_0 = 1$ and $c_n = 0$, for all $n \geq 1$.

Therefore, if $a_0 = 0$ then $c_0 = 0$ and hence the Cauchy product $f(x)g(x)$ can never equal 1. Now, let $a_0 \neq 0$. Then, we show that the coefficients b_n's can be recursively obtained as follows:

$b_0 = \dfrac{1}{a_0}$ as $1 = c_0 = a_0 b_0$.

$b_1 = \dfrac{-1}{a_0}\cdot(a_1 b_0)$ as $0 = a_0 b_1 + a_1 b_0$.

$b_2 = \dfrac{-1}{a_0}\cdot(a_2 b_0 + a_1 b_1)$ as $0 = a_0 b_2 + a_1 b_1 + a_2 b_0$. And in general, if we have already computed the values of b_k, for $k \leq r$, then

$b_{r+1} = \dfrac{-1}{a_0}\cdot(a_{r+1} b_0 + a_r b_1 + \cdots + a_1 b_r)$ as $0 = a_{r+1} b_0 + a_r b_1 + \cdots + a_1 b_r + a_0 b_{r+1}$.

■

Let us now look at the composition of two formal power series. Recall that if $f(x) = \sum_{n\geq 0} a_n x^n$ and $g(x) = \sum_{n\geq 0} b_n x^n$ then the composition $(f\circ g)(x) = f(g(x)) = \sum_{n\geq 0} c_n x^n$. Therefore, observe that the composition of two formal power series may not be defined (just to compute the constant term of the composition, one may have to look at an infinite sum). For example, let $f(x) = e^x$ and $g(x) = x + 1$. Note that $g(0) = 1 \neq 0$. Here, $(f\circ g)(x) = f(g(x)) = f(x+1) = e^{x+1}$. So, as function $f\circ g$ is well defined, but there is no formal procedure to write e^{x+1} as $\sum_{k\geq 0} a_k x^k$ and hence e^{x+1} is not a formal power series.

The next result gives the condition under which the composition $(f\circ g)(x)$ is well defined.

Proposition 5.1.7. *Let $f(x) = \sum_{n\geq 0} a_n x^n$ and $g(x) = \sum_{n\geq 0} b_n x^n$ be two formal power series. Then the composition $(f\circ g)(x)$ is well defined if either f is a polynomial or $b_0 = 0$.*

Moreover, suppose that $a_0 = 0$ and both $(f\circ g)(x)$ and $(g\circ f)(x)$ are well defined. Then $(f\circ g)(x) = x = (g\circ f)(x)$ under the condition that $b_0 = 0, a_1 \neq 0$ and $b_1 \neq 0$.

Proof. Let $(f\circ g)(x) = f(g(x)) = \sum_{n\geq 0} c_n x^n$ and suppose that either f is a polynomial or $b_0 = 0$. Then to compute $c_k = [x^k]\,(f\circ g)(x)$, for $k \geq 0$, one just needs to consider the terms $a_0 + a_1 g(x) + a_2 (g(x))^2 + \cdots + a_k (g(x))^k$. Hence, each c_k is a real number and $(f\circ g)(x)$ is well defined. This completes the proof of the first portion. The proof of the other part is left to the readers. ■

We now define the formal differentiation of elements of $\mathcal{P}(x)$.

Definition 5.1.8 (Differentiation)**.** *Let $f(x) = \sum_{n\geq 0} a_n x^n \in \mathcal{P}(x)$. Then the formal differentiation of $f(x)$, denoted $Df(x)$, is defined by*

$$Df(x) = a_1 + 2a_2x + \cdots + na_nx^{n-1} + \cdots = \sum_{n\geq 1} na_nx^{n-1}.$$

With the definitions as above, the following proposition can be easily proved. So, we omit the proof.

Proposition 5.1.9. *Let $f(x) = \sum_{n\geq 0} a_n x^n \in \mathcal{P}(x)$. Then $f(x) = a_0$, a constant, whenever $Df(x) = 0$. Also, $f(x) = a_0e^x$ whenever $Df(x) = f(x)$.*

Before proceeding further, let us look at some important examples, where the above results can be used.

Example 5.1.10. *Using the generalized binomial theorem (see Theorem 2.3.1), recall that the formal power series*

1. *$\sum_{n\geq 0} x^n$ equals $\dfrac{1}{1-x}$.*

2. *$\sum_{n\geq 0} \binom{n+r-1}{n} x^n$ equals $\dfrac{1}{(1-x)^r}$.*

3. *Determine a closed form expression for $\sum_{n\geq 0} nx^n \in \mathcal{P}(x)$.*

   Solution: *As $\dfrac{1}{1-x} = \sum_{n\geq 0} x^n$, one has $\dfrac{1}{(1-x)^2} = D\left(\dfrac{1}{1-x}\right) = D\left(\sum_{n\geq 0} x^n\right) = \sum_{n\geq 0} nx^{n-1}$. Thus, the closed form expression is $\dfrac{x}{(1-x)^2}$.*

   This can also be computed as follows:
   Let $S = \sum_{n\geq 0} nx^n = x + 2x^2 + 3x^3 + \cdots$. Then $xS = x^2 + 2x^3 + 3x^4 + \cdots$. Hence, $(1-x)S = x + x^2 + x^3 + \cdots = x(1 + x + x^2 + \cdots) = \dfrac{x}{1-x}$. Thus, $S = \dfrac{x}{(1-x)^2}$.

4. *Let $f(x) = \sum_{n\geq 0} a_n x^n \in \mathcal{P}(x)$. Determine $\sum_{k=0}^{n} a_k$.*
   Solution: *Recall that the Cauchy product of $f(x) = \sum_{n\geq 0} a_n x^n$ and $g(x) = \sum_{n\geq 0} b_n x^n$ equals $\sum_{n\geq 0} c_n x^n$, where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$, for $n \geq 0$. Therefore, to get $c_n = \sum_{k=0}^{n} a_k$, one needs $b_k = 1$, for all $k \geq 0$. That is, $\sum_{k=0}^{n} a_k = c_n = [x^n]\left(f(x) \cdot \dfrac{1}{1-x}\right)$.*

   Hence, the Cauchy product helps us in computing the sum of the first N coefficients of a formal power series, for any $N \geq 1$.

5. *Determine the sum of the squares of the first N positive integers.*

Solution: *Using Example 5.1.10.2, observe that* $k = [x^{k-1}]\left(\frac{1}{(1-x)^2}\right)$. *Therefore, using Example 5.1.10.4, one has*

$$\sum_{k=1}^{N} k = [x^{N-1}]\left(\frac{1}{(1-x)^2}\cdot\frac{1}{1-x}\right) = [x^{N-1}]\,\frac{1}{(1-x)^3} = \binom{N-1+3-1}{N-1} = \frac{N(N+1)}{2}.$$

6. *Determine a closed form expression for* $\sum_{k=1}^{N} k^2$.

Solution: *Using Example 5.1.10.3, observe that* $\sum_{k\geq 0} kx^k = \frac{x}{(1-x)^2}$. *Therefore, using the differentiation operator, one obtains*

$$\sum_{k\geq 0} k^2x^k = x\left(\sum_{k\geq 0} k^2x^{k-1}\right) = xD\left(\frac{x}{(1-x)^2}\right) = \frac{x(1+x)}{(1-x)^3}. \tag{5.2}$$

Thus, by Example 5.1.10.4

$$\begin{aligned}\sum_{k=1}^{N} k^2 &= [x^N]\left(\frac{x(1+x)}{(1-x)^3}\cdot\frac{1}{1-x}\right) = [x^{N-1}]\left(\frac{1}{(1-x)^4}\right) + [x^{N-2}]\left(\frac{1}{(1-x)^4}\right)\\ &= \binom{N-1+4-1}{N-1} + \binom{N-2+4-1}{N-2}\\ &= \frac{N(N+1)(2N+1)}{6}.\end{aligned}$$

7. *Determine a closed form expression for* $\sum_{k=1}^{N} k^3$.

Solution: *Using Equation (5.2), observe that* $\sum_{k\geq 0} k^2x^k = \frac{x(1+x)}{(1-x)^3}$. *So,*

$$\sum_{k\geq 0} k^3x^k = x\left(\sum_{k\geq 0} k^3x^{k-1}\right) = xD\left(\frac{x(1+x)}{(1-x)^3}\right) = \frac{x(1+4x+x^2)}{(1-x)^4}.$$

Thus, by Example 5.1.10.4

$$\begin{aligned}\sum_{k=1}^{N} k^3 &= [x^N]\left(\frac{x(1+4x+x^2}{(1-x)^4)}\cdot\frac{1}{1-x}\right)\\ &= [x^{N-1}]\left(\frac{1}{(1-x)^5}\right) + [x^{N-2}]\left(\frac{4}{(1-x)^5}\right) + [x^{N-3}]\left(\frac{1}{(1-x)^5}\right)\\ &= \binom{N-1+5-1}{N-1} + 4\binom{N-2+5-1}{N-2} + \binom{N-3+5-1}{N-3}\\ &= \left(\frac{N(N+1)}{2}\right)^2.\end{aligned}$$

Hence, we observe that we can inductively use this technique to get a closed form expression for $\sum_{k=1}^{N} k^r$, *for any positive integer* r.

8. *Determine a closed form expression for* $\sum_{n\geq 0} \frac{n^2+n+6}{n!}$.

   Solution: *As we need to compute the infinite sum, Cauchy product cannot be used. Also, one needs to find a convergent series, which when evaluated at some* x_0*, gives the required expression. Therefore, recall that the series* $e^x = \sum_{n\geq 0} \frac{x^n}{n!}$*, converges for all* $x \in \mathbb{R}$ *and evaluating* e^x *at* $x = 1$*, gives* $\sum_{n\geq 0} \frac{1}{n!} = e$*. Similarly,* $\frac{n}{n!} = [x^n]\,(xD(e^x)) = [x^n]\,(xe^x)$ *and* $\frac{n^2}{n!} = [x^n]\,(xD(xDe^x)) = [x^n]\left((x+x^2)e^x\right)$*. Thus,*

$$\sum_{n\geq 0} \frac{n^2+n+6}{n!} = (x+x^2)e^x + xe^x + 6e^x\Big|_{x=1} = 9e.$$

9. *Let* n *and* r *be two fixed positive integers. Then determine the number of non-negative integer solutions to the system* $x_1 + x_2 + \cdots + x_n = r$*?*

   Solution: *Recall that this number was already computed in Lemma 2.1.2 and equals* $\binom{r+n-1}{r}$.

   In this chapter, we can think of the problem as follows: the above system can be interpreted as coming from the monomial x^r*, where* $r = x_1+x_2+\cdots+x_n$*. That is, the problem reduces to finding the coefficients of* y^{x_k} *of a formal power series, for non-negative integers* x_k*'s. Now, recall that the terms* y^{x_k} *appear with a coefficient* 1 *in the expression* $\frac{1}{1-y} = \sum_{i\geq 0} y^i$*. Hence, the question reduces to computing*

$$[y^r]\left(\frac{1}{(1-y)(1-y)\cdots(1-y)}\right) = [y^r]\,\frac{1}{(1-y)^n} = \binom{r+n-1}{r}.$$

We now look at some examples that may require the use the package "MATHEMATICA" or "MAPLE" to obtain the exact answer. So, in the examples that we give below, we are interested in getting a formal power series and then its coefficients give the answer to the questions raised.

Example 5.1.11. 1. *Let* n *and* r *be two fixed positive integers. Then determine the number of non-negative integer solutions to the system* $x_1 + 2x_2 + \cdots + nx_n = r$*?*

   Solution: *Note that using the ideas in Example 5.1.10.9, one needs to consider the formal power series* $\sum_{i\geq 0} x^{ki}$ *that equals* $\frac{1}{1-x^k}$*. Hence, the question reduces to computing the coefficient of* x^r *in*

$$\frac{1}{(1-x)(1-x^2)\cdots(1-x^n)}.$$

2. *Determine the number of solutions in non-negative integer to the system* $x_1+x_2+\cdots+x_5 = n$ *such that* $x_1 \geq 4$*,* $x_4 \leq 10$ *and for* $r \neq 1, 4$*,* x_r *is a multiple of* r*.*

   Solution: *Note that the condition* $x_1 \geq 4$ *corresponds to looking at* x^k*, for* $k \geq 4$*, forcing us to look at the formal power series* $\sum_{k\geq 4} x^k$*. Similarly,* $x_4 \leq 10$ *gives the formal power*

series $\sum_{k=0}^{10} x^k$ *and the condition* x_r *is a multiple of* r*, for* $r \neq 1, 4$*, gives the formal power series* $\sum_{k\geq 0} x^{rk}$*. So, we are interested in computing the coefficient of* x^n *in the product*

$$\left(\sum_{k\geq 4} x^k\right)\cdot\left(\sum_{k=0}^{10} x^k\right)\cdot\left(\sum_{k\geq 0} x^{2k}\right)\cdot\left(\sum_{k\geq 0} x^{3k}\right)\cdot\left(\sum_{k\geq 0} x^{5k}\right) = \frac{x^4(1-x^{11})}{(1-x)^2(1-x^2)(1-x^3)(1-x^5)}.$$

3. *Determine the number of ways in which* 100 *voters can cast their* 100 *votes for* 10 *candidates such that no candidate gets more than* 20 *votes.*
   Solution: *Note that we are assuming that the voters are identical. So, we need to solve the system in non-negative integers to the system* $x_1 + x_2 + \cdots + x_{10} = 100$*, with* $0 \leq x_i \leq 20$*, for* $1 \leq i \leq 10$*. So, we need to find the coefficient of* x^{100} *in*

$$\begin{aligned}\left(\sum_{k=1}^{20} x^k\right)^{10} &= \frac{(1-x^{21})^{10}}{(1-x)^{10}} = \left(\sum_{i=0}^{10}(-1)^i\binom{10}{i}x^{21i}\right)\cdot\left(\sum_{j\geq 0}\binom{10+j-1}{j}x^j\right)\\ &= \sum_{i=0}^{4}(-1)^i\binom{10}{i}\cdot\binom{109-21i}{9}.\end{aligned}$$

Exercise 5.1.12. *Let* m, n*, and* r *be fixed positive integers. Then prove that the following problems are equivalent?*

1. *Determine the number of solutions in non-negative integer to the system*

$$x_1 + x_2 + \cdots + x_n = r, \quad \textit{with} \quad m \leq x_i \leq 2m?$$

2. *Determine the number of ways of putting* r *indistinguishable balls into* n *distinguishable boxes so that the number of balls in each box is a number between* m *and* $2m$ *(endpoints included)?*

3. *What is the coefficient of* x^r *in the formal power series* $\dfrac{x^{mn}\,(1-x^{m+1})^n}{(1-x)^n}$*?*

Before moving to the applications of generating functions/formal power series to the solution of recurrence relations, let us list a few well known power series. The readers are requested to get a proof of their satisfaction.

$$
\begin{array}{lcllcl}
e^x & = & \sum_{k\geq 0} \frac{x^k}{k!} & \log(1-x) & = & -\sum_{k\geq 1} \frac{x^k}{k},\ |x|<1 \\
(1+x)^a & = & \sum_{r\geq 0} \binom{a}{k} x^k,\ |x|<1 & \frac{1}{1-x} & = & \sum_{k\geq 0} x^k,\ |x|<1 \\
\frac{1}{(1-x)^a} & = & \sum_{k\geq 0} \binom{a+k-1}{k} x^k,\ |x|<1 & \frac{1}{\sqrt{1-4x}} & = & \sum_{k\geq 0} \binom{2k}{k} x^k,\ |x|<\frac{1}{4} \\
\frac{x^{-r}}{(1-x)^n} & = & \sum_{k\geq -r} \binom{n}{r+k} x^k,\ |x|<1 & \frac{x^n}{(1-x)^{n+1}} & = & \sum_{k\geq 0} \binom{n}{k} x^n,\ n\geq 0, |x|<1 \\
\sin(x) & = & \sum_{r\geq 0} \frac{(-1)^r x^{2r+1}}{(2r+1)!} & \cos(x) & = & \sum_{r\geq 0} \frac{(-1)^r x^{2r}}{(2r)!} \\
\sinh(x) & = & \sum_{r\geq 0} \frac{x^{2r+1}}{(2r+1)!} & \cosh(x) & = & \sum_{r\geq 0} \frac{x^{2r}}{(2r)!}
\end{array}
$$

$$
\begin{aligned}
\frac{1-\sqrt{1-4x}}{2x} &= \sum_{k\geq 0} \frac{1}{k+1}\binom{2k}{k} x^k,\ |x|<\frac{1}{4} \\
\frac{1}{\sqrt{1-4x}}\left(\frac{1-\sqrt{1-4x}}{2x}\right)^n &= \sum_{k\geq 0} \binom{2k+n}{k} x^k,\ |x|<\frac{1}{4}.
\end{aligned}
$$

5.2 Applications to Recurrence Relation

This section contains the applications of generating functions to solving recurrence relations. Let us try to understand it using the following examples.

Example 5.2.1. *1. Determine a formula for the numbers $a(n)$'s, where $a(n)$'s satisfy the recurrence relation $a(n) = 3\,a(n-1) + 2n$, for $n \geq 1$ with $a(0) = 1$.*

Solution: *Define $A(x) = \sum_{n\geq 0} a(n)x^n$. Then using Example 5.1.10.3, one has*

$$
\begin{aligned}
A(x) &= \sum_{n\geq 0} a(n)x^n = a_0 + \sum_{n\geq 1} a(n)x^n = 1 + \sum_{n\geq 1} (3\,a(n-1) + 2n)\,x^n \\
&= 3x\sum_{n\geq 1} a(n-1)x^{n-1} + 2\sum_{n\geq 1} nx^n + 1 = 3xA(x) + 2\frac{x}{(1-x)^2} + 1.
\end{aligned}
$$

So, $A(x) = \dfrac{1+x^2}{(1-3x)(1-x)^2} = \dfrac{5}{2(1-3x)} - \dfrac{1}{2(1-x)} - \dfrac{1}{(1-x)^2}$. Thus,

$$
a(n) = [x^n]A(x) = \frac{5}{2}3^n - \frac{1}{2} - (n+1) = \frac{5\cdot 3^n - 1}{2} - (n+1).
$$

2. Determine a generating function for the numbers $f(n)$ that satisfy the recurrence relation

$$
f(n) = f(n-1) + f(n-2), \quad \text{for} \quad n \geq 2 \quad \text{with} \ \ f(0) = 1 \ \ \text{and} \ f(1) = 1.
$$

Hence or otherwise find a formula for the numbers $f(n)$.

Solution: *Define* $F(x) = \sum_{n\geq 0} f(n)x^n$. *Then one has*

$$\begin{aligned} F(x) &= \sum_{n\geq 0} f(n)x^n = 1 + x\sum_{n\geq 2}\left(f(n-1)+f(n-2)\right)x^n \\ &= 1 + x + x\sum_{n\geq 2} f(n-1)x^{n-1} + x^2\sum_{n\geq 2} f(n-2)x^{n-2} = 1 + xF(x) + x^2F(x). \end{aligned}$$

Therefore, $F(x) = \dfrac{1}{1-x-x^2}$. *Let* $\alpha = \dfrac{1+\sqrt{5}}{2}$ *and* $\beta = \dfrac{1-\sqrt{5}}{2}$. *Then it can be checked that* $(1-\alpha x)(1-\beta x) = 1-x-x^2$ *and*

$$F(x) = \frac{1}{\sqrt{5}}\left(\frac{\alpha}{1-\alpha x} - \frac{\beta}{1-\beta x}\right) = \frac{1}{\sqrt{5}}\left(\sum_{n\geq 0}\alpha^{n+1}x^n - \sum_{n\geq 0}\beta^{n+1}x^n\right).$$

Therefore,

$$f(n) = [x^n]F(x) = \frac{1}{\sqrt{5}}\sum_{n\geq 0}\left(\alpha^{n+1} - \beta^{n+1}\right).$$

As $\beta < 0$ *and* $|\beta| < 1$, *we observe that* $f(n) \approx \dfrac{1}{\sqrt{5}}\left(\dfrac{1+\sqrt{5}}{2}\right)^{n+1}$.

Remark 5.2.2. *The numbers* $f(n)$, *for* $n \geq 0$ *are called* FIBONACCI NUMBERS. *It is related with the following problem:* Suppose a couple bought a pair of rabbits (each one year old) in the year 2001. If a pair of rabbits starts giving birth to a pair of rabbits as soon as they grow 2 years old, determine the number of rabbits the couple will have in the year 2025.

3. *Determine a formula for the numbers* $a(n)$*'s, where* $a(n)$*'s satisfy the recurrence relation* $a(n) = 3\,a(n-1) + 4a(n-2)$, *for* $n \geq 2$ *with* $a(0) = 1$ *and* $a(1) = c$, *a constant.*
Solution: *Define* $A(x) = \sum_{n\geq 0} a(n)x^n$. *Then*

$$\begin{aligned} A(x) &= \sum_{n\geq 0} a(n)x^n = a_0 + a_1x + \sum_{n\geq 2} a(n)x^n = 1 + cx + \sum_{n\geq 2}\left(3\,a(n-1) + 4\,a(n-2)\right)x^n \\ &= 1 + cx + 3x\sum_{n\geq 2} a(n-1)x^{n-1} + 4x^2\sum_{n\geq 2} a_{n-2}x^{n-2} \\ &= 1 + cx + 3x(A(x) - a_0) + 4x^2A(x). \end{aligned}$$

So, $A(x) = \dfrac{1+(c-3)x}{(1-3x-4x^2)} = \dfrac{1+(c-3)x}{(1+x)(1-4x)}$.

(a) If $c = 4$ *then* $A(x) = \dfrac{1}{1-4x}$ *and hence* $a_n = [x^n]\,A(x) = 4^n$.

(b) If $c \neq 4$ *then* $A(x) = \dfrac{1+c}{5}\cdot\dfrac{1}{1-4x} + \dfrac{4-c}{5}\cdot\dfrac{1}{1+x}$ *and hence*
$a_n = [x^n]\,A(x) = \dfrac{(1+c)\,4^n}{5} + \dfrac{(-1)^n(4-c)}{5}$.

4. *Determine a sequence,* $\{a(n) \in \mathbb{R} : n \geq 0\}$, *such that* $a_0 = 1$ *and* $\sum_{k=0}^{n} a(k)a(n-k) = \binom{n+2}{2}$, *for all* $n \geq 1$.

Solution: *Define* $A(x) = \sum_{n\geq 0} a(n)x^n$. *Then, using the Cauchy product, one has*

$$A(x)^2 = \sum_{n\geq 0}\left(\sum_{k=0}^{n} a(k)a(n-k)\right)x^n = \sum_{n\geq 0}\binom{n+2}{2}x^n = \frac{1}{(1-x)^3}.$$

Hence, $A(x) = \frac{1}{(1-x)^{3/2}}$ *and thus* $a(n) = \binom{-3/2}{n} = \dfrac{3\cdot 5\cdot 7\cdots(2n+1)}{2^n\, n!}$, *for all* $n \geq 1$.

5. *Determine a generating function for the numbers* $f(n,m),\ n,m \in \mathbb{Z}, n,m \geq 0$ *that satisfy*

$$\begin{aligned} f(n,m) &= f(n-1,m) + f(n-1,m-1),\ (n,m) \neq (0,0) \qquad \textit{with} \\ &\quad f(n,0) = 1, \ \textit{ for all } n \geq 0 \ \textit{ and } f(0,m) = 0, \ \textit{for all } m > 0. \end{aligned} \tag{5.1}$$

Hence or otherwise, find a formula for the numbers $f(n,m)$.

Solution: *Note that in the above recurrence relation, the value of* m *need not be* $\leq n$.

Method 1: *Define* $F_n(x) = \sum_{m\geq 0} f(n,m)x^m$. *Then, for* $n \geq 1$, *Equation (5.1) gives*

$$\begin{aligned} F_n(x) &= \sum_{m\geq 0} f(n,m)x^m = \sum_{m\geq 0}\left(f(n-1,m) + f(n-1,m-1)\right)x^m \\ &= \sum_{m\geq 0} f(n-1,m)x^m + \sum_{m\geq 0} f(n-1,m-1)x^m \\ &= F_{n-1}(x) + xF_{n-1}(x) = (1+x)F_{n-1}(x) = \cdots = (1+x)^n F_0(x). \end{aligned}$$

Now, using the initial conditions, $F_0(x) = 1$ *and hence* $F_n(x) = (1+x)^n$. *Thus,*

$$f(n,m) = [x^m](1+x)^n = \binom{n}{m} \quad \textit{if} \quad 0 \leq m \leq n \quad \textit{and} \quad f(n,m) = 0, \quad \textit{for} \quad m > n.$$

Method 2: *Define* $G_m(y) = \sum_{n\geq 0} f(n,m)y^n$. *Then, for* $m \geq 1$, *Equation (5.1) gives*

$$\begin{aligned} G_m(y) &= \sum_{n\geq 0} f(n,m)y^n = \sum_{n\geq 0}\left(f(n-1,m) + f(n-1,m-1)\right)y^n \\ &= \sum_{n\geq 0} f(n-1,m)y^n + \sum_{n\geq 0} f(n-1,m-1)y^n \\ &= yG_m(y) + yG_{m-1}(y). \end{aligned}$$

Therefore, $G_m(y) = \dfrac{y}{1-y}G_{m-1}(y)$. *Now, using initial conditions,* $G_0(y) = \dfrac{1}{1-y}$ *and hence* $G_m(y) = \dfrac{y^m}{(1-y)^{m+1}}$. *Thus,* $f(n,m) = [y^n]\dfrac{y^m}{(1-y)^{m+1}} = [y^{n-m}]\dfrac{1}{(1-y)^{m+1}} = \binom{n}{m}$, *whenever* $0 \leq m \leq n$ *and* $f(n,m) = 0$, *for* $m > n$.

6. Determine a generating function for the numbers $S(n,m),\ n,m \in \mathbb{Z}, n,m \geq 0$ that satisfy

$$\begin{aligned} S(n,m) &= mS(n-1,m) + S(n-1,m-1),\ (n,m) \neq (0,0) \qquad \textit{with} \\ & \qquad S(0,0) = 1, S(n,0) = 0, \ \textit{for all}\ n > 0\ \textit{and}\ S(0,m) = 0, \ \textit{for all}\ m > 0. \end{aligned} \tag{5.2}$$

Hence or otherwise find a formula for the numbers $S(n,m)$.

Solution: *Define $G_m(y) = \sum_{n\geq 0} S(n,m)y^n$. Then, for $m \geq 1$, Equation (5.2) gives*

$$\begin{aligned} G_m(y) &= \sum_{n\geq 0} S(n,m)y^n = \sum_{n\geq 0} \left(mS(n-1,m) + S(n-1,m-1)\right) y^n \\ &= m\sum_{n\geq 0} S(n-1,m)y^n + \sum_{n\geq 0} S(n-1,m-1)y^n \\ &= myG_m(y) + yG_{m-1}(y). \end{aligned}$$

Therefore, $G_m(y) = \dfrac{y^m}{1-my} G_{m-1}(y)$. Using initial conditions, $G_0(y) = 1$ and hence

$$G_m(y) = \frac{y^m}{(1-y)(1-2y)\cdots(1-my)} = y^m \sum_{k=1}^{m} \frac{\alpha_k}{1-ky}, \tag{5.3}$$

where $\alpha_k = \dfrac{(-1)^{m-k}k^m}{k!\,(m-k)!}$, for $1 \leq k \leq m$. Thus,

$$\begin{aligned} S(n,m) &= [y^n]\left(y^m \sum_{k=1}^{m} \frac{\alpha_k}{1-ky}\right) = \sum_{k=1}^{m} [y^{n-m}] \frac{\alpha_k}{1-ky} \\ &= \sum_{k=1}^{m} \alpha_k k^{n-m} = \sum_{k=1}^{m} \frac{(-1)^{m-k}k^n}{k!\,(m-k)!} \\ &= \frac{1}{m!} \sum_{k=1}^{m} (-1)^{m-k} k^n \binom{m}{k} = \frac{1}{m!} \sum_{k=1}^{m} (-1)^k (m-k)^n \binom{m}{k}. \end{aligned} \tag{5.4}$$

Therefore, $S(n,m) = \dfrac{1}{m!} \sum\limits_{k=1}^{m} (-1)^k (m-k)^n \binom{m}{k}$ and $m!\, S(n,m) = \sum\limits_{k=1}^{m} (-1)^k (m-k)^n \binom{m}{k}$.

The above expression was already obtained earlier (see Equation (2.1) and Exercise 6). This identity is generally known as the Stirling's Identity.

Observation:

(a) $H_n(x) = \sum\limits_{m\geq 0} S(n,m)x^m$ is not considered. But verify that

$$H_n(x) = (x + xD)^n \cdot 1 \quad \textit{as} \quad H_0(x) = 1.$$

Therefore, $H_1(x) = x,\ H_2(x) = x + x^2, \cdots$. Hence, it is difficult to obtain a general formula for its coefficients. But it is helpful in showing that the numbers $S(n,m)$, for fixed n, first increase and then decrease (commonly called unimodal*). The same holds for the sequence of binomial coefficients $\left\{\binom{n}{m}, m = 0, 1, \ldots, n\right\}$.*

(b) *Since there is no restriction on the non-negative integers n and m, the expression Equation (5.4) is also valid for $n < m$. But, in this case, we know that $S(n,m) = 0$. Hence, verify that* $\sum_{k=1}^{m} \frac{(-1)^{m-k}\, k^{n-1}}{(k-1)!\,(m-k)!} = 0$, *whenever* $n < m$.

7. Bell Numbers: *For a positive integer n, the n^{th} Bell number, denoted $b(n)$, is the number of partitions of the set $\{1, 2, \ldots, n\}$. Therefore, by definition, $b(n) = \sum_{m=1}^{n} S(n,m)$, for $n \geq 1$ and by convention (see Stirling Numbers), $b(0) = 1$. Thus, for $n \geq 1$,*

$$\begin{aligned} b(n) &= \sum_{m=1}^{n} S(n,m) = \sum_{m\geq 1} S(n,m) = \sum_{m\geq 1}\sum_{k=1}^{m} \frac{(-1)^{m-k}\, k^{n-1}}{(k-1)!\,(m-k)!} \\ &= \sum_{k\geq 1} \frac{k^n}{k!} \sum_{m\geq k} \frac{(-1)^{m-k}}{(m-k)!} = \frac{1}{e}\sum_{k\geq 1}\frac{k^n}{k!}. \end{aligned} \tag{5.5}$$

As, $b(n)$ has terms of the form $\frac{k^n}{k!}$, we compute its exponential generating function (see Exercise 2.1.20.9). Thus, if $B(x) = \sum_{n\geq 0} b(n)\frac{x^n}{n!}$ then

$$\begin{aligned} B(x) &= 1 + \sum_{n\geq 1} b(n)\frac{x^n}{n!} = 1 + \sum_{n\geq 1}\left(\frac{1}{e}\sum_{k\geq 1}\frac{k^n}{k!}\right)\frac{x^n}{n!} \\ &= 1 + \frac{1}{e}\sum_{k\geq 1}\frac{1}{k!}\sum_{n\geq 1} k^n\frac{x^n}{n!} = 1 + \frac{1}{e}\sum_{k\geq 1}\frac{1}{k!}\sum_{n\geq 1}\frac{(kx)^n}{n!} \\ &= 1 + \frac{1}{e}\sum_{k\geq 1}\frac{1}{k!}\left(e^{kx} - 1\right) = 1 + \frac{1}{e}\sum_{k\geq 1}\left(\frac{(e^x)^k}{k!} - \frac{1}{k!}\right) \\ &= 1 + \frac{1}{e}\left(e^{e^x} - 1 - (e-1)\right) = e^{e^x - 1}. \end{aligned} \tag{5.6}$$

Recall that e^{e^x-1} is a valid formal power series (see Remark 5.1.5). Now, let us derive the recurrence relation for $b(n)$'s. Taking the natural logarithm on both the sides of Equation (5.6), one has $Ln\left(\sum_{n\geq 0} b(n)\frac{x^n}{n!}\right) = e^x - 1$. Now, differentiation with respect to x gives $\frac{1}{\sum_{n\geq 0} b(n)\frac{x^n}{n!}} \cdot \sum_{n\geq 0} b(n)\frac{x^{n-1}}{(n-1)!} = e^x$. *Therefore, after cross multiplication and an multiplication with x, implies*

$$\sum_{n\geq 1}\frac{b(n)x^n}{(n-1)!} = xe^x\sum_{n\geq 0} b(n)\frac{x^n}{n!} = x\left(\sum_{m\geq 0}\frac{x^m}{m!}\right)\cdot\left(\sum_{n\geq 0} b(n)\frac{x^n}{n!}\right).$$

Thus,

$$\frac{b(n)}{(n-1)!} = [x^n]\sum_{n\geq 1}\frac{b(n)x^n}{(n-1)!} = [x^n]\, x\left(\sum_{m\geq 0}\frac{x^m}{m!}\right)\cdot\left(\sum_{n\geq 0} b(n)\frac{x^n}{n!}\right) = \sum_{m=0}^{n-1}\frac{1}{(n-1-m)!}\cdot\frac{b(m)}{m!}.$$

Hence, it follows that $b(n) = \sum_{m=0}^{n-1} \binom{n-1}{m} b(m)$, *for* $n \geq 1$ *with* $b(0) = 1$.

8. *Determine the number of ways of arranging n pairs of parentheses (left and right) such that at any stage the number of right parentheses is always less than or equal to the number of left parentheses.*
Solution: *Recall that this number equals* C_n, *the* n^{th} *Catalan number (see Page 44). Let us obtain a recurrence relation for these numbers and use it to get a formula for* C_n*'s.*

Let P_n *denote the arrangements of those n pairs of parentheses that satisfy "at any stage, the number of right parentheses is always less than or equal to the number of left parentheses". Also, let* Q_n *denote those elements of* P_n *for which, "at the* $2k$*-th stage, for* $k < n$, *the number of left parentheses is strictly greater than the number of right parentheses".*

We now claim that $Q_n = 1$ *and for* $n \geq 2$, $Q_n = P_{n-1}$.
Clearly $Q_1 = 1$. *Note that, for* $n \geq 2$, *any element of* Q_n, *necessarily starts with two left parentheses and ends with two right parentheses. So, if we remove the first left parenthesis and the last right parenthesis then one obtains as element of* P_{n-1}. *In a similar way, if we add one left parenthesis at the beginning and a right parenthesis at the end of an element of* P_{n-1}, *one obtains an element of* Q_n. *Hence,* $Q_n = P_{n-1}$.

Let $n \geq 2$ *and consider an element of* P_n. *Then, for some* k, $1 \leq k \leq n$, *the first* k *pairs of parentheses will form an element of* Q_k *and the remaining* $(n-k)$ *pairs of parenthesis will form an element of* P_{n-k}. *Hence, if we take* $P_0 = Q_1 = 1$, *one has* $P_n = \sum_{k=1}^{n} Q_k P_{n-k} = \sum_{k=1}^{n} P_{k-1} Q_{n-k}$, *for* $n \geq 2$. *Now, define* $P(x) = \sum_{n \geq 0} P_n x^n$. *Then*

$$\begin{aligned} P(x) &= \sum_{n\geq 0} P_n x^n = 1 + \sum_{n\geq 1} P_n x^n = 1 + \sum_{n\geq 1} \left(\sum_{k=1}^{n} P_{k-1} Q_{n-k} \right) x^n \\ &= 1 + x \left(\sum_{k\geq 1} P_{k-1} x^{k-1} \sum_{n\geq k} P_{n-k} x^{n-k} \right) = 1 + x \left(P(x) \sum_{k\geq 1} P_{k-1} x^{k-1} \right) \\ &= 1 + x \left(P(x) \right)^2 . \end{aligned}$$

Thus, $xP(x)^2 - P(x) + 1 = 0$ *and* $P(x) = \dfrac{1 \pm \sqrt{1-4x}}{2x}$. *Therefore, using* $P_0 = 1$, *one*

obtains $P(x) = \dfrac{1-\sqrt{1-4x}}{2x}$. *Hence,*

$$\begin{aligned} P_n &= [x^n]P(x) = \frac{1}{2}\cdot[x^{n+1}]\left(1-\sqrt{1-4x}\right) \\ &= -\frac{1}{2}\cdot\frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\left(\frac{1}{2}-2\right)\cdots\left(\frac{1}{2}-n\right)}{(n+1)!}(-4)^{n+1} \\ &= 2(-4)^n\cdot\frac{1\cdot(-1)\cdot(-3)\cdot(-5)\cdots(1-2n)}{2^{n+1}(n+1)!} = 2^n\frac{1\cdot3\cdot5\cdots(2n-1)}{(n+1)!} \\ &= \frac{1}{n+1}\binom{2n}{n}, \quad \textit{the } n^{th} \textit{ Catalan Number.} \end{aligned}$$

We end this section with a set of exercises.

Exercise 5.2.3. *1. In each of the following, obtain a recurrence relation for the numbers a_n's. Hence or otherwise, determine a formula for the a_n's.*

(a) Let a_n denote the number of ways of climbing a staircase in which "at each step, the climber climbs either one stair or two stairs.

(b) Let a_n denote the number of sequences of length n that consist of the digits $0, 1, 2, \ldots, 9$ that do not have two consecutive appearances of 9's.

(c) Let a_n denote the number of ways in which a person can spend n rupees to buy either a toffee worth 1 rupee or a chocolate worth 2 rupees or an ice-cream worth 3 rupees.

2. Let $f(n,k)$ denote the number of k-element subsets that can be selected from the set $\{1, 2, \ldots, n\}$ that do not contain two consecutive integers. Determine a recurrence relation for $f(n,k)$'s and hence determine the value of $f(n,k)$.

3. Suppose the numbers $\{1, 2, \ldots, n\}$ are arranged in a round table. Let $g(n,k)$ denote the number of k-element subsets that can be selected from this round table with the condition that no two consecutive integers appear. Find a recurrence relation for $g(n,k)$'s and hence determine the value of $g(n,k)$. [Hint: $g(n,k) = f(n-1,k) + f(n-3,k-1)$.]

4. Let $a(0) = a(1) = 1$ and let $a(n) = a(n-1) + (n-1)a(n-2)$, for $n \geq 2$. Compute the exponential generating function $A(x) = \sum\limits_{n\geq 0} a(n)\dfrac{x^n}{n!}$. Hence or otherwise, compute $a(n)$'s.

5. Let $a(0) = 1$ and let $\sum\limits_{k=0}^{n} a(k)a(n-k) = 1$, for all $n \geq 1$. Determine the sequence $a(n)$.

5.3 Applications to Generating Functions

The ideas learnt in the previous sections will be used to get closed form expressions for sums arising out of binomial coefficients. To do so, recall the list of formal power series that appear on Page 100.

Example 5.3.4. *1. Find a closed form expression for the numbers* $a(n) = \sum_{k\geq 0} \binom{k}{n-k}$.

Solution: *Define* $A(x) = \sum_{n\geq 0} a(n)x^n$. *Then*

$$\begin{aligned} A(x) &= \sum_{n\geq 0} a(n)x^n = \sum_{n\geq 0}\left(\sum_{k\geq 0}\binom{k}{n-k}\right)x^n = \sum_{k\geq 0}\left(\sum_{n\geq 0}\binom{k}{n-k}x^n\right) \\ &= \sum_{k\geq 0} x^k\left(\sum_{n\geq k}\binom{k}{n-k}x^{n-k}\right) = \sum_{k\geq 0} x^k(1+x)^k = \sum_{k\geq 0}\big(x(1+x)\big)^k = \frac{1}{1-x(1+x)}. \end{aligned}$$

Therefore, Example 5.2.1.2 implies $a(n) = [x^n]A(x) = [x^n]\dfrac{1}{1-x(1+x)} = F_n$, *the n-th Fibonacci number.*

2. Find a closed form expression for the polynomials $a(n,x) = \sum_{k=0}^{\lfloor \frac{n}{2}\rfloor}\binom{n-k}{k}(-1)^k x^{n-2k}$.

Solution: *Define* $A(x,y) = \sum_{n\geq 0} a(n,x)y^n$. *Then*

$$\begin{aligned} A(x,y) &= \sum_{n\geq 0} a(n,x)y^n = \sum_{n\geq 0}\left(\sum_{k=0}^{\lfloor \frac{n}{2}\rfloor}\binom{n-k}{k}(-1)^k x^{n-2k}\right)y^n \\ &= \sum_{k\geq 0}(-1)^k y^{2k}\left(\sum_{n\geq 2k}\binom{n-k}{k}(xy)^{n-2k}\right) \\ &= \sum_{k\geq 0}(-1)^k y^{2k}(xy)^{-k}\left(\sum_{t\geq k}\binom{t}{k}(xy)^t\right) \\ &= \sum_{k\geq 0}(-y^2)^k(xy)^{-k}\frac{(xy)^k}{(1-xy)^{k+1}} = \frac{1}{1-xy}\cdot\sum_{k\geq 0}\left(\frac{-y^2}{1-xy}\right)^k \\ &= \frac{1}{1-xy}\cdot\frac{1}{1-\frac{-y^2}{1-xy}} = \frac{1}{1-xy+y^2} = \frac{1}{(1-\alpha y)(1-\beta y)}, \end{aligned}$$

where $\alpha = \dfrac{x+\sqrt{x^2-4}}{2}$ *and* $\beta = \dfrac{x-\sqrt{x^2-4}}{2}$. *Thus,*

$$\begin{aligned} a(n,x) &= [y^n]A(x,y) = [y^n]\frac{1}{1-xy+y^2} = [y^n]\,\frac{1}{\alpha-\beta}\left(\frac{\alpha}{1-\alpha y} - \frac{\beta}{1-\beta y}\right) \\ &= \frac{1}{\alpha-\beta}\left(\alpha^{n+1}-\beta^{n+1}\right) \\ &= \frac{1}{\sqrt{x^2-4}}\left(\left(\frac{x+\sqrt{x^2-4}}{2}\right)^{n+1} - \left(\frac{x-\sqrt{x^2-4}}{2}\right)^{n+1}\right). \end{aligned}$$

Since α *and* β *are the roots of* $y^2 - xy + 1 = 0$, $\alpha^2 = \alpha x - 1$ *and* $\beta^2 = \beta x - 1$. *Therefore, verify that the* $a(n,x)$*'s satisfy the recurrence relation* $a(n,x) = xa(n-1,x) - a(n-2,x)$, *for* $n \geq 2$, *with initial conditions* $a(0,x) = 1$ *and* $a(1,x) = x$.

Let $A = (a_{ij})$ be an $n \times n$ matrix, with $a_{ij} = 1$, whenever $|i-j| = 1$ and 0, otherwise. Then A is an adjacency matrix of a tree T on n vertices, say $1, 2, \ldots, n$ with the vertex i being adjacent to $i+1$, for $1 \le i \le n-1$. It can be verified that if $a(n,x) = \det(xI_n - A)$, the characteristic polynomial of A, then $a(n,x)$'s satisfy the above recurrence relation. The polynomials $a(n,2x)$'s are also known as CHEBYSHEV'S *polynomial of second kind.*

Let us now substitute different values for x to obtain different expressions and then use them to get binomial identities.

(a) *Let $x = z + \frac{1}{z}$. Then $\sqrt{x^2-4} = z - \frac{1}{z}$ and we obtain $a(n, z+\frac{1}{z}) = \frac{z^{2n+2}-1}{(z^2-1)z^n}$. Hence, we have $\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n-k}{k}(-1)^k \left(z+\frac{1}{z}\right)^{n-2k} = \frac{z^{2n+2}-1}{(z^2-1)z^n}$. Or equivalently,*

$$\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n-k}{k}(-1)^k \left(z^2+1\right)^{n-2k} z^{2k} = \frac{z^{2n+2}-1}{z^2-1}.$$

(b) *Writing x in place of z^2, we obtain the following identity.*

$$\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n-k}{k}(-1)^k (x+1)^{n-2k} x^k = \frac{x^{n+1}-1}{x-1} = \sum_{k=0}^{n} x^k. \tag{5.7}$$

(c) *Hence, equating the coefficient of x^m in (5.7) gives the identity*

$$\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n-k}{k}\binom{n-2k}{m-k} = \begin{cases} 1, & \text{if } 0 \le m \le n; \\ 0, & \text{otherwise.} \end{cases}$$

(d) *Substituting $x = 1$ in (5.7) gives $\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n-k}{k} 2^{n-2k} = n+1$.*

3. *Determine the generating function for the numbers $a(n,y) = \sum_{k \ge 0} \binom{n+k}{2k} y^k$.*

Solution: *Define $A(x,y) = \sum_{n \ge 0} a(n,y)x^n$. Then*

$$\begin{aligned} A(x,y) &= \sum_{n \ge 0} \left(\sum_{k \ge 0} \binom{n+k}{2k} y^k \right) x^n = \sum_{k \ge 0} \left(\frac{y}{x}\right)^k \sum_{n \ge k} \binom{n+k}{2k} x^{n+k} \\ &= \sum_{k \ge 0} \left(\frac{y}{x}\right)^k \frac{x^{2k}}{(1-x)^{2k+1}} = \frac{1}{1-x} \sum_{k \ge 0} \left(\frac{yx}{(1-x)^2} \right)^k \\ &= \frac{1-x}{(1-x)^2 - xy}. \end{aligned}$$

(a) *Verify that if we substituting $y = -2$ then*

$$\sum_{k \ge 0} \binom{n+k}{2k} (-2)^k = [x^n]\, A(x,-2) = [x^n]\, \frac{1-x}{(1+x^2)} = (-1)^{\lceil n/2 \rceil}.$$

(b) Verify that if we substituting $y = -4$ then

$$\sum_{k\geq 0}\binom{n+k}{2k}(-4)^k = [x^n]\ A(x,-4) = [x^n]\ \frac{1-x}{(1+x)^2} = (-1)^n(2n+1).$$

(c) Let $f(n) = \sum_{k\geq 0}\binom{n+k}{2k}2^{n-k}$ and let $F(z) = \sum_{n\geq 0} f(n)z^n$. Then verify that

$$F(z) = A(2z, \frac{1}{2}) = \frac{1-2z}{(1-z)(1-4z)} = \frac{2}{3}\cdot\frac{1}{1-4z} + \frac{1}{3}\cdot\frac{1}{1-z}.$$

Hence, $f(n) = [z^n]F(z) = \dfrac{2\cdot 4^n}{3} + \dfrac{1}{3} = \dfrac{2^{2n+1}+1}{3}$.

We end this chapter with the following set of exercises.

Exercise 5.3.5. *1. Let n be a non-negative integer and define $a(n,y) = \sum_{k\geq 0}\binom{n}{k}\binom{2k}{k}y^k$ and $A(x,y) = \sum_{n\geq 0} a(n,y)x^n$. Then prove that $A(x,y) = \dfrac{1}{\sqrt{1-x}}\cdot\dfrac{1}{\sqrt{1-x-4xy}}$. Hence, respectively replace y with $\dfrac{-1}{4}$ and $\dfrac{-1}{2}$, to obtain the following identities:*

(a) $\sum_{k\geq 0}\binom{n}{k}\binom{2k}{k}(-4)^{n-k} = \binom{2n}{n}$.

(b) $\sum_{k\geq 0}\binom{n}{k}\binom{2k}{k}(-2)^{n-k} = \begin{cases}\binom{n}{n/2}, & \text{if } n \text{ is even;}\\ 0, & \text{otherwise.}\end{cases}$
This identity is popularly known as the (Reed Dawson's Identity)

2. Let $m, n \in \mathbb{N}$. Then prove that $\sum_{k\geq 0}\binom{n+k}{m+2k}\binom{2k}{k}\dfrac{(-1)^k}{k+1} = \binom{n-1}{m-1}$.

3. Fix positive integers $m, n \in \mathbb{N}$ and define $a(n,y) = \sum_{k\geq 0}\binom{m}{k}\binom{n+k}{k}y^k$, $b(n,y) = \sum_{k\geq 0}\binom{m}{k}\binom{n}{k}y^k$ and $c(n,y) = \sum_{k\geq 0}\binom{m}{k}\binom{m+n-k}{k}y^k$. Compute $A(x,y)\sum_{n\geq 0} a(n,y)x^n$, $B(x,y) = \sum_{n\geq 0} b(n,y)x^n$ and $C(x,y) = \sum_{n\geq 0} c(n,y)x^n$ to show the following:

(a) $\sum_{k\geq 0}\binom{m}{k}\binom{n+k}{m} = \sum_{k\geq 0}\binom{m}{k}\binom{n}{k}2^k = \sum_{k\geq 0}\binom{m}{k}\cdot\binom{m+n-k}{m}$.

(b) $\sum_{k\geq 0}\binom{m}{k}\binom{n}{k}(-1)^k = \sum_{k\geq 0}\binom{m}{k}\cdot\binom{m+n-k}{m}(-2)^k$.

(c) $\binom{m+n}{n} = [x^n]\ (1+x)^{m+n} = [x^n]\ (1+x)^m(x+1)^n = \sum_{k\geq 0}\binom{m}{k}\binom{n}{k} = [x^n]\ \dfrac{1}{(1-x)^{m+1}}$.

(d) $\sum_{k\geq 0}\binom{m}{k}\cdot\binom{m+n-k}{m}(-1)^k = 1$.

(e) $\sum_{k\geq 0}\binom{m}{k}\cdot\binom{n+k}{m}(-1)^k = (-1)^m$.

4. Prove that $\sum_{k\geq 0}\binom{m+1}{k}\binom{m+n-k}{m}(-1)^k = \begin{cases}1, & \text{if } n = 0,\\ 0, & \text{if } n > 0.\end{cases}$

5. *Prove that* $\sum_{k\geq 1}(-1)^k\binom{n}{k}\binom{k-1}{m} = (-1)^{m+1}$.

6. *Determine whether or not the following statement is correct.*

$$\sum_{k=1}^{n}\frac{(-1)^{k+1}}{k}\binom{n}{k} = \sum_{k=1}^{n}\frac{1}{k}.$$

7. *Determine the exponential generating function for the numbers* $a(n)$*'s that appeared in Exercise 14.*

Notes: Most of the ideas for this chapter have come from book [8].

Bibliography

[1] J. Cofman, "Catalan Numbers for the Classroom?", *Elem. Math.*, 52 (1997), 108 - 117.

[2] D. I. A. Cohen, *Basic Techniques of Combinatorial Theory*, John Wiley and Sons, New York, 1978.

[3] William Dunham, *Euler, The Master of Us All*, Published and Distributed by The Mathematical Association of America, 1999.

[4] G. E. Martin, *Counting: The Art of Enumerative COmbinatorics*, Undergraduate Texts in Mathematics, Springer, 2001.

[5] R. Merris, *Combinatorics,* 2^{th} *edition*, Wiley-Interscience, 2003.

[6] J. Riordan, *Introduction to Combinatorial Analysis*, John Wiley and Sons, New York, 1958.

[7] R. P. Stanley, *Enumerative Combinatorics, vol. 2*, Cambridge University Press, 1999.

[8] H. S. Wilf, *Generatingfunctionology*, Academic Press, 1990.