TEXTS AND READINGS
IN MATHEMATICS 64

Combinatorial Techniques

Sharad S. Sane

HINDUSTAN
BOOK AGENCY

TEXTS AND READINGS IN MATHEMATICS **64**

Combinatorial Techniques

Combinatorial Techniques

Sharad S. Sane
Indian Institute of Technology
Mumbai

HINDUSTAN
BOOK AGENCY

Published by

Hindustan Book Agency (India)
P 19 Green Park Extension
New Delhi 110 016
India

email: info@hindbook.com
www.hindbook.com

ISBN 978-93-80250-48-9 ISBN 978-93-86279-55-2 (eBook)
DOI 10.1007/978-93-86279-55-2

Paper cover Edition, 2016

ISBN 978-93-80250-84-7

To my wife Suneeta

Contents

Preface

The idea of writing a text book on combinatorics has been on my mind for a very long time. The question was that of judgment in deciding which topics are to be considered basic and must be included as also the organization of the material. I hope that the reader will find both the content and the organization of material in this book sufficiently interesting and coherent.

Whether combinatorics consists of merely techniques (read, a slightly derogatory, tricks) or also encompasses a sequence of theorems that can properly be perceived to be purporting some deep theory is not clear. The truth lies somewhere in between. To a section of the mathematical community, combinatorics is nothing more than a pastime of solving puzzles. Cleverness is an acknowledged essential ingredient of creating all mathematics. That appears to be more so in case of combinatorics where cleverness is seemingly the only tool required to achieve the goal. However, "Mathematics is not cleverness" declares the objecting section of the mathematical community. When a mathematician, working in some other area of mathematics uses the expression, "that is combinatorial", it sometimes means clever but at times also means that it amounts to messy computations devoid of any theory and an ennui inducing exercise. It is the author's intention to show to the reader that combinatorial techniques can be well studied and that these techniques, are far more systematic and sophisticated than the puzzles and tricks that they engulf and encompass.

In terms of the material presented here, I have loosely followed the contents of the texts by Brualdi, Liu, Krishnamurty and Cohen [13, 36, 35, 18]. In terms of style as well as contents, I am highly impressed by the book on combinatorics by van Lint and Wilson [57]. A large part of modern combinatorics seems to have its origin in the gambling problems of the last few centuries European society, particularly in the work of Laplace and de Moivre. Questions in probability, therefore, form a right setting for combinatorial problems, both in terms of understanding and historical perspective. An algebraization of the discipline was obtained in the concept of generating functions championed by Euler, the founder of modern combinatorics. This paved a way for the systematic use of algebra in combinatorics. To many people, algebra itself is a discrete discipline. Algebra has played such a major role in modern combinatorics that the present state of knowledge and direction in combinatorics make it appear as if it were all the time a subtopic of algebra. This, to me, appears a serious shortcoming on the pedagogical aspects of the combinatorial discipline. It is not my intention to suggest

here that the role of algebra has been overestimated. On the contrary, algebra makes things very systematic and smooth. My assertion mainly pertains to the pedagogy of combinatorics, where, I believe that overuse of algebraic language pushes the reader into a jungle of symbols and the sophisticated write-up drives him away rather than inviting him to the plain cleverness of combinatorics. This also amounts to losing some historical perspective. A lack of knowledge of the framework of discrete probability on the part of a mathematician and more importantly on the part of a combinatorialist is a serious shortcoming. I have been driven by the article of Mumford [43] where a strong appeal in favour of stochasticness is made. Wherever possible, the book will also try to connect the material under consideration with other areas of mathematics particularly number theory, analysis and topology. The rigid distinction between pure and applied (or applicable) mathematics as well as the distinction between discrete and continuous mathematics is fast vanishing and in this regard, I am highly impressed by Concrete Mathematics of Knuth, Graham Knuth and Patashnik [27].

I have been associated with the Mathematics Olympiad activity in India for more than two decades. That has certainly influenced my choice of problems and exercises, some of which are borrowed from the Olympiad contests including the International Mathematical Olympiad, the IMO. However, the major contribution to the contents of this book has come from teaching the combinatorics course at the University of Mumbai besides at the University of Florida, Central Michigan University and the Michigan Technological University. Besides the books mentioned in the earlier paragraphs, I have been impressed by the article on Polya theory by de Bruijn, [4] and the essays in [33].

Every chapter discusses a famous and important result in Combinatorics which demonstrates the tremendous power of some of the very simple ideas. Organization of chapters in the book is as follows. The first six chapters could form a semester course at an undergraduate level. These include the basics of counting parameters (Chapter 1), listing combinatorial objects (Chapter 2), combinatorics of permutations (Chapter 3) and the basic inclusion exclusion principle (Chapter 4). Chapter 5 and Chapter 6 deal with probability and random variables respectively. Chapters 7, 8 and 9 have material for the Olympiad level audience and contain a large number of exercises. Many situations of the occurrence of parity arguments in combinatorics are discussed in Chapter 7. This chapter also includes the Gauss quadratic reciprocity law. Chapter 8 is on pigeonhole principle and after discussing Ramsey theory, this chapter also includes various Erdös-Szekeres theorems that use pigeonhole principle in some form. Chapter 9 deals with geometric results that have combinatorial flavour and includes the Euler equation, classification of regular polytopes, tilings and Sperner's result on triangulations. The first nine chapters can form a semester course at a slightly advanced level. Chapter 10 deals with Stirling and Catalan numbers. Chapter 11 is on recurrence relations and Chapter 12 deals with generating functions. Besides standard material on generating functions, this chapter discusses at some length the coin exchange theorem and the Dirichlet generating functions. I have separated the partition theory of integers by making it an independent Chapter 13. This chapter ends with the Euler pentagonal theorem and the material here should be useful to people interested

in combinatorial number theory. Chapter 14 can be viewed as a forerunner to Chapter 15 but it can also be of independent interest since it shows the use of group action as a major tool in finite group theory. This chapter includes the class equation, Sylow theorems, automorphisms of symmetric groups and the classification of finite subgroups of the orthogonal groups in 2 and 3 dimensions. Chapter 15 deals with Polya's theory of enumeration where a large number of examples are discussed and it also includes de Bruijn's generalization of Polya theory. Chapter 16 deals with the systems of distinct representatives and includes the Birkhoff von Neumann theorem on doubly stochastic matrices and Dilworth's theorem on posets. Finally, I would like to emphasize that wherever possible, I have tried to draw a comparison between the combinatorics of infinite and finite. This is particularly visible in Chapters 7 and 9 where an idea of infinite version of Ramsey results and the Euclidean Ramsey theory respectively are discussed as well in the last chapter 16 where the Rado selection principle is discussed.

An elementary first level course could include the first eight chapters in that order. Chapters 3 through 11 are more suited for Mathematical Olympiad students. A course with emphasis on generating functions and recurrence relations should follow Chapters 3, 4, 10, 11, 12 and 13. Advanced Chapters in the text are Chapters 12 through 16. The author believes that the strength of the text also lies in the very large number of exercises at the end of each chapter. The exercises are at various levels of difficulty and range from very simple to more advanced such as Euler convergence (Exercise 6.31) and Conway's soldiers in the desert (Exercise 11.45).

I trust that the book will prove useful and interesting to a wide range of mathematical and non-mathematical community.

Sharad S. Sane

Acknowledgements

I thank all the referees of the manuscript of the text for their careful reading and detailed comments that led to considerable improvement in the final form of the book. It is a pleasure to thank a number of senior colleagues in the Mathematics Department of Mumbai University. They include M.G.Nadkarni, Nirmala Limaye, Prafullata Chawathe and Anjana Wirmani-Prasad. Special thanks to R.C.Cowsik. Discussing mathematics with all these people has been a pleasure. I also wish to thank Rajendra Pawale, Santosh Shende and Anand Kumtha, who made careful reading of parts of the manuscript of the book. University Mathematics Department library was the main source of material and I have extensively used it in the book writing project. I wish to thank colleagues C.R. Pranesachar and B.J.Venkatachala of the Mathematics Olympiad cell at the Indian Institute of Science, Bangalore for helpful discussions.

Some friends have been instrumental in pushing me to complete the book writing project and I wish to thank them. They include Ravindra Bapat of the Indian Statistical Institute, Delhi, S.A. Katre of Pune University and Sham Navathe of Computer Science Department at Georgia Tech. Sham Navathe had warned me decades ago that ninety percent of text book writing projects that authors have on their minds never see the completion. Special thanks are due to Jet Wimp for making a copy of his review [62] in the Mathematical Intelligencer (excerpts used in Chapter 8) available to me.

My father was an embodiment of hard work and perseverance and my mother that of cleverness; both would have liked to see the completion of the book writing project.

Sharad S. Sane

Chapter 1

Basic counting

1.1 Introduction

A large part of combinatorics is concerned with counting. As such this is not a very difficult activity, at least in principle. However, lack of clarity can very much make the counting obscure. The material covered in this chapter forms a basis for all the other chapters in this book because it sets the basic counting parameters required throughout the book. We begin by introducing the following elementary principles.

A man wants to travel from place A to place B either by a bus or by a train. He knows that there are five different buses he can choose from and three different trains that he take to go from A to B. Obviously then, there are eight different ways in which he can travel. If there are three different sized apples and two different sized mangoes on a table, then the number of ways of picking up one fruit among these is $3 + 2 = 5$. We have:

Addition Principle: If the first box contains m objects and the second box contains n objects, then the number of ways of choosing one object from either of the two boxes is $m + n$.

In a purely set-theoretic language, the addition principle tells us about the order of the union of two disjoint sets X and Y, if we know the number of objects in each of them. The main point to note here is that the sets must be disjoint. For example, The order of the set of numbers less than 30 that are either prime or perfect squares is obtained by finding the possibilities of the occurrences of the two events separately and then adding them and hence the answer is $10 + 5 = 15$. However if we wish to find the order of the set of numbers below 30 that are divisible by 2 or 3, the required number is not $14 + 9$.

Multiplication Principle: If the first box contains m objects and the second box contains n objects, then the number of ways of choosing a pair of objects, the first from the first box and the second from the second box is mn.

While the addition principle gives union of two disjoint sets, the multiplication principle gives the (order of) the Cartesian product of two sets. For example, suppose that in order to go from A to B, one must pass through C. If there are five ways of going

from A to C and three ways of going from C to B, then there are fifteen ways of going from A to B.

There is a slightly more profound but equally basic technique that is applied in combinatorics. Suppose in a classroom where the students occupy seats on the benches arranged in m rows and n columns, not all the seats on all the benches are filled. If you wish to count the total number of students in the class, then there are two ways of doing it. You could fix a row and carefully count the number of students in that row. Do this for each row and then sum over all the rows. You could then ask your friend to do the same thing fixing a column first and then summing over all the columns (do not forget to count yourself if you are occupying a seat on some bench!). The two numbers must be equal. This elementary observation leads to:

Two-Way Counting : Let S be a subset of the Cartesian product of two sets A and B. Let, for a in A, R_a denote the subset $\{y \in B : (a, y) \in S\}$. Similarly, for b in B, let C_b denote the subset $\{x \in A : (x, b) \in S\}$. Then

$$\sum_{a \in A} |R_a| = \sum_{b \in B} |C_b|$$

All throughout this book, we denote the *order or cardinality* of a set T by $|T|$. Since most of the sets we deal with are *finite* the word *size* will also be used to denote this number. Also, the Cartesian product of two sets X and Y (denoted by $X \times Y$) consists of all the pairs (x, y) for which x is in X and y is in Y.The innocuous technique of two-way counting has a large number of applications in branches of Combinatorics such as *graph theory and design theory*. We illustrate this with an application.

Example 1.1.1. A graph G is a pair (V, E) where V is the set of vertices and E the set of edges. An edge is an unordered pair of vertices. We denote an edge e by the pair (xy) or (yx) with the understanding that the edge e is an edge between the two vertices x and y. In that case, we say that e is incident with x (and also with y). An edge of the form (xx) is called a loop . If we have two edges e and e' such that both are equal to (xy), then e, e' are called double edges (or multiple edges, in general) between x and y. For x in V, $d(x)$, the degree of the vertex , is simply the number of edges (xy) with y in V. Here, we count a loop (xx) twice and hence it contributes two to the degree of x. In the first part in Figure 1.1, we have a loop at x, two loops at y and a double edge between x and z. A graph G is called a *simple graph* if it has no loops or multiple edges. The second part in Figure 1.1 is a simple graph. In the first graph, the degrees of x, y, z, w are $6, 5, 2, 1$ respectively while the second graph has 5 vertices of degree 2 each and one vertex with degree 0.

If we sum over all the degrees in the graph G, then the result must be an even number since this simply counts each edge two times. We thus can not have a graph on 7 vertices with each vertex of degree 3.

We use the term an n-set to mean a set of order n. Likewise a k-subset will mean a subset of order k (similar term also applies to a superset). A *k-permutation* of a set S is an *ordered* k-tuple of elements of S. Thus a k-permutation is a sequence $(x_1, x_2, \ldots, x_k)$ where the k elements are all different (and come from S) and with

an understanding that $(x_2, x_1, \ldots, x_k)$ is not the same as $(x_1, x_2, \ldots, x_k)$. How many k-permutations does an n-set have? The first element can be chosen in n ways since any of the n elements can be the first element. Having chosen the first element, the second can be chosen in $n-1$ ways (out of the remaining $n-1$ elements). Proceeding in this manner, we have the last, i.e., the k-th element chosen in $n-(k-1)$ ways.

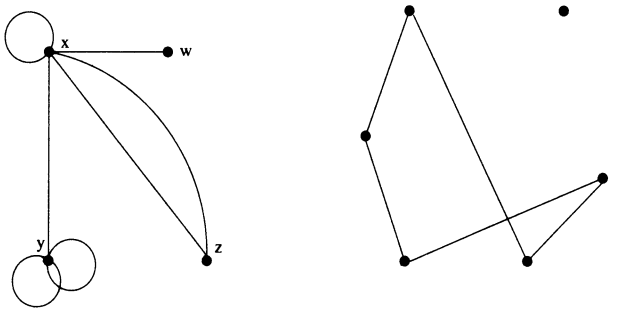

Figure 1.1: A graph and a simple graph

Lemma 1.1.2. *The number of k-permutations of an n-set is given by:*

$$P(n,k) = n(n-1)\cdots(n-k+1)$$

A slight trouble with the argument is what happens if k exceeds n. The answer then should be zero, which is what it obtains in the formula given in Lemma 1.1.2. The most important case (in Lemma 1.1.2) occurs here when k equals n. In that case we call an n-permutation of an n-set simply a permutation (of an n-set S) and $P(n,n)$ is denoted by $n!$ (pronounced n factorial; the exclamation is probably due to the fact that $n!$ is very large even for moderate values of n) which is equal to the product of all the integers from 1 to n.

Example 1.1.3. If a class has 15 students and if 5 prizes (first, second, third, fourth and fifth) are to be given to some 5 students among them, then the number of ways of doing this is $P(15,5)$ which equals $3,603,600$. However, if the class has 10 boys and 5 girls, then the number of ways in which three prizes are to be separately given to boys and two prizes separately given to girls is equal to $P(10,3) \times P(5,2) = 14,400$ by using the multiplication principle.

Notice that a permutation is an ordered selection of objects (from a set). In a similar manner, an *unordered* selection of k objects from a set of n objects is called a *k-combination*. We denote the number of k-combinations of an n-set by $C(n,k)$. If n is a positive integer, then $C(n,0)$ equals 1 and $C(n,k)$ is 0 for all $k > n$. Since a set, by definition, is an unordered collection, a k-combination of a set is simply a k-subset of the given n-set.

Example 1.1.4. Consider the problem in 1.1.3 and suppose that we are merely interested in selecting three boys and two girls to distribute the prizes (without ranking them). Then the required number is $C(10,3) \times C(5,2) = 1,200$.

Lemma 1.1.5. *Let n be a positive integer and let k be a non-negative integer. Then*

$$C(n,k) = \frac{P(n,k)}{P(k,k)} = \frac{n(n-1)\cdots(n-k+1)}{k!}$$

Proof Make a two-way counting. How many k-permutations does an n-set have? We first select a k-subset in $C(n,k)$ ways and then order (permute) the elements of the chosen k-subset in $P(k,k)$ ways. But this number is just $P(n,k)$ which obtains the required formula using Lemma 1.1.2. Also, the statement trivially holds if $k > n$. □

A large number of problems of (elementary) combinatorics are solved by determining whether the required answer is a combination or a permutation. Since $n! = n \times (n-1) \times \cdots \times 2 \times 1$, we write $[n]_k$ to denote $P(n,k) = n(n-1)\cdots(n-k+1)$ and call it the *falling factorial*. By convention, $0! = 1$. We then have:

$$C(n,k) = \frac{[n]_k}{k!} = \frac{[n]_k(n-k)!}{k!(n-k)!} = \frac{n!}{k!(n-k)!}$$

Observe the arithmetical symmetry between k and $n-k$ in the above formula. This, of course, is a reflection of the fact that $C(n,k)$ and $C(n,n-k)$ are the same numbers:

Lemma 1.1.6. *If k is a non-negative integer and n a positive integer such that $k \leq n$, then $C(n,k) = C(n,n-k)$.*

Proof Let S be an n-set. In choosing a k-subset A of S, we are also rejecting an $(n-k)$-subset of S, namely the $(n-k)$-subset B which is the complement of A in S. It is like this. Any time you select k elements of S, ask your friend to select the remaining $n-k$ elements of S. The process is reversible. If your friend picks up $n-k$ elements, then you pick up the remaining k elements. This sets up a bijection (one-to one correspondence) between the set of k-subsets (that you select) and the set of $n-k$-subsets (that your friend selects). Hence we have $C(n,k) = C(n,n-k)$. □

1.2 Bijections

If we have an alphabet consisting of only three letters a, b and c and we wish to form words of length 5 using the three letters of our alphabet, then the number of ways of doing this is $3 \times 3 \times 3 \times 3 \times 3 = 3^5 = 243$. Similarly the number of binary sequences (i.e., sequences of 0's and 1's) of length n is 2^n. Interestingly, this also indirectly counts the total number of subsets of a set of order n. Let $T = \{1,2,\ldots,n\}$. *Throughout this book, we use the symbol $[n]$ to mean the set $\{1,2,\ldots,n\}$.* For each subset X of T associate a binary sequence b_X of length n where b_X has 1 in the position i if i is in X and has 0 in the i-th position if i is not in X. For example, if $n = 5$, and if X, Y, Z are respectively the empty set, the set $\{2,3,5\}$, and the set $\{1,2,3,4,5\}$, then the corresponding sequences b_X, b_Y and b_Z are 00000, 01101 and 11111 respectively.

This procedure obtains a pairing of the members of the set of all the subsets of T with the members of the set consisting of all the binary sequences of length n. Since the latter set has size 2^n, we readily have:

Corollary 1.2.1. *The total number of subsets of a set of order n is 2^n.*

As the discussion preceding Corollary 1.2.1 shows, more is true than what is stated in Corollary 1.2.1. Namely, we actually have an explicit bijection between the two sets: The set of all the subsets of an n-set and the set of all the binary sequences of length n. The proof technique used here is called *proof by bijection*. This is among the most powerful and elementary techniques of combinatorics and a substantial part of combinatorial activity today is devoted to finding ingenious bijections to count the number of objects under certain stipulations. We recall that a bijection is an injective (one-to-one) and surjective (onto) function. An interested reader is encouraged to look up the book by Stanton and White [50], where a large number of results are proved using bijections. One important point to observe here is that the *proofs that use bijection are genuinely combinatorial in nature.* An example of this is furnished in the proof of Lemma 1.1.6. We content ourselves with giving some obvious bijections.

Theorem 1.2.2. *Let m and n be positive integers. Then there exists a bijection between any two of the following sets.*

(a) The set of all the functions from an n-set to an m-set.

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

(c) The set of n-tuples (sequences) with entries from an m-set.

(d) The set consisting of all the ways of distributing n distinct objects into m distinct boxes (or cells).

In each case, the cardinality of the set in question is m^n.

Proof Let $D = \{1, 2, \ldots, n\}$ be the n-set and let $R = \{r_1, r_2, \ldots, r_m\}$. Given any function f from D to R, we write down the sequence $(f(1), f(2), \ldots, f(n))$ of length n with entries from R. This process is reversible and clearly sets a bijection between the set in (a) with those in (b) or (c). For (d), let the m distinct boxes be denoted by $B_1, B_2, \ldots, B_m$ and let the n distinct objects be denoted by the elements of D. Given a function f from D to R, put object i in the box B_j if and only if $f(i) = r_j$. This obtains the required bijection. □

Recall that an injective function is a function f for which $f(a) = f(b)$ implies $a = b$.

Theorem 1.2.3. *Let m and n be positive integers. Then there exists a bijection between any two of the following sets.*

(a) The set of all the injective functions from an n-set to an m-set.

(b) The set of words of length n on an alphabet of m letters with the condition that the word consists of distinct letters.

(c) *The set of n-tuples (sequences) with distinct entries from an m-set.*

(d) *The set consisting of all the ways of distributing n distinct objects into m distinct boxes (or cells) with the condition that no box holds more than one object.*

(e) *The set of all the n-permutations of an m-set.*

In each case, the cardinality of the set in question is $P(m,n) = [m]_n$.

Proof Let $D = \{d_1, d_2, \ldots, d_n\}$ and let R be the set of integers $1, 2, \ldots, m$. Any n-permutation of R say, $(a_1, a_2, \ldots, a_n)$ may be viewed as an injective function from D to R which sends d_1 to a_1, d_2 to a_2, $\cdots$, d_n to a_n (note that a_i's are distinct and are ordered). This sets up a bijection between the sets in (e) and (a). For (b), we may think of an alphabet precisely consisting of the elements of R. Since (b) stipulates that words must consist of distinct letters, we have a bijection between the set in (a) and the set in (b). The bijection between (c) and (a) is similar and is left to the reader. For (d), treat the elements of D as the objects, and let $B_1, B_2, \ldots, B_m$ be the set of m (distinct) boxes. Given any permutation, i.e., a function say $(a_1, a_2, \ldots, a_n)$ as above, put the object d_i in the k-th box B_k iff $a_i = k$. *All throughout this book, the term iff is used to mean "if and only if".* It is easy to check that this sets a bijection between the set in (d) and the one in (a) (or (e)). □

Definition 1.2.4. For a real number α and any non-negative integer k,

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}$$

Pronounce $\binom{\alpha}{k}$ as α choose k. By convention, we let $\binom{\alpha}{0}$ equal 1. Thus, if α is a positive integer, then $\binom{\alpha}{k}$ is same as $C(\alpha, k)$. As an example,

$$\binom{-1/2}{3} = \frac{-1/2 \times -3/2 \times -5/2}{6} = -\frac{3 \times 5}{3 \times 2^4}$$

equals $\frac{-5}{16}$. For reasons that will become clear after we prove the following theorem (Theorem 1.2.5), *the numbers $\binom{\alpha}{k}$ are called binomial coefficients. The two-line notation given in Definition 1.2.4 that we follow throughout this book is the modern notation and the reader is strongly urged to follow it.* At school level, one encounters expressions such as $(x+y)^2 = x^2+2xy+y^2$ and $(x+y)^3 = x^3+3x^2y+3xy^2+y^3$. Observe that the right hand sides of these equations are homogeneous polynomials in x and y, i.e., polynomials in which in each summand, the exponents of x and y add to a fixed number. This fixed number is called the degree of the homogeneous polynomial. We have the following theorem which dates back to at least the medieval times and was known to various civilizations.

Theorem 1.2.5. *(The Binomial Theorem) Let n be a non-negative integer. Then:*

$$(x+y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + + \binom{n}{n}y^n$$

Proof We first apply the multiplication principle. Indeed, $(x + y)^n$ is simply $(x + y)(x + y) \cdots (x + y)$, n times. Multiplying out all the terms amounts to doing the following. There are n boxes with each box containing both x and y. We choose x's from some boxes and y's from the remaining. If we choose x's from $n - k$ boxes and y's from the remaining k boxes, then that obtains a summand $x^{n-k}y^k$. Thus we already know that the resulting polynomial is homogeneous of degree n. It just remains to find the coefficient of the term $x^{n-k}y^k$. Obviously, we have to choose $n - k$ boxes for x (and hence the remaining k boxes for y). Since there are n boxes in all, the coefficient of $x^{n-k}y^k$ must be $\binom{n}{k}$. This completes the proof of the binomial theorem. □

For a brief history of the binomial theorem, refer to Knuth [34]. In the statement of the binomial theorem (Theorem 1.2.5), we did not specify as to what x and y are. In some sense, the reason for choosing to do so is that *it does not matter*. In other words given the usual laws of addition and multiplication (of natural numbers), the binomial theorem (Theorem 1.2.5) is an always true statement. This is expressed in the mathematical parlance by saying that the binomial theorem is a 'formal identity' (or a combinatorial identity). This, in particular, means that the statement of binomial theorem is true if we let the variables (x and y) to be any real or complex numbers. But still more important is the fact that the theorem is a formal identity. To be mathematically more precise, the binomial theorem holds over any commutative ring with identity. All through combinatorics, we shall have occasions to meet many such formal identities, a combinatorial theory of which will be reasonably formulated and formalized in Chapter 12 on generating functions.An uninteresting proof of the binomial theorem (Theorem 1.2.5)involves induction on n via the use of Pascal identity.

Theorem 1.2.6. *(Pascal identity) Let n and k be positive integers. Then*

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

Proof Let X be an n-set and fix an element x of X. and let Y denote the set $X - \{x\}$. For any k-combination (i.e. a k-subset) A of X, either x is in A or x is not in A. In the first case if B is the set $A - \{x\}$, then B is a $(k - 1)$-subset of Y and hence can be chosen in $\binom{n-1}{k-1}$ ways, while in the second case, A is itself a k-subset of Y and hence can be chosen in $\binom{n-1}{k}$ ways. The proof is complete by invoking the addition principle. □

It is also easy to see that Pascal's identity (Theorem 1.2.6) follows from the binomial theorem (Theorem 1.2.5) (if we have proved the latter without using the former as we did). Pascal's identity gives rise to the famous Pascal triangle, initial portion of which is drawn below. Each entry is obtained from the two entries directly above it as given Pascal's identity. This obtains all the binomial coefficients $\binom{n}{k}$, where n runs from 1 to 6 and the horizontal lines correspond to a fixed value of n. Note that Pascal triangle is an infinite triangle; only a finite portion of this triangle (from 1 to 6) is shown in

Figure 1.2. A more familiar form of the binomial theorem (Theorem 1.2.5) is:

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$$

which is obtained by letting $y = 1$ in the binomial theorem. By making the substitution $x = 1$ in the above expression, we obtain:

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

Stated in other words, the above identity tells us that a set of order n has 2^n subsets in all. We have already seen this in Corollary 1.2.1. Again, the substitution $x = -1$ yields

$$\sum_{k} \binom{n}{2k} = \sum_{k} \binom{n}{2k+1}$$

Thus, in any set the number of subsets of odd order is the same as the number of subsets of even order.

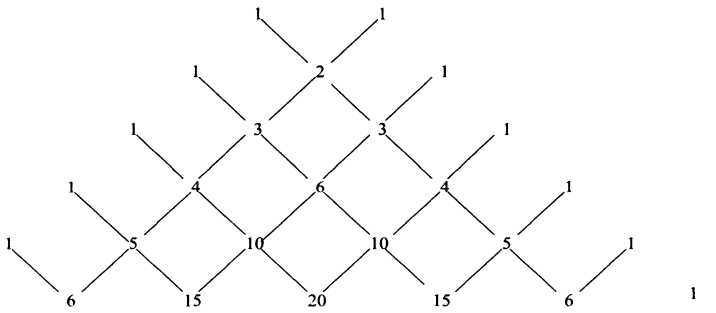

Figure 1.2: Pascal Triangle

A large number of identities involving binomial coefficients are actually proved either using a known combinatorial identity or a known polynomial expression such as the binomial theorem (Theorem 1.2.5) and manipulating it. For example, to prove that

$$\binom{n}{1} - 2\binom{n}{2} + 3\binom{n}{3} - 4\binom{n}{4} + \cdots + (-1)^{n-1} n \binom{n}{n} = 0$$

we observe that each summand on the left hand side (ignoring the sign) has the form

$$\begin{aligned} k\binom{n}{k} &= k\frac{n}{k}\binom{n-1}{k-1} \\ &= n\binom{n-1}{k-1} \end{aligned}$$

Hence the left hand side reduces to the alternating sum

$$n \times \sum_{i=0}^{n-1} (-1)^i \binom{n-1}{i} = 0$$

using the binomial theorem (Theorem 1.2.5). Similarly, to find $n \sum \frac{1}{k+1}\binom{n}{k}$, we take the familiar form of the binomial theorem and integrate both sides (as polynomials in x) w.r.t. x from 0 to 1.

Example 1.2.7. A binary block code C of length n is a set of binary sequences of length n. For two words $x = x_1x_2\cdots x_n$ and $y = y_1y_2\cdots y_n$ we define the *Hamming distance (or distance)* between x and y to be the number of i's where x_i and y_i differ. For example, let $x = 1000110$ and $y = 1101001$. Then the Hamming distance between x and y which we write as $d(x, y)$ is 5. The distance d is a metric on the set of all the binary words of length n. For a word x, how many words are exactly at distance k from x? Since any such word differs from x at exactly k places, and since the length of a word is n, there are precisely $\binom{n}{k}$ words that are at distance k from x. The codewords are transmitted over a communication channel (which is like a telephone line). Since the channel is noisy (prone to make mistakes), the word that is received at the other end of the channel may not be the same as the transmitted word x. The person (or a device, called the decoder) at other side of the channel has, however, a list of all the codewords, i.e., the list of all the words in C. He tries to match the received word with a codeword that is nearest to it. If y is the received word and z is the nearest codeword (word in the code C), then the decoder interprets this as "z must have been sent". It then follows that if it is known beforehand that the channel makes no more than r errors (i.e. at most r of the n positions can have 0's and 1's interchanged by the channel) and if any word is at distance less than or equal to r from at most one codeword, then we will be able to correct all the errors and recover the codeword that was sent. This is called the *nearest neighbor decoding.* (In the above case, $d(x, y) \leq r$, and hence, if z and x are different, then $d(y, z) \geq r + 1$, so that the decoder will recover x and not z). For this to happen, we must have the following condition satisfied: The distance between any two codewords must be at least $2r + 1$. If we draw balls of radius r around each codeword, then no word should belong to two such balls. We call a code C with words of length n an (n, r)-code if any two balls of radius r drawn with centers at two codewords are disjoint from each other, i.e., contain no word in common. For a good code, we should have r as large as possible and also $|C|$ as large as possible since having more codewords amounts to being able to send more information. The following gives an upper bound (called the *Hamming bound*) on the number of codewords in an (n, r)-code C.

$$|C| \leq \frac{2^n}{\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{r}}$$

For a proof, observe that there are 2^n words in all (some of which are codewords). If x is a codeword, then a ball of radius r drawn around x contains no codeword other than x. Since the number of words at distance at most r from x is given by the denominator

of the right hand side of the above expression and since the total number of words is 2^n, a two-way counting produces the desired result.

Definition 1.2.8. Let r be a real number. *By* $\lfloor r \rfloor$, *we mean the largest integer less than or equal to* r *and by* $\lceil r \rceil$, *we mean the smallest integer greater than or equal to* r.

Thus, $\lfloor \pi \rfloor = 3$ while $\lceil \pi \rceil = 4$ and $\lfloor 7 \rfloor = \lceil 7 \rceil = 7$.

Theorem 1.2.9. *Fix a positive integer* n. *Then the number of odd binomial coefficients among the* $n+1$ *numbers:* $\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n-1}, \binom{n}{n}$ *is a power of* 2.

Though a shorter proof of this theorem will be given as an exercise in Chapter 12 (Exercise 12.55), we prefer to give the following elementary and elegant proof, which we break into seven parts.

(a) Let p be a prime number. We say that p divides an integer u exactly m times if p^m divides u but p^{m+1} does not. Let n be natural number and let p a prime. Suppose p divides $n!$ exactly m times. Then

$$m = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots$$

To prove this, write down $n!$ as the product $1 \times 2 \times \cdots \times (n-1) \times n$. In that expression, the multiples of p are $p, 2p, \ldots$ and hence $\left\lfloor \frac{n}{p} \right\rfloor$ in number. The terms such as p^2 should actually be counted two times and they have already been counted once. Again the multiples of p^2 are $p^2, 2p^2, \ldots$ and hence $\left\lfloor \frac{n}{p^2} \right\rfloor$ in number. For these numbers that have been already counted once, we need to add one for each such and hence we obtain the second summand. Then on to multiples of p^3 and continuing in this fashion, we obtain the required result.

(b) Given a non-negative integer n, an expression of the form $a_r a_{r-1} \cdots a_1 a_0$ is called the binary representation of n if

$$n = \sum_{i=0}^{i=r} a_i 2^i$$

and provided that each a_i is 0 or 1. For the sake of uniqueness, We also stipulate that a_r is not 0. It is then easily seen (exercise) that the binary expansion of any non-negative integer is uniquely determined by that integer.

(c) Let a, b, c be non-negative integers and let u be a positive integer. Suppose $a = b + c$. What can we say about $\lfloor \frac{a}{u} \rfloor$, $\lfloor \frac{b}{u} \rfloor$ and $\lfloor \frac{c}{u} \rfloor$? We leave it to the reader to check that

$$\left\lfloor \frac{a}{u} \right\rfloor = \left\lfloor \frac{b}{u} \right\rfloor + \left\lfloor \frac{c}{u} \right\rfloor$$

except when the remainders obtained on dividing b by u and c by u add to a number greater than or equal to u in which case the L.H.S. exceeds the R.H.S. by 1.

(d) Turning back to the binary representation in (b), if n has binary representation $a_r a_{r-1} \cdots a_1 a_0$,then $\lfloor \frac{n}{2^j} \rfloor$ has binary representation $a_r a_{r-1} \cdots a_j$ (check this).

(e) Now let n be a positive integer and let m and k be non-negative integers with $n = m + k$. Since $\binom{n}{k} = \frac{n!}{k!m!}$, the number $\binom{n}{k}$ will be odd (in view of (a) and (c)) if and only if

$$\left\lfloor \frac{n}{2^j} \right\rfloor = \left\lfloor \frac{k}{2^j} \right\rfloor + \left\lfloor \frac{m}{2^j} \right\rfloor$$

for all j. For example, with $j = 0$, if k and m are both odd, then $\binom{n}{k}$ will have a power of 2 surviving and hence will not be an odd integer.

(f) If we now let $a_r a_{r-1} \cdots a_1 a_0$, $b_r b_{r-1} \cdots b_1 b_0$ and $c_r c_{r-1} \cdots c_1 c_0$ to be respectively the binary representations of n, k, m (where, we may add zeros to the left of the representation so as to make them all of the same length) then using (d) and (e) the binomial coefficient $\binom{n}{k}$ will be odd if and only if in writing

$$a_r a_{r-1} \cdots a_j = b_r b_{r-1} \cdots b_j + c_r c_{r-1} \cdots c_j$$

with binary (i.e. modulo 2) addition, there is no carry at any stage of addition. (For example, in the usual 10-based system, a child with no knowledge of 'carry', will make the correct addition of 23 and 45 but will not be able to add 47 and 38 correctly!)

(g) Let the binary representations of n and k be as given above. Given n, if we have to choose k so that the binary coefficient $\binom{n}{k}$ is odd, then we must have a "no carry" situation at each level of addition in (f). Hence if a_i is 0, then b_i must be 0 (and hence c_i is 0), while if a_i is 1, then b_i can be either 1 or 0 (and correspondingly c_i will be 0 or 1 respectively to make the correct binary addition). In any case, each such i gives two choices for b_i. Using the multiplication principle, the number of odd binomial coefficients is a multiple of some 1's and some 2's and is therefore a power of 2. □

For the sake of completeness, we give the following general form of the binomial theorem called Newton's binomial theorem. This can be proved using Taylor's theorem (in fact, it is just a power series expansion with a suitable radius of convergence).

Theorem 1.2.10. *Let α be real number and let x be a real number with absolute value less than* 1. *Then*

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$$

1.3 Counting objects with repetitions

In the example of apples and mangoes given earlier (at the beginning of the present chapter), we had three apples (of different sizes) and two mangoes (of different sizes).

If the three apples were all identical and the two mangoes can not be distinguished from each other, and if we still have to pick up one fruit among these, in how many ways can that be done? The answer is not 5 but is 2 since picking up either of the two mangoes or any one of the three apples makes no difference. In mathematical terms, we say that we are counting objects with repetition or objects are drawn with replacement or we are dealing not with a set of elements but a multi-set. A multi-set $M = \{a_1.x_1, a_2.x_2, \ldots, a_n.x_n\}$ is a collection of distinct objects $x_1, x_2, \ldots, x_n$ with the object x_i occurring a_i times, $i = 1, 2, \ldots, n$. Here a_i's are all non-negative integers and we say that a_i is the multiplicity of the object x_i. The set $S = \{x_1, x_2, \ldots, x_n\}$ will be called the underlying set of the multi-set M. We also allow the possibility of some or all a_i's equal to ∞. This is to be understood as 'the object is available in unlimited supply'. We talk of permutations and combinations of a multi-set in the same way as that of a set. Initially, it might appear that the formulas for permutations and combinations of a multi-set might be as easy as those for sets and might even have straightforward relationships of the kind given in the two Lemmas 1.1.2 and 1.1.5. It is not that simple. For example, if $M = \{4.x_1, 2.x_2, \infty.x_3\}$ then to pick up 10 objects of M, we need to consider different cases: we must pick up j copies of x_2, where j equals 0, 1 or 2. The number of 10-combinations of M is then 15. As another example, the number of r-permutations of the multi-set $M = \{\infty.x\}$ is just one for any value of r.

Theorem 1.3.1. *Let M be a multi-set consisting of r distinct objects, each with infinite multiplicity. Then the total number of d-permutations of M is r^d.*

Proof This simply amounts to counting the number of sequences of length d on an alphabet that consists of r distinct letters. We never run short of any letter since it is allowed to repeat any number of times. Hence by Theorem 1.2.2 (c), the required number is r^d. □

In defining the binomial coefficient $\binom{n}{k}$, we actually count the number of ways of putting n distinct objects into two distinct boxes labeled B_1 and B_2 such that the first box holds k objects and the second box holds the remaining $n-k$ objects. Generalizing this situation prompts us to make the following definition of multinomial coefficient: The number of ways of putting n distinct objects in r distinct boxes $B_1, B_2, \ldots, B_r$ such that the i-th box B_i holds n_i objects is called a multinomial coefficient and is denoted by $\binom{n}{n_1, n_2, \ldots, n_r}$. Necessarily then, $n_1 + n_2 + \cdots + n_r = n$. Thus $\binom{n}{k} = \binom{n}{k,\ n-k}$.

Theorem 1.3.2. *Let S be an n-set and suppose the n objects in S are to be put in r distinct boxes $B_1, B_2, \cdots, B_r$ such that the i-th box B_i contains n_i objects with $n_1 + n_2 + \cdots + n_r = n$. Then the number of ways of doing this is equal to*

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}$$

Proof Though a direct proof can be given, we prefer to make an induction on r, purely for pedagogical reasons. For $r = 2$, Lemma 1.1.5 and the discussion following

it show that the statement of the theorem is true. Let $r \geq 3$. Out of n objects, we may first choose n_1 objects to put in the first box B_1 in $C(n, n_1) = \binom{n}{n_1}$ ways and then try to put the remaining $n - n_1$ objects in the other $r - 1$ boxes. Having performed the first task, that can be done in $\binom{n-n_1}{n_2, n_3, \ldots, n_r}$ ways. Now using induction on r,

$$\binom{n - n_1}{n_2, n_3, \ldots, n_r} = \frac{(n - n_1)!}{n_2! \cdots n_r!}$$

Hence the required number equals

$$\frac{n!}{n_1!(n - n_1)!} \times \frac{(n - n_1)!}{n_2! \cdots n_r!} = \frac{n!}{n_1! n_2! \cdots n_r!}$$

□

Corollary 1.3.3. *Let M be a multi-set consisting of r distinct objects $x_1, x_2, \ldots, x_r$ such that the i-th object x_i has multiplicity n_i. Let $n = n_1 + n_2 + \cdots + n_r$. Then the total number of n-permutations of M is*

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}$$

Proof We set up a bijection between the required set of all the ways of putting the elements of S in r boxes and all the n-permutations of the multi-set

$$M = \{n_1.x_1, n_2.x_2, \ldots, n_r.x_r\}$$

First number the elements of S from 1 to n. If the element i is put in the box B_j, then make an n-permutation in which the i-th place is occupied by x_j. Conversely, given an n-permutation of M, if we find the i-th place occupied by x_j then put the element i in the box B_j. Hence the result is proved using bijection and Theorem 1.3.2. □

It is also possible to prove Corollary 1.3.3 without using Theorem 1.3.2.

Theorem 1.3.4. *(The multinomial theorem) Let n be a non-negative integer. Then:*

$$(x_1 + x_2 + \cdots + x_r)^n = \sum_{n_1 + n_2 + \cdots + n_r = n} \binom{n}{n_1, n_2, \ldots, n_r} {x_1}^{n_1} {x_2}^{n_2} \cdots {x_r}^{n_r}$$

where the sum is taken over all the non-negative integers n_i that satisfy the given stipulation.

Proof We sketch three different proofs.
First Proof Make induction on r using the fact that the assertion holds for $r = 2$, which is the binomial theorem (Theorem 1.2.5).
Second Proof Make induction on n and use Pascal type identities.
Third (direct) Proof Expand directly imitating the proof of the binomial theorem (Theorem 1.2.5) and use Theorem 1.3.2. □

Theorem 1.3.5. *Let M be a multi-set with r distinct objects $x_1, x_2, \ldots, x_r$ each with infinite multiplicity. Let $\overline{C}(r,k)$ denote the number of k-combinations of M. Then $\overline{C}(r,k) = \binom{k+r-1}{k}$.*

Proof Every k-combination is uniquely determined by a sequence $b_1, b_2, \ldots, b_r$ where b_i's are all non-negative integers and $b_1 + b_2 + \cdots + b_r = k$. For each k-combination of M make a binary sequence of length $k + r - 1$ as follows. At the beginning, write b_1 zeros and follow this by a 1, then write b_2 zeros and then write a 1 and so on. There will be a final 1 separating the b_{r-1} zeros and the last b_r zeros. Thus the binary sequence will consist of exactly $b_1 + b_2 + \cdots + b_r = k$ zeros and $r - 1$ ones. Conversely given a binary sequence of length $k + r - 1$ consisting of k zeros and $r - 1$ ones, we read the number of zeros to the left of the first one and call it b_1, then the number of zeros between the first one and the second and call it b_2 and so on. Finally the number of zeros to the right of the last one is b_r. For example, with $r = 4$ and $k = 6$, the sequence $2, 1, 1, 2$ gives the binary sequence 001010100 while the binary sequence 100011000 must have come from $(0, 3, 0, 3)$. Since the number of binary sequences of length $k + r - 1$ with exactly k zeros equals $C(k + r - 1.k)$, the result is proved using bijection. □

Let S be an r-set. We call S an ordered set if there exists an order $<$ on the elements of S which is a total order (or a chain). That is, the elements of S can be written in the form $c_1, c_2 \ldots, c_r$ where $c_1 < c_2 < \cdots < c_r$. A sequence $(x_1, x_2, \ldots, x_k)$ with entries from the set S is said to be a monotonically increasing sequence if $x_1 \leq x_2 \leq \cdots \leq x_k$. We then have

Theorem 1.3.6. *The following sets are in bijective correspondence.*

(a) The set of all increasing sequences of length k on an ordered set with r elements.

(b) The set of all the ways of putting k identical objects into r distinct boxes.

(c) The set of all the k-combinations of a multi-set with r distinct elements.

In all the three cases, the cardinality of the set is $C(k + r - 1, k)$.

While we leave the proof of the above theorem to the reader, to conclude this section, we also note some interesting connections.

$$\overline{C}(r,k) = \binom{k+r-1}{k} = (-1)^k \binom{-r}{k}$$

Now let $[r]^k$ denote the product $r(r+1)\cdots(r+k-1)$. This called the *rising factorial*. We then have

$$\overline{C}(r,k) = \frac{[r]^k}{k!}$$

Theorem 1.3.7. *The number of ways of putting k identical objects into r distinct boxes with each box containing at least one object is $\binom{k-1}{r-1}$.*

Proof Follow the proof of Theorem 1.3.5. We have to find the totality of binary sequences of length $k + r - 1$ with no two 1's adjacent and the sequences do not

begin or end with a 1. Given such a binary sequence remove one zero between every two adjacent 1's and also one zero each from the left and the right ends. This leaves us with a binary sequence of length $k-1$ with exactly $r-1$ ones. The process is reversible. Since the number of binary sequences of length $k-1$ with $r-1$ ones is $\binom{k-1}{r-1}$, the assertion is proved. □

Example 1.3.8. *An application to Statistical Mechanics:* In Statistical Mechanics, one encounters the situation of putting k particles into r distinct energy levels. The particles can thus be considered as objects and the different energy levels as distinct boxes or cells. Three different situations (statistics) are obtained by making three different assumptions. These are

(a) Maxwell-Boltzman: Here the particles are all distinct and any number of particles can be put in any of the r boxes. The number of possibilities (as given by Theorem 1.3.1) is r^k.

(b) Bose-Einstein: Here the particles are all identical and any number of particles of particles can be put in any of the r boxes. The number of possibilities (as given by Theorem 1.3.6) is $\binom{k+r-1}{k}$.

(c) Fermi-Dirac: Here the particles are all identical but no box can hold more than one particle. The number of possibilities is $\binom{r}{k}$.

1.4 Two-way counting revisited: the de Bruijn-Erdös Theorem

In this last part of the discussion on basic counting techniques, we give a somewhat sophisticated and deep application of two-way counting. The result known as the de Bruijn-Erdös theorem was proved by the authors using repeated applications of two-way counting. This theorem first appears in the literature in 1948. However, Erdös knew its proof ten years prior to its appearance in print. But he did not publish it at that time because "It was considered relatively less important to do mathematics of that sort!" All the combinatorial proofs of the de Bruijn-Erdös theorem tend to be somewhat messy. A short proof (due to Conway) given in van Lint and Wilson [57] is discussed here.

A (finite) incidence structure $\mathbf{I}$ is a pair $\mathbf{I} = (\mathbf{P}, \mathbf{L})$ where $\mathbf{P}$ is (finite) set called the set of points and $\mathbf{L}$ is a set of subsets of $\mathbf{P}$. Each member of $\mathbf{L}$ is called a line. A *linear space* is an incidence structure in which every pair of points is contained in a unique line (thus *no* two lines intersect in two or more points). To avoid obvious trivialities, we stipulate that every line has size at least two but no line contains all the points (in that case there will be no other line). Letting the number of points to be v and the number of lines to be b, what relationship do these two integers have in general?

There are two special linear spaces of interest. A linear space is called a *near pencil* if some line has size $v-1$ (and hence necessarily other lines have size two each). In this

case, there is just one more point outside the line of size $v-1$ and this point is on $v-1$ lines each of size two. We thus have $v = b$. Clearly a near pencil can be constructed for all the values of $v \geq 3$. A more special linear space is what is called a *projective plane*. Here all the lines have the same size $n+1$ for some $n \geq 2$ and further any two lines intersect each other. This object is called a projective plane of order n. Figure 1.3 shows a projective plane of order two and a near-pencil with $v = 6$.

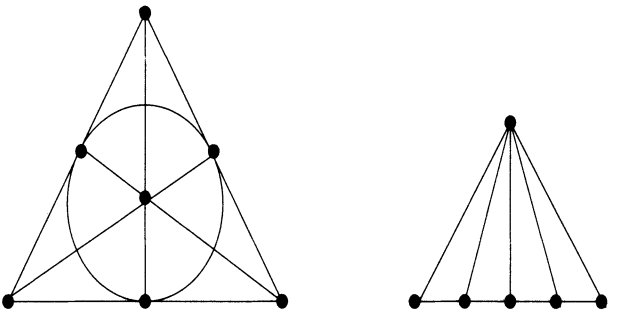

Figure 1.3: Fano Plane and Near-Pencil

Answering the question posed in the previous paragraph, we have the following theorem of de Bruijn and Erdös:

Theorem 1.4.1. *In a linear space* **I**, *we must have* $v \leq b$ *with equality iff the linear space is a near pencil or a projective plane.*

Proof All the combinatorial proofs to the above theorem depend on one or more applications of two-way counting. We first make some elementary observations. For a point x of the linear space **I** under consideration, let r_x denote the number of lines of **I** that contain x and for a line L of **I**, let k_L denote the number of points contained in L. We proceed through the following claims.

(a) If $x \notin L$, then $r_x \geq k_L$ with equality iff each line containing x meets L.

Proof For each y on L we must have a line containing both x and y and no two such lines can be identical for then the line L will intersect some line on x in two points. Further equality can hold iff there is no line containing x which is disjoint from L.

(b) Every two lines meet each other (i.e. are not disjoint) if for all $x \notin L, r_x = k_L$. This is obvious from claim (a).

(c) We have

$$\sum_{y \in \mathbf{P}} r_y = \sum_{L \in \mathbf{L}} k_L$$

Proof Make a two-way counting on the set S consisting of all the pairs (y, L) where y is a point on the line L.

At this point, we assume that $b \leq v$ and then prove that this leads to $b = v$ and $\mathbf{I}$ *must be a near-pencil or a projective plane.*

(d) Let L and M be two lines such that $L \cup M = \mathbf{P}$. Then we have

$$L \cap M \neq \emptyset; \;\; b = v$$

and $\mathbf{I}$ is indeed a near pencil.

Proof If $L \cap M \neq \emptyset$ then let $\{p\} = L \cap M$. It then follows that L and M are the only lines containing p and hence every other line Z must intersect each one of L and M in a single point. In particular, we must have $|Z| = 2$ for such a line Z. Let $\alpha = |L| \geq |M| = \beta \geq 2$. We see that if $L \cap M = \emptyset$, then $v = \alpha + \beta$ and $b = 2 + \alpha\beta$ while if $L \cap M \neq \emptyset$, then $v = \alpha + \beta - 1$ and $b = 2 + (\alpha - 1)(\beta - 1)$. In the former case, we get (using $b \leq v$), $(\alpha - 1)(\beta - 1) + 1 \leq 0$ which is absurd and in the latter case, we get $(\alpha - 2)(\beta - 2) \leq 0$ showing that $\beta = 2$ and hence $\alpha = v - 1$ showing that we have a near pencil as desired.

(e) Let $\mathbf{I}$ be not a near pencil. Let every two lines have a non-empty intersection. Then $\mathbf{I}$ is a projective plane.

Proof Let L and M be two lines. Using claim (d), we see that there is a point $x \notin L \cup M$. Then the hypothesis implies that $k_L = r_x = k_M$ showing that all the lines have the same number of points $k = n + 1$ where $n \geq 2$. Then all the points x also satisfy $r_x = n + 1$ and we have a projective plane as desired.

We now finish the proof of the theorem using the following clever argument of Conway as quoted in the book [57] by van Lint and Wilson. Using (b) and (d), it will suffice to show that $r_x = k_L$ for all the pairs (x, L) with x not on L. To that end, fix such a pair (x, L) with x not on L. Then (a) shows that $r_x \geq k_L$ and hence because of the assumption $vr_x \geq bk_L$, i.e. $-vr_x \leq -bk_L$, i.e. $v(b - r_x) \leq b(v - k_L)$. Thus

$$\frac{1}{v(b - r_x)} \geq \frac{1}{b(v - k_L)}$$

Then sum both sides of the inequality over all the elements of the set T consisting of all the pairs (x, L) with x not on the line L. Consider the L.H.S. of the above inequality. Fix a point x and sum the expression over all the lines not containing x. These are $b - r_x$ in number, and then summing over all the points x must sum the L.H.S. to 1. Use two-way counting to change the order of summation and sum the expression on the R.H.S. over all the elements of T to obtain 1 again. We thus have:

$$1 = \sum_{x \in \mathbf{P}} \sum_{x \notin L} \frac{1}{v(b - r_x)} \geq \sum_{L \in \mathbf{L}} \sum_{x \notin L} \frac{1}{b(v - k_L)} = 1$$

Therefore equality must hold everywhere and $r_x = k_L$ for all the pairs x, L with x not all L and we are done. □

A slightly different proof of Theorem 1.4.1 will be given in Exercise 16.19.