Chapter 4 - COUNTING AND BASIC COMBINATORICS

4.1 PRINCIPLE OF INCLUSION AND EXCLUSION

We want to develop tools, based on Set Theory, which will enable us to solve the following type of counting problem (as well as more general ones).

Example 1. Of 100 students surveyed, 43 have taken course A, 55 have taken course B, 30 have taken course C, 18 have taken both course A and course B, 13 have taken both courses A and C, 15 have taken B and C, and 8 took all three courses. How many of the 100 students did not take any of the three courses?

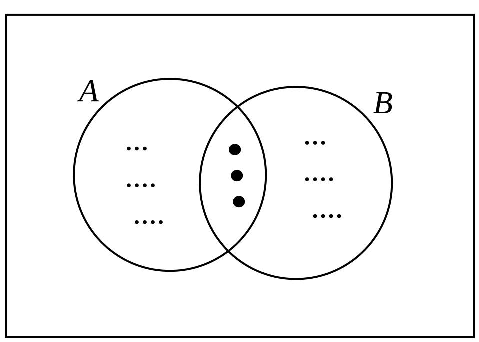
We begin with some elementary observations the first of which is obvious.

Claim 2. If two finite sets *A* and *B* are disjoint then $\boxed{|A\cup B| = |A| + |B|}$.

When $A\cap B \neq \emptyset$, the above formula is wrong because the total on the right-hand side counts the elements of $A\cap B$ twice, once as elements of *A* and once as elements of *B*, whereas these elements are counted only once on the left-hand side. The next claim fixes this problem.

Claim 3. For all sets *A* and *B*, $\boxed{|A\cup B| = |A| + |B| - |A\cap B|}$.

In this formula each element of the union $A\cup B$, counted once on the left-hand side, is also counted once on the right-hand side. Venn diagram depicts the situation well – the elements in the intersection of the two sets, shown in the figure as boldface points, are counted twice. All the other points are counted once if they are *A* or in *B*, and they are not counted at all if they outside of $A\cup B$.



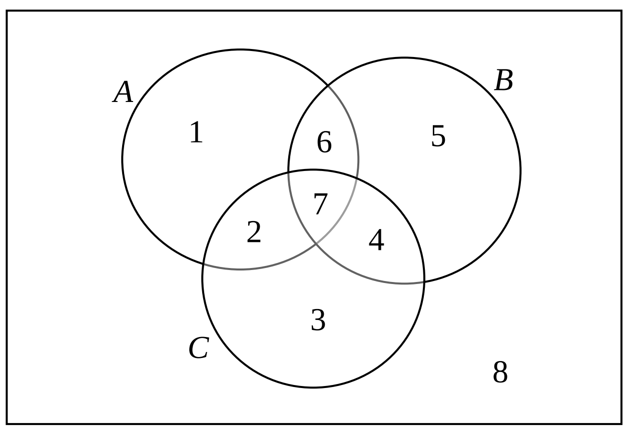
Example 4. Fifty students in a certain dormitory spent an evening on the following activities:

(a) 10 went to sleep early, (b) 30 went to see the movie "Amadeus," and
(c) 30 went to a bar.

Students that went to sleep did not partake in any other activities. We want to know how many students watched the movie and also went to the bar?

Solution. We use Claim 3, with A the students who went to watch the movie and B the students who went to the bar. It is given that $|A| = |B| = 30$ and $|A \cup B| = 50 - 10 = 40.$ Hence, $|A \cap B| = |A| + |B| - |A \cup B| = 30 + 30 - 40 = 20.$ Drawing a Venn diagram could help seeing all the details.

Lets consider the general case of three sets (like in Example 1). The Venn diagram of the general situation is shown below (where the numbers represent the regions and not the number of elements in that region):



We want to count $|A \cup B \cup C|$. The first approximation is $S_1 = |A| + |B| + |C|$ which involves counting elements in the various areas of the Venn diagram given above as follows:

(a) elements in areas 1, 3, and 5 are counted once,
(b) elements in areas 2, 4, and 6 are counted twice,
(c) elements in area 7 are counted thrice, and
(d) elements in area 8 are not counted.

The second approximation attempts to fix the count for those elements that were counted twice in S_1. This done by subtracting $S_2 = |A \cap B| + |A \cap C| + |B \cap C|$ from S_1. Thus, the second approximation is $S_1 - S_2$ and it correctly (i.e., just once) counts the elements in areas: 1, 3, 5, 2, 4, and 6. However, the elements in area 7 are now not counted at all! We fix this by simply adding $|A \cap B \cap C|$ to the second approximation. This reasoning proves the following result.

Theorem 5. Let *A*, *B*, and *C* be finite sets. Then

$$\boxed{|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|}.$$

The formula of Theorem 5, is called the ***Principle of Inclusion-Exclusion*** (for three sets), abbreviated as PIE_3. It is also known as the ***Sieve Principle***.

We are now in a position to solve the problem formulated in Example 1.

Example 6 (Example 1 all over again). Of 100 students surveyed, 43 have taken course A, 55 have taken course B, 30 have taken course C, 18 have taken both course A and course B, 13 have taken both courses A and C, 15 have taken B and C, and 8 took all three courses. How many of the 100 students did not take any of the three courses?

Solution. Lets use *A*, *B*, and *C* for the sets of students that take the corresponding class. Thus,

$$|A| = 43, \quad |B| = 55, \quad |C| = 30, \quad |A \cap B| = 18,$$
$$|A \cap C| = 13, \quad |B \cap C| = 15, \quad |A \cap B \cap C| = 8.$$

Using PIE_3, we obtain $|A \cup B \cup C| = (43 + 55 + 30) - (18 + 13 + 15) + 8 = 128 - 46 + 8 = 90.$ It follows that 10 students did not enroll in any of the three courses.

In the general case we have *n* finite sets whose various “overlappings” have some given cardinalities. We want to find the cardinality of the union of these sets or possibly the cardinality of some other combination of these sets. The corresponding general principle is denoted by PIE_n.

Theorem 7. [PIE_n] Let $A_1, A_2, ..., A_n$ be finite sets. Then,

$$
\boxed{\begin{aligned}
|\bigcup_{k=1}^{n} A_k| &= \sum_{i=1}^{n} |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| - .... + (-1)^{n-1} |\bigcap_{k=1}^{n} A_k| \\
&= \sum_{I \subseteq \{1,2,...,n\}} (-1)^{|I|+1} |\bigcap_{k \in I} A_k|
\end{aligned}}
$$

Proof. For the proof it suffices to show that each element of the union on the left-hand side (which is obviously counted there just once) is contributing exactly 1 in the count given in the right-hand side. Let *x* an arbitrary element in the union and suppose that it belongs to precisely *m* of the sets $A_1, A_2, ..., A_n$. Then,

1. x contributes $m = \binom{m}{1}$ in the first sum on the right-hand side,
2. x contributes $\binom{m}{2}$ in the second sum on the right-hand side,
3. x contributes $\binom{m}{3}$ in the second sum on the right-hand side,

and generally, x contributes $\binom{m}{k}$ in the k^{th} sum on the right-hand side. Hence the total contribution of x to the count on the right-hand side is:

$$
\binom{m}{1} - \binom{m}{2} + \binom{m}{3} - ... + (-1)^{m-1} \binom{m}{m}
$$

which, by a result proved as a consequence of the Binomial Theorem (Example 17 in the handout on Induction), is equal to $\binom{m}{0} = 1.$ ■

Corollary 8. Let $A_1, A_2, ..., A_n$ be subsets of the finite set *X*. Then

$$
\boxed{\begin{aligned}
|X - \bigcup_{k=1}^{n} A_k| &= |X| - |\bigcup_{k=1}^{n} A_k| \\
&= |X| - \sum_{i=1}^{n} |A_i| + \sum_{i<j} |A_i \cap A_j| - \sum_{i<j<k} |A_i \cap A_j \cap A_k| + ... + (-1)^{n} |\bigcap_{k=1}^{n} A_k|
\end{aligned}}
$$

Example 9. Consider the 4-digit decimal integers between 0000 and 9999. How many among these do not have k in the k^{th} place (counting from the left), $k = 1, 2, 3, 4$?

Solution. We count the number integers that violate the requirements. Denote by A_k the subset of these numbers that do have a k in the k^{th} place, $k = 1, 2, 3, 4.$ We easily see that:

$$|A_k| = 1000, \quad |A_k \cap A_l| = 100, \quad |A_k \cap A_l \cap A_m| = 10, \quad |A_1 \cap A_2 \cap A_3 \cap A_4| = 1$$

where all the indices are appropriately distinct and in the correct range. It follows, using PIE_4 that the number of integers that violate at least one of the requirements is

$$|A_1 \cup A_2 \cup A_3 \cup A_4| = 4 \cdot 1000 - 6 \cdot 100 + 4 \cdot 10 - 1 = 3439,$$

and so the number of integers that do satisfy the requirements is $10000 - 3439 = 6561.$

4. 2. GENERALIZED PIGEONHOLE PRINCIPLE

Lets first recall the "standard" pigeonhole principle mentioned in the context of our study of properties of functions: when placing more than n pigeons in n pigeonholes, at least one pigeonhole must contain two or more pigeons. Or in terms of functions: any function from a set with more than n elements to an n-set is not 1-1.

Here is a neat application of the standard pigeonhole principle.

Example 10. [Rosen, pg. 351] Show that among arbitrary $n+1$ positive integers, none exceeding $2n$, there must be an integer that divides one of the other integers.

Solution. Write each of the $n+1$ integers $a_1, a_2, ..., a_{n+1}$ as a product of a power of 2 times an odd integer: $a_i = 2^{k_i} q_i$ for $i = 1, 2, ..., n+1,$ where $k_i \geq 0$ is an integer and q_i is odd. The integers $q_1, q_2, ..., q_{n+1}$ are all positive odd integers and all are $< 2n.$ Since there are only n positive odd integers $< 2n$, by the pigeonhole principle, for some $i \neq j$, $q_i = q_j$. Let this common value be denoted by q. Thus $a_i = 2^{k_i} q$ and $a_j = 2^{k_j} q.$ It follows that if $k_i < k_j$ then $a_i \mid a_j$, and if $k_j < k_i$ then $a_j \mid a_i.$

To introduce the generalized principle we start with an example.

Example 11. How many positive integers in $[1000]$ are not divisible by 5 or by 9?

Solution. All the multiples of 5 are precisely the positive integers divisible by 5; denote this set by M_5. Clearly, $|M_5| = \left\lfloor \frac{1000}{5} \right\rfloor = 200.$ By exactly the same argument, for the set of multiples of 9 we have $|M_9| = \left\lfloor \frac{1000}{9} \right\rfloor = 111.$ Now use PIE_2 to obtain:

$$|M_5 \cup M_9| = |M_5| + |M_9| - |M_5 \cap M_9| = 200 + 111 - |M_5 \cap M_9|.$$

However, 5 and 9 are *relatively prime* (a.k.a. *co-prime*), which, in our case, simply means that an integer is divisible by 5 and by 9 precisely when it is divisible by their product 45. Hence, $|M_5 \cap M_9| = |M_{45}| = \left\lfloor \frac{1000}{45} \right\rfloor = 22.$ So $|M_5 \cup M_9| = 311 - 22 = 289.$ Finally, by Corollary 8, the number of integers not divisible by 5 or by 9 is $1000 - 289 = 711.$

We now formalize this more general type of pigeonhole principle.

Theorem 12. (***Generalized Pigeonhole Principle***) If n objects are placed in k boxes, then there is at least one box with at least $\left\lceil n/k \right\rceil$ objects.

Proof. If all the boxes contained at most $\left\lceil n/k \right\rceil - 1$ objects then the total number of objects would be at most $k\left(\left\lceil n/k \right\rceil - 1\right) < k\left(\left(n/k + 1\right) - 1\right) = n$, which contradicts the assumption that there is a total of n objects. ■

Example 13. [Rosen, pg. 352] In a group of 6 people each pair either *likes* each other or *hates* each other. Prove that this group contains a 3-subset of people who mutually like each other or who mutually hate each other.

Solution. Fix an arbitrary individual in the group, say John, and consider his relation to the other 5 people. There are 5 such relationships and each is of one of two kinds (either *likes* or *hates*). So by the generalized pigeonhole principle there must be either three other people that John likes or three other people that John hates. WLOG suppose there are three of these other people that John likes. If any two of these like each other, then, together with John they form a 3-subset of people who mutually like each other. On the other hand, if no two of them like each other they form a 3-subset of people who mutually hate each other.

4.3. PIEs AND PERMUTATIONS → DERANGEMENTS

A bijection of a finite set into itself is called a ***permutation*** of the given set. We consider the set $[n] \triangleq \{1, 2, ..., n\}$. The set of all permutations of $[n]$ is commonly denoted by S_n. We know that $\boxed{|S_n| = n!}$. If $\pi \in S_n$ is a permutation and $\pi(i) = i$ for some $i \in [n]$, we say that π *fixes* *i*. π is called a *derangement* if $\pi(i) \neq i$, $\forall i \in [n]$. Obviously, a permutation is a derangement if, and only if, it does not fix any $i \in [n]$. Denote by $D_n \subseteq S_n$ the set of all derangements on $[n]$. In the following example we will calculate the cardinality $|D_n|$.

Example 14. [*The confused secretary problem*.] A secretary has *n* letters and *n* addressed envelopes. In how many ways can the secretary put every letter in a wrong envelope? This problem is the *derangement problem* (in disguise) and it should be clear why: each wrong letter-to-envelope assignment is in fact a derangement. Denote by F_i the subset of S_n comprising all the permutations that fix *i* (these permutations can, but don't have to, fix other $j \in [n]$). It is immediately seen that regardless of *i*, $|F_i| = (n-1)!$ How many different (but not necessarily disjoint) F_i's are there? Answer: as many as there are different $i's$, i.e. $n = \binom{n}{1}$. Next, lets count how many permutations fix *i* and *j* (distinct); obviously these are precisely those in the set $F_i \cap F_j$ and, clearly, $|F_i \cap F_j| = (n-2)!$ regardless of the specific $i \neq j$. In addition, the number of such intersections is the same as the number of 2-subsets $\{i, j\} \subseteq [n]$; i.e., $\binom{n}{2}$. More generally, the number of permutations that fix specific *k* elements of $[n]$ is $(n-k)!$ and the set of these permutations can be expressed as an intersection of the *k* corresponding F_i's. Moreover, the number of ways to pick k F_i's out of possible *n* is, as we know, the binomial coefficient $\binom{n}{k}$. Now, the union $F_1 \cup F_2 \cup ... \cup F_n$ includes precisely all and only those permutations that fix at least one element of $[n]$, viz., the complement $S_n - D_n$. Applying PIE_n, we can compute the cardinality of $F_1 \cup F_2 \cup ... \cup F_n$:

$$
\begin{aligned}
|F_1 \cup \ldots \cup F_n| &= \sum_{i=1}^{n} |F_i| - \sum_{i<j} |F_i \cap F_j| + \sum_{i<j<k} |F_i \cap F_j \cap F_k| + \ldots + (-1)^{n-1} |F_1 \cap \ldots \cap F_n| \\
&= \binom{n}{1} \cdot (n-1)! + \ldots + (-1)^{k-1} \binom{n}{k} \cdot (n-k)! \ + \ldots + (-1)^{n-1} \binom{n}{n} \cdot (n-n)! \\
&= \frac{n!}{1!(n-1)!}(n-1)! + \ldots + (-1)^{k-1} \frac{n!}{k!(n-k)!}(n-k)! + \ldots + (-1)^{n-1} \frac{n!}{n!(n-n)!}(n-n)! \\
&= n!\left[1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + (-1)^{n-1} \frac{1}{n!}\right].
\end{aligned}
$$

It follows that the cardinality of D_n is

$$
|D_n| = |S_n| - |F_1 \cup F_2 \cup \ldots \cup F_n| = n! - n!\left[1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + (-1)^{n-1} \frac{1}{n!}\right],
$$

and hence

$$
\boxed{|D_n| = n!\left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots + (-1)^{n} \frac{1}{n!}\right]}.
$$

4.4. <u>THERE IS MORE ABOUT PERMUTATIONS</u>

We first present two easy examples illustrating a basic, but important, counting principle.

Example 15. A bank password consists of two letters of the English alphabet followed by two digits (0 to 9). How many different passwords are there?

Solution. Obviously, since there are 26 letters in the English alphabet, the first two symbols of the password can be chosen in $26 \cdot 26 = 676$ ways. This depends crucially on the implicit assumption that the two choices of letters are <u>*independent*</u> of each other. Next, the two digits can be each chosen in 10 ways, resulting 100 ways to pick the two digits. Hence the total number of passwords is 67,600.

Example 16. A coin is tossed four times and the result of each toss is recorded. How many different sequences of heads (H) and tails (T) are possible?

Solution. Assuming independence of the tosses and no “complete bias” of any of the coins, the total number of 4-sequences is $2 \cdot 2 \cdot 2 \cdot 2 = 16$.

The principle at work in these two examples is the so called <u>Product Rule</u>: if a task T can be broken (decomposed) into a sequence of *independent* tasks T_1 , $T_2, \dots, T_n$ and if the number of ways to perform task T_i is t_i $(1 \le i \le n),$ then the number of ways to perform the task T is $t = t_1 \cdot t_2 \cdot \dots \cdot t_n$.

A permutation was defined above as a bijective function of a set onto itself. Sometimes it is useful to think of permutations as "sequential arrangements" of the elements of the set. Thus, for the set $A = \{a, b, c\}$ there are 6 such arrangements: (*abc*), (*acb*), (*bac*), (*bca*), (*cab*), and (*cba*). These arrangements can be viewed as permutations in the obvious way once we pick a fixed ordering of *A*. So, taking the usual alphabetic order, the arrangement (*bac*) corresponds to the permutation: $f(a) = b,\ f(b) = a,\ f(c) = c.$ Moreover, it should be quite clear that as far as counting permutations of *A* goes, it is equivalent to counting permutations of the set $[3] = \{1, 2, 3\}$. It follows that the number of permutations of *A* is the same as the cardinality of S_3, i.e. 3! = 6. More generally, $|S_n| = n!$, which we have defined as the number of permutations of $[n]$, <u>can equivalently be viewed as the number of ways to sequentially arrange elements of any *n*-set</u>.

The permutations as discussed above are called <u>*n*-permutations</u> (of *n* objects) because they involve arrangements of all the *n* elements of the given *n*-set. We now generalize this concept. Let $0 \le r \le n.$ An <u>*r*-permutation</u> (of *n* objects) is an arrangement of *r* objects out of the given *n*-set.

Example 17. Lets list all 3-permutations of $[4] = \{1, 2, 3, 4\}$:

(123), (132), (213), (231), (312), (321),
(124), (142), (214), (241), (412), (421),
(134), (143), (314), (341), (413), (431),
(234), (243), (324), (342), (423), (432).

The number of r-permutations of $[n]$ (or any other *n*-set) is denoted by $P(n, r)$. We have observed that $P(n, n) = n!$ We also make the convention that $P(n, 0) = 1.$ The general result is the following:

Theorem 18. $\forall 0 \le r \le n$, $\boxed{P(n,r) = \frac{n!}{(n-r)!}}$.

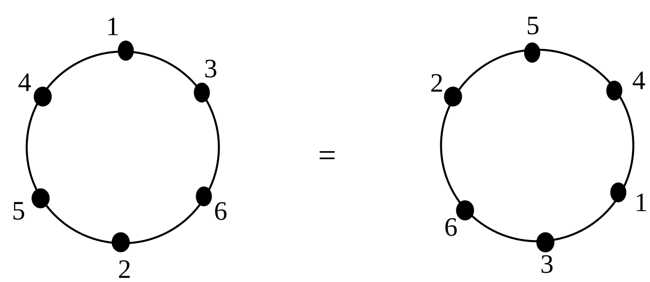
Proof. We count in how many ways we can construct an *r*-permutation of $[n]$. The number of choices for the first position in the *r*-permutation is *n*, the number of choices for the second position is $n-1$ (because one of the elements of $[n]$ has been used up), and generally, the number of choices for the k^{th} position is $n-(k-1)$ (because $k-1$ of the elements of $[n]$ has been used up). By the product rule we have:

$$\begin{aligned} P(n,r) &= n(n-1)\ldots(n-(k-1))\ldots(n-(r-1)) \\ &= n(n-1)\ldots(n-(k-1))\ldots(n-(r-1))\frac{(n-r)!}{(n-r)!} \\ &= \frac{n!}{(n-r)!} \end{aligned}$$

■

The following example is interesting.

Example 19. In how many ways we can seat *n* guests around a round table? Two seatings are considered the same if every guest has the same neighbors on her/his left and her/his right. In other words, we count as the same two seatings that can be obtained from each other by “rotation” of the *n* guests clockwise around the table.



Solution. Since the seatings have “rotational symmetry” we can take an arbitrary guest and seat him in some fixed place. This place will be the “zero point” in that the other guests are seated “relative to him”; this guest’s seating is used to “break the symmetry.” Now the task is to seat the remaining $n-1$ guests and this can be achieved in $(n-1)!$ ways.

Example 20. How many permutations of $[n]$ are there in which 1 and 2 are adjacent?

Solution. There are $2(n-1)!$ permutations because we can remove 1 and 2 and add (12) as a new object. This gives a total of $n-1$ object that can be permuted, as we know, in $(n-1)!$ ways. In each of these permutations we can "unpack" (12) in two ways: either as 12 or as 21, giving the claimed result.

Example 21. How many permutations of $[n]$ are there in which 1 and 2 are not adjacent?

Solution. Since 1 and 2 are either adjacent or not, Example 20 implies that the required number is: $n!-2(n-1)!=(n-2)(n-1)!$.

Example 22. Suppose a running contest includes 8 runners. The winner gets a gold medal, the runner up receives a silver medal, and the third-place finisher receives a bronze medal. In how many different ways can the medals be awarded?

Solution. This is just the number of 3-permutations of 8 elements: $P(8,3)=8\cdot 7\cdot 6=336$.

4.5. <u>PERMUTATIONS WITH REPETITIONS</u>

Let *S* be a set with *n* objects divided into (disjoint) *m* subsets $S=S_1\cup S_2\cup ...\cup S_m$ where $|S_i|=k_i\geq 0$ and $k_1+k_2+...+k_m=n$.

Example 23. *S* consists of 9 balls of which 4 are black, 3 are white, and 2 are green. In this case $n=9, m=3, k_1=4, k_2=3,$ and $k_3=2$.

We say that two permutations are "*the same*," or *equivalent*, if they differ only in the placement of elements in the same group.

Example 24. [cont.] Let the 9 balls be: $b_1,b_2,b_3,b_4,w_1,w_2,w_3,g_1,g_2$. Then the permutations

$$b_2, b_3, w_1, g_2, b_4, g_1, w_3, b_1, w_2$$
$$b_3, b_1, w_2, g_2, b_2, g_1, w_1, b_4, w_3$$
$$b_2, b_1, w_1, g_1, b_3, g_2, w_3, b_4, w_2$$

are equivalent to each other but not to $b_1, g_1, b_2, b_3, g_2, w_1, w_2, b_4, w_3$.

Theorem 25. The number of different permutations of an *n*-set *S*, k_i of which are of "type *i*," $i = 1, 2, \ldots, m$, where $k_1 + k_2 + \ldots + k_m = n$, is

$$\boxed{\binom{n}{k_1, k_2, \ldots, k_m} \triangleq \frac{n!}{k_1!k_2!\ldots k_m!}}.$$

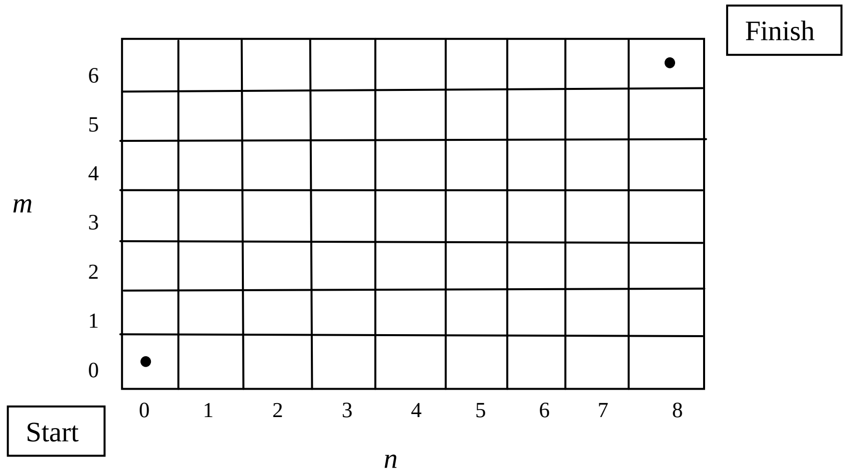
Remark. The expression on the right-hand side is often referred to as a ***multinomial coefficient***. The expression on the left-hand side is just a notation emphasizing the similarity of the expression on the right-hand side to the binomial coefficient. In fact, it is easy to see that

$$\binom{n}{k} = \binom{n}{k, n-k} = \frac{n!}{k!(n-k)!}.$$

Moreover, this result generalizes the known fact that there are *n*! permutations of an *n*-set. In this case all k_i are equal to 1 and $m = n$.

Proof. Let *K* be the number of (inequivalent) permutations as required by the statement of the theorem. From each such permutation we can obtain $k_1!$ (equivalent) permutations by permuting elements of type 1 between themselves. From each of these we can construct $k_2!$ (equivalent) permutations by permuting elements of type 2 between themselves. Continuing in this way we see that from each of the *K* inequivalent permutations we can obtain $k_1!k_2!\ldots k_m!$ (equivalent but) distinct permutations of the set *S*. This gives a total of $K k_1!k_2!\ldots k_m!$ permutations of *S* which must be equal to *n*!. ■

Example 26. Consider the following $m \times n$ grid:



How many paths are there from the $(0,0)$ cell to the $(m-1, n-1)$ cell? In each step a path can only move either up or to the right.

Solution. One has to travel $m+n$ "streets," m of these are up (north) and n are right (east). Moreover, any sequence of m "norths" and n "easts" takes the traveler from Start to Finish and various permutations of the $m+n$ directions give different ways to get from Start to Finish. Hence the total is the multinomial coefficient $(m+n)!/m!n!$ which is precisely equal to the binomial coefficient $\binom{m+n}{n}$.

The reason for the name "multinomial coefficients" is that these expressions occur as coefficient in the *Multinomial Theorem* in the same manner as the binomial coefficients occur in the Binomial Theorem. Here is the statement of the Multinomial Theorem which is a very nice generalization of the Binomial Theorem.

Theorem 27. [***Multinomial Theorem***] Let $n \in \mathbb{N}$. Then

$$\left(x_1 + x_2 + \ldots + x_m\right)^n = \sum_{k_1+\ldots+k_m=n} \frac{n!}{k_1!k_2!\ldots k_m!} x_1^{k_1} x_2^{k_2} \ldots x_m^{k_m} .$$

(The k_i's in the sum are *non-negative*.)

4.5. COMBINATIONS

Let the set S have n elements and let $k > 0$ be an integer. We have already considered the question of the number of k-subsets of S and even have a notation for it: $\binom{n}{k}$. Another notation for this is $C(n,k)$. Any k-subset of S is called a *k-combination* of S. We have observed that when $k > n$, $C(n,k) = 0$, and have proved that when $0 \le k \le n$,

$$\boxed{C(n,k) \equiv \binom{n}{k} = \frac{n!}{k!(n-k)!}}$$

Example 28. All the 2-combinations of the set $S = \{a,b,c,d\}$ are: $\{a,b\},\{a,c\},\{a,d\}$, $\{b,c\},\{b,d\}$, and $\{c,d\}$. Their number is, as we know $C(4,2) = 4!/2!2! = 6$.

Example 29. Given a basketball team of 10 players, in how many ways we can pick a starting lineup? Obviously, it is $C(10,5) = 252$.

Note that our discussion of the binomial coefficients, Pascal's Triangle, and the Binomial Theorem are all very relevant to the study of k-combinations. Whereas in many proofs of these results we used mathematical induction, here we are more interested in what is known as ***combinatorial proofs*** which, usually, give more insight into the reasons why a particular result is true.

Example 30. [Example 25, revisited] Recall that in Example 25 we calculated the number paths from the $(0,0)$ cell to the $(m-1, n-1)$ cell in the planar grid, see the Figure on page 13. We have shown that it is precisely equal to the binomial coefficient $\binom{m+n}{n}$. In the argument given in the solution, we have stated that the different paths differ only in where the path goes north and when it moves east. Hence, each such path can be represented as a sequence (array) of $n+m$ entries where each entry is either "north" or "east"; the number of such sequences is therefore equal to the number of ways to choose n locations for "north"'s in such an array. This is simply "$n+m$ choose n", as claimed.

Example 31. Let $n \geq m \geq k \geq 0$ be integers. Show, $\boxed{\binom{n}{m}\binom{m}{k} = \binom{n}{k}\binom{n-k}{m-k}}$.

Solution. This identity can be easily proved using the definition of the binomial coefficients. However, a combinatorial proof is instructive. For that we need to find is some quantity *that is counted in two different ways on the two sides of the identity*. In this case we count the number of ways to choose an *m*-subset (of the larger *n*-set) and a *k*-subset of the chosen *m*-subset. This precisely what is counted on the left-hand side of the given identity. However, we can count the same thing differently:

(a) first pick a *k*-subset of the larger *n*-set, and then
(b) pick a *m*-superset of the *k*-subset.

Task (a) can be performed in $\binom{n}{k}$ ways. Once these *k* elements are removed from further consideration and only $n-k$ elements are subject to further possible choice, we need to pick $m-k$ more elements (in order to make up the required *m*-subset). This latter task can be performed in $\binom{n-k}{m-k}$ ways.

Example 32. Let m, n, k be natural numbers. We give a *combinatorial proof* of the following identity:

$$\boxed{\binom{m+n}{k} = \binom{m}{0}\binom{n}{k} + \binom{m}{1}\binom{n}{k-1} + \ldots + \binom{m}{k}\binom{n}{n} = \sum_{i=0}^{k}\binom{m}{i}\binom{n}{k-i}}.$$

Take an arbitrary (but fixed) set *S* of $m+n$ elements. Partition it arbitrarily into two disjoint subsets *A*, containing *m* elements, and *B*, containing *n* elements. Now we want to pick *k* elements (this is represented by the quantity on the left-hand side of the equality). To account for the right-hand side we can proceed as follows:

1. we can choose no elements from *A* and *k* elements from *B* – this can be done in $\binom{m}{0}\binom{n}{k}$ ways; or
2. we can choose one element from *A* and $k-1$ elements from *B* – this can be done in $\binom{m}{1}\binom{n}{k-1}$ ways;

....

....

3. we can choose k elements from A and no elements from B – this can be done in $\binom{m}{k}\binom{n}{0}$ ways; ∎

4.6. PLACING BALLS IN BOXES

In this section we study the number of ways of placing n balls into b distinct boxes under various restrictions on the identity of the objects and the way that the balls are place in the boxes.

1. *The n balls are distinct and their order inside the boxes doesn't matter*: since there are no restrictions on the placement of the balls, each one of the n balls can be placed in b ways giving a total of $\boxed{\underbrace{b \cdot b \cdot \ldots \cdot b}_{n \text{ times}} = b^n}$.
2. *The n balls are distinct and the balls placed in each box are ordered*: we think about such a placement as a permutation of n distinct objects (the n balls) and $b-1$ identical objects, called "walls" (denoted by the letter w). The reason that these are equivalent ways of counting is that (given that the boxes are distinct) we can arbitrarily designate them as box 1, box 2, etc. Then, given any distribution of the n balls, in which the balls in each box are ordered, we can first list the balls in box 1, in the given order, then all the balls in the second box, in the given order, and so on; to distinguish between the balls in the different boxes, we put a wall between the groups of ball placed in different consecutive boxes. To separate the sequence into b boxes we need $b-1$ walls. For example, with $n=6$ and $b=3$, a distribution:

   box 1: 3, 5, 1, 2;   box 2: *empty*;   box 3: 6, 4 .

   corresponds uniquely to the permutation (with w standing for the wall symbol):

   3, 5, 1, 2, w, w, 6, 4 .

   The idea, of course, is that the subsequences between two consecutive w's corresponds to a single box and the assumption that the boxes are distinct is reflected in their (arbitrary but fixed) ordering from left to right. Conversely, any permutation of the six distinct balls and two identical w's corresponds to a distribution of distinct balls in three distinct boxes with balls in each box ordered.

For example, the permutation: 1, w, 5, 3, w, 2, 6, 4 corresponds (uniquely) to the following distribution

<u>box 1</u>: 1 ; <u>box 2</u>: 5, 3 ; <u>box 3</u>: 2, 6, 4 .

In the general case we have $n+b-1$ object of which $b-1$ are identical. The $n+b-1$ objects can be permuted in $(n+b-1)!$ ways (assuming, for a moment that the $b-1$ walls are distinct. But these walls can be permuted in $(b-1)!$ ways without changing the actual placement of the balls, hence the answer is

$$\boxed{\frac{(n+b-1)!}{(b-1)!}}\,.$$

3. <u>*The n balls are identical and their order inside the boxes doesn't matter*</u>: Follows from the reasoning of part 2. There the walls were identical and we divided by $(b-1)!$ to account for that. Now the balls are identical as well and, by similar argument we need to divide the count by $n!$ Hence the result is

$$\boxed{\frac{(n+b-1)!}{(b-1)!n!} = \binom{n+b-1}{n}}\,.$$

4. <u>*The n balls are identical and no box contains more than 1 ball*</u>: This amounts to choosing n boxes out of b available (into which we distribute the n identical balls). This is simply the binomial coefficient $\binom{b}{n}$. Note that if the number of balls is larger than the number of boxes, $n > b,$ that we obtain 0 which is the correct answer.

5. <u>*The n balls are identical but no box can be left empty*</u>: Since the balls are identical, we take b balls and put one in each box. This can be done in one way if $n \ge b,$ and 0 ways otherwise. Now we have $n-b$ identical balls and there are no longer any restrictions on their distribution. This is precisely case 3 and we can use the formula obtained there:

$$\boxed{\binom{(n-b)+b-1}{n-b} = \binom{n-1}{n-b} = \binom{n-1}{b-1}}\,.$$

4.7. COUNTING PARTITIONS

Consider the task of counting the partitions of the set $[n] = \{1, 2, ..., n\}$ (or any other set of *n* elements). This problem is very interesting but turns out to be quite involved. The *Bell*[1] *number* B_n is the number of partitions of the set $[n]$ into non-empty subsets. For example, $B_3 = 5$ because there are 5 partitions of $[3] = \{1, 2, 3\}$:

$$\{\{1,2,3\}\}; \quad \{\{1,2\},\{3\}\}; \quad \{\{1,3\},\{2\}\}; \quad \{\{1\},\{2,3\}\}; \quad \{\{1\},\{2\},\{3\}\}.$$

The Bell numbers satisfy an interesting inductive (recursive) specification

Claim 33. The Bell numbers satisfy $B_0 = B_1 = 1$ and $\boxed{\forall n \geq 1,\ B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_{n-k}}$.

Proof. Consider the object $n+1$. For each $0 \leq k \leq n$, $n+1$ can be joined by *k* other elements to form a block in the partition. These *k* elements can be chosen in $\binom{n}{k}$ ways. The remaining blocks can be partitioned in any one of the B_{n-k} ways (independently of the block that contains $n+1$). ■

Remark. Interestingly, there is no known closed-form expression for B_n. It is possible to show that

$$B_n = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!}.$$

Also, we can define an "exponential generating function," which is an infinite series in which the B_n's are the coefficients: $\sum_{n=0}^{\infty} B_n \frac{x^n}{n!}$. Amazingly, it turns out that this infinite series actually is equal to the function $e^{e^x - 1}$.

Consider now counting only those partitions of $[n] = \{1, 2, ..., n\}$ that have exactly *k* parts, i.e. each partition consists of *k* disjoint subsets. We denote this number by $S(n,k)$. Looking back at the set of partitions of $[3]$:

---

[1] **Eric Temple Bell** (1883, Scotland - 1960, Watsonville, California) was a mathematician and science fiction author.

$$\{\{1,2,3\}\}; \quad \{\{1,2\},\{3\}\}; \quad \{\{1,3\},\{2\}\}; \quad \{\{1\},\{2,3\}\}; \quad \{\{1\},\{2\},\{3\}\}$$

we see that $S(3,1)=1$, $S(3,2)=3$, and $S(3,3)=1$. Clearly, since every partition of $[n]$ contains some (at least 1 and at most *n*) number of parts,

$$\boxed{B_n = \sum_{k=1}^{n} S(n,k)}.$$

The numbers $S(n,k)$ are called *Stirling*[2] *numbers of the second kind*.

Theorem 34. Let $n \ge k \ge 1$ be integers. Then, $S(n,1)=S(n,n)=1$ and

$$\boxed{S(n,k) \ = \ S(n-1,k-1)+kS(n-1,k)}.$$

Proof. The claim $S(n,1)=S(n,n)=1$ is easy: the only partition with one part is $\{[n]\}$ and the only partition with *n* parts is $\{\{i\} \mid i\in[n]\}$. For the general case, consider the element *n* and split all the partitions of $[n]$ with exactly *k* parts into 2 disjoint subsets: (a) those in which $\{n\}$ is a part, and (b) those in which *n* is with other elements. The number of partitions of $[n]$ of type (a) is clearly $S(n-1,k-1)$. To count the number of partitions of type (b), consider the set of all partitions of $[n-1]$ with exactly *k* parts. From each such partition we can generate *k* distinct partitions of $[n]$ (with exactly *k* parts) by adding *n* into any of the *k* available parts. Thus, the number of partitions of type (b) is $kS(n-1,k)$ and the formula in the claim is proved. ■

Thanks to the inductive formula of Theorem 34, we can construct a “Stirling triangle” similar to Pascal’s:

$$\begin{array}{c}
1 \\
1 \quad 1 \\
1 \quad 3 \quad 1 \\
1 \quad 7 \quad 6 \quad 1 \\
1 \quad 15 \quad 25 \quad 10 \quad 1 \\
1 \quad 31 \quad 90 \quad 65 \quad 15 \quad 1
\end{array}$$

[2] **James Stirling** (1692-1770) was a Scottish mathematician whose most important work *Methodus Differentialis* in 1730 is a treatise on infinite series, summation, interpolation and quadrature.

<u>EXERCISES</u>

(01) [Some easy counting.]

(a) Four new students have to be assigned a tutor. There are seven possible tutors and none of them will accept more than one student. In how many ways can the tutor assignment be done?

(b) Five Professors of Anthropology wish to build their new homes on the Triangle Island whose shape is a precise equilateral triangle with each side being two miles long. Is it possible for them to find sites on the island so that their houses are **more than one mile apart**. Briefly explain your reasons.

(c) The rules for the ISU intramural 5-on-5 soccer competition specify that the members of each team must have their birthdays in the same month. How many Computer Science students are needed in order to **guarantee** that they can enter a team into the competition?

(d) Suppose that the curriculum committee of the Sociology department at NU decided, for administrative convenience, that each freshman must take precisely **four** out of the seven courses available to freshmen. The teachers of the seven courses report the following numbers of registrants: 52, 30, 30, 20, 25, 12, and 18. Is there a problem with counting or registration? Explain.

(e) In Dr. Frank Jones' "Algorithms" class, 32 of the students are boys. Each boy knows 5 of the girls in the class and each girl knows 8 of the boys. "Knowing" is assumed to be mutual. How many girls are in the class?

(f) Keys are made by cutting incisions of various depths in a number of **positions** in a blank key. Suppose there are eight possible depths of incision. We need to be able to make one million different key types; for simplicity of calculation assume that we need at least 2^{20} different keys (which is just a little over a million). How many different positions are required and sufficient?

(02) We have seen in class that the number of ways to seat n people around a round table is $(n-1)!$ How many ways we can seat n men and n women around a round table if every woman must sit between two men (and vice versa). Explain.

(03) In how many ways it is possible to seat $n+m$ people around a round table that can seat n of them and on a bench that can seat the remaining m of them? Explain.

(04)[Lewis Carroll[3]] In a terrible battle all the soldiers were wounded. 70% lost an arm (A), 50% lost an eye (E), 40% lost a leg (L), 30% lost an arm and an eye, 20% lost an arm and a leg, and 20% lost an eye and a leg. What percent of the soldiers lost an arm, an eye, and a leg?

(05) Let $X \triangleq \{n \in \mathbb{Z} \,|\, 1 \le n \le 6{,}000{,}000\}$.

(a) How many elements of X are not divisible by 2 or 3?
(b) How many elements of X are not divisible by 2, 3 or 5?

(06) (a) Suppose that there are (i) 200 faculty members who speak French, (ii) 50 that speak Russian, (iii) 100 that speak Spanish, (iv) 20 that speak French and Russian, (v) 60 that speak French and Spanish, (vi) 35 that speak Russian and Spanish, and (vii) 10 that speak French, Russian and Spanish. How many faculty members speak either French, or Russian or Spanish?

(b) Consider the set of all 4-digit decimal numbers, from 0000 to 9999. How many of these do not have i in the ith position, i = 1,2,3,4?

[For instance, the number 8238 has 2 in the second position and therefore should not be counted. On the other hand, the number 5555 should be counted.]

(07) Prove two proofs of: $\boxed{\binom{n+1}{m+1} = \sum_{k=m}^{n} \binom{k}{m}}$. (a) one using repeatedly the ***Pascal identity*** $\binom{n+1}{m+1} = \binom{n}{m+1} + \binom{n}{m}$, and (b) use a purely combinatorial argument.

(08) (a) What is the number of permutations of the word: ATANASOFF?
(b) What is the number of permutations of the word: BBBBB?
(c) How many 11-letter words can be constructed using all and only the letters of the word CADABRAABRA?

(09) [Gabriel Carroll] We have k switches arranged in a row, each of which can point left, right, up or down. Whenever three successive switches point in distinct directions, all three can be simultaneously turned to point in the fourth direction (in a single "superswitch" operation). For example, let $k = 4$. A possible configuration could be: ***uldd***. Since the first three switches point in different directions, the superswitch operation

[3] **Charles Lutwidge Dodgson** (1832 – 1898), better known by the pseudonym **Lewis Carroll**, was an English author, mathematician, logician, Anglican deacon and a photographer. His most famous writings are *Alice's Adventures in Wonderland* and its sequel *Through the Looking-Glass*, as well as the poems "The Hunting of the Snark" and "Jabberwocky", all examples of the genre of literary nonsense.

will result a new configuration: ***rrrd***. We want to prove that no matter what the initial configuration is, this operation cannot be repeated infinitely many times.

For each configuration of switches define the "height" of the configuration to be the product of all the values *n* such that switches $n-1$ and n point in the same direction, or 1 if there is no such *n*. E.g., the height of configuration ***uldd*** is 4, whereas the height of configuration ***rrrd*** is $2 \cdot 3 = 6$.

(a) Prove that each superswitch operation increases the height.
(b) Give an upper bound on the height of any switch configuration.
(c) Use (a) and (b) to conclude that the superswitch operation cannot be repeated infinitely many times.

(10) [Chess] (a) In how many different ways can we place three distinct chess pieces (say a pawn (p), a knight (k), and a rook (r)) on an 8×8 chessboard so that no two of the pieces occupy the same row or column?

(b) In how many ways can we place two identical chess pieces (say two white rooks) on an 8×8 chessboard so that they do not occupy the same row or column?

(11) Prove the Multinomial Theorem 27.

(12) Use the fact that the multinomial coefficients are non-negative integers to prove that

(a) $2^n \mid (2n)!$ and that for $n \geq 2$ the quotient is even.
(b) $(n!)^{n+1} \mid (n^2!)!$

(13) Prove that for all integers $0 \leq a,b,c$ with $a+b+c=n$,

$$\binom{n}{a,b,c} = \binom{n-1}{a-1,b,c} + \binom{n-1}{a,b-1,c} + \binom{n-1}{a,b,c-1}.$$

(14) (a) How many solutions the equation $x+y+z+u=10$ has in non-negative integers $x,y,z,u \geq 0$. Explain.

(15) In a souvenir store there are *k* kinds of postcards (unlimited supply of each kind). You want to send postcards to *n* friends.

(a) In how many different ways you can do it?
(b) Like in (a), but you want to send them different cards?
(c) What if you want to send two different cards to each of your friends (but different friends can get the same cards)?

(16) Use combinatorial argument to prove the identity

$$\boxed{\sum_{1\le k\le n}\binom{n}{k}k!kn^{n-k-1} \;=\; n^n}.$$

(17) Prove that $\forall n \ge 0\ \exists! c > b > a \ge 0$ (all integers) such that $n = \binom{a}{1} + \binom{b}{2} + \binom{c}{3}$.

(18) In reference to Stirling numbers of the second kind, $S(n,k)$, solve the following.

(a) Compute all the six partitions of the set $[4] = \{1,2,3,4\}$ into 3 parts. As an example, one such partition is $\pi_1 = \{\{1,2\},\{3\},\{4\}\}$.

(b) Calculate the next line of the "Stirling triangle" given in the text.

(c) Prove that $S(n,2) = 2^{n-1} - 1$.

(d) Prove that $S(n,n-1) = \binom{n}{2}$.

(e) Prove that $S(n,3) = \frac{1}{2}(3^{n-1}+1) - 2^{n-1}$.

(f) Show that $S(n,k) = \dfrac{1}{k!}\sum \binom{n}{n_1, n_2, ..., n_k}$, where the sum is taken over all k-tuples $(n_1, n_2, ..., n_k)$ of positive integers such that $n_1 + n_2 + ... + n_k = n$.

(g) Show that if the sum in (f) is taken over all k-tuples $(n_1, n_2, ..., n_k)$ of non-negative integers such that $n_1 + n_2 + ... + n_k = n$, then the sum is equal to k^n.

(19) Denote $\mathfrak{I} \triangleq \left\{ f \,\middle|\, f : [n] \xrightarrow{onto} [k] \right\}$, the set of all function from an n-set $[n]$ to k-set $[k]$. Prove that: $\boxed{|\mathfrak{I}| \;=\; k!S(n,k)}$.