4

Inclusion-Exclusion and Related Techniques

This chapter studies combinatorial techniques that are related to the arithmetic operation of subtraction: involutions, inclusion-exclusion formulas, and Möbius inversion. Involutions allow us to give bijective proofs of identities involving both positive and negative terms. The inclusion-exclusion formula extends the sum rule 1.2 to a rule for computing $|A_1 \cup A_2 \cup \cdots \cup A_m|$ in the case where the sets A_i need not be pairwise disjoint. This formula turns out to be a special case of the general Möbius inversion formula for posets, which has many applications in number theory and algebra as well as combinatorics.

4.1 Involutions

In Chapter 2, we saw how to use bijections to prove combinatorial identities. Many identities involve a mixture of positive and negative terms. One can use *involutions* to furnish combinatorial proofs of such identities. We illustrate the idea using the following binomial coefficient identity.

4.1. Theorem. For all $n \geq 1$, $\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0$.

Proof. The result can be proved algebraically by using the binomial theorem 2.14 to expand the left side of $(-1+1)^n = 0$. To prove the identity combinatorially, let X be the set of all subsets of $\{1, 2, \ldots, n\}$. For each $S \in X$, we define the *sign* of S to be $\operatorname{sgn}(S) = (-1)^{|S|}$. Since there are $\binom{n}{k}$ subsets S of size k, and $\operatorname{sgn}(S) = (-1)^k$ for all such subsets, we see that

$$\sum_{S\in X} \operatorname{sgn}(S) = \sum_{k=0}^{n}(-1)^k\binom{n}{k}.$$

Thus we have found a combinatorial model for the left side of the desired identity, which involves *signed objects.*

Now, define a function $I : X \to X$ as follows. Given $S \in X$, let $I(S) = S \cup \{1\}$ if $1 \notin S$, and let $I(S) = S \sim \{1\}$ if $1 \in S$. Observe that $I(I(S)) = S$ for all $S \in X$; in other words, $I \circ I = \mathrm{id}_X$. Thus, I is a bijection that is equal to its own inverse. Furthermore, since $|I(S)| = |S| \pm 1$, $\operatorname{sgn}(I(S)) = -\operatorname{sgn}(S)$ for all $S \in X$. It follows that I pairs each positive object in X with a negative object in X. Consequently, the number of positive objects in X equals the number of negative objects in X, and so $\sum_{S\in X} \operatorname{sgn}(S) = 0$. □

The general setup for involution proofs is described as follows.

4.2. Definition: Involutions. An *involution* on a set X is a function $I : X \to X$ such that $I \circ I = \mathrm{id}_X$. Equivalently, I is a bijection on X and $I = I^{-1}$. Given an involution I, the *fixed point set* of I is the set $\mathrm{Fix}(I) = \{x \in X : I(x) = x\}$, which may be empty. If $\operatorname{sgn} : X \to \{+1, -1\}$ is a function that attaches a sign to every object in X, we say that I is a *sign-reversing* involution (relative to sgn) iff for all $x \in X \sim \mathrm{Fix}(I)$, $\operatorname{sgn}(I(x)) = -\operatorname{sgn}(x)$.

4.3. Involution Theorem. Given a finite set X of signed objects and a sign-reversing involution I on X,

$$\sum_{x\in X}\text{sgn}(X)=\sum_{x\in\text{Fix}(I)}\text{sgn}(X).$$

Proof. Let $X^+=\{x\in X\sim\text{Fix}(I):\text{sgn}(x)=+1\}$ and $X^-=\{x\in X\sim\text{Fix}(I):\text{sgn}(x)=-1\}$. By definition, I restricts to X^+ and X^- to give functions $I^+:X^+\rightarrow X^-$ and $I^-:X^-\rightarrow X^+$ that are mutually inverse bijections. Therefore, $|X^+|=|X^-|$ and

$$\begin{aligned}\sum_{x\in X}\text{sgn}(X) &= \sum_{x\in X^+}\text{sgn}(x)+\sum_{x\in X^-}\text{sgn}(x)+\sum_{x\in\text{Fix}(I)}\text{sgn}(x)\\ &= |X^+|-|X^-|+\sum_{x\in\text{Fix}(I)}\text{sgn}(x)=\sum_{x\in\text{Fix}(I)}\text{sgn}(x).\end{aligned}$$ □

As a first illustration of the involution theorem, we prove a variation of 4.1.

4.4. Theorem. For all $n\geq 1$,

$$\sum_{k=0}^{n}(-1)^k\binom{2n}{k}=(-1)^n\binom{2n-1}{n}.$$

Proof. Let X be the set of all subsets of $\{1,2,\ldots,2n\}$ of size at most n, and let the sign of a subset T be $(-1)^{|T|}$. The left side of the desired identity is $\sum_{T\in X}\text{sgn}(T)$. Next, define an involution I on X as follows. If $T\in X$ and $1\in T$, let $I(T)=T\sim\{1\}$. If $T\in X$ and $1\notin T$ and $|T|<n$, let $I(T)=T\cup\{1\}$. Finally, if $T\in X$ and $1\notin T$ and $|T|=n$, let $I(T)=T$. One checks immediately that I is a sign-reversing involution. The fixed points of I are the n-element subsets of $\{1,2,\ldots,2n\}$ not containing 1. There are $\binom{2n-1}{n}$ such subsets, and each of them has sign $(-1)^n$. So $\sum_{T\in\text{Fix}(I)}\text{sgn}(T)$ is the right side of the desired identity. □

4.5. Theorem. For all $n\geq 0$,

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}^2=\begin{cases}0 & \text{if } n \text{ is odd;}\\ (-1)^{n/2}\binom{n}{n/2} & \text{if } n \text{ is even.}\end{cases}$$

Proof. Let X be the set of all pairs (S,T), where S and T are subsets of $\{1,2,\ldots,n\}$ of the same size. Define $\text{sgn}(S,T)=(-1)^{|S|}$. Then the left side of the desired identity is $\sum_{(S,T)\in X}\text{sgn}(S,T)$. We define an involution I on X as follows. Given $(S,T)\in X$, let i be the least integer in $\{1,2,\ldots,n\}$ (if there is one) such that either $i\notin S$ and $i\notin T$, or $i\in S$ and $i\in T$. In the former case, let $I(S,T)=(S\cup\{i\},T\cup\{i\})$; in the latter case, let $I(S,T)=(S\sim\{i\},T\sim\{i\})$; if no such i exists, let $I(S,T)=(S,T)$. It is routine to check that I is a sign-reversing involution; in particular, the designated integer i in the definition of $I(S,T)$ is the same as the i used to calculate $I(I(S,T))$. By the involution theorem,

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}^2=\sum_{(S,T)\in\text{Fix}(I)}(-1)^{|S|}.$$

Note that $(S,T)\in\text{Fix}(I)$ iff for every $i\leq n$, i lies in exactly one of the two sets S or T. This can only happen if n is even and $|S|=|T|=n/2$ and $S=\{1,2,\ldots,n\}\sim T$. So the fixed point set is empty if n is odd. If n is even, we can construct an arbitrary element of $\text{Fix}(I)$ by choosing any subset S of size $n/2$ and letting T be the complementary subset of $\{1,2,\ldots,n\}$. Since there are $\binom{n}{n/2}$ choices for S, each with sign $(-1)^{n/2}$, the formula in the theorem is proved. □

4.6. Example: Stirling Numbers. Recall that $s(n,k) = (-1)^{n-k}c(n,k)$, where $c(n,k)$ is the number of permutations of an n-element set whose functional digraph consists of k cycles (§3.6). We will show that

$$\sum_{k=1}^{n} s(n,k) = \chi(n=1) \qquad (n \geq 1).$$

Both sides are 1 when $n = 1$, so assume $n > 1$. Let X be the set of all permutations of $\{1,2,\ldots,n\}$. If $w \in X$ is a permutation with k cycles, define $\operatorname{sgn}(w) = (-1)^k$. Now $\sum_{w \in X} \operatorname{sgn}(w) = (-1)^n \sum_{k=1}^{n} s(n,k)$, so it suffices to define a sign-reversing involution I on X with no fixed points. Given $w \in X$, the numbers 1 and 2 either appear in the same cycle of w or in different cycles. If 1 and 2 are in the same cycle, let the elements on this cycle (starting at 1) be

$$(1, x_1, x_2, \ldots, x_k, 2, y_1, y_2, \ldots, y_j),$$

where $j, k \geq 0$. Define $I(w)$ by replacing this cycle by the two cycles

$$(1, x_1, x_2, \ldots, x_k)(2, y_1, y_2, \ldots, y_j)$$

and leaving all other cycles the same. Similarly, if 1 and 2 are in different cycles of w, write these cycles as

$$(1, x_1, x_2, \ldots, x_k)(2, y_1, y_2, \ldots, y_j)$$

and define $I(w)$ by replacing these two cycles by the single cycle

$$(1, x_1, x_2, \ldots, x_k, 2, y_1, y_2, \ldots, y_j).$$

It is immediate that $I \circ I = \mathrm{id}_X$, I is sign-reversing, and I has no fixed points.

We can modify the preceding involution to obtain a combinatorial proof of the identity

$$\sum_{k \geq 0} s(i,k)S(k,j) = \chi(i=j),$$

which we proved algebraically in part (d) of 2.77. If $i < j$, then for every k, either $s(i,k) = 0$ or $S(k,j) = 0$. So both sides of the identity are zero in this case. If $i = j$, the left side reduces to $s(i,i)S(i,i) = 1 = \chi(i = j)$. If $j = 0$, the identity is true. So we may assume i and j are fixed numbers such that $i > j > 0$. Let X be the set of pairs (w, U), where w is a permutation of $\{1,2,\ldots,i\}$ (viewed as a functional digraph) and U is a set partition of the set of cycles in w into j blocks. If w has k cycles, let $\operatorname{sgn}(w,U) = (-1)^k$. Then

$$\sum_{(w,U)\in X} \operatorname{sgn}(w,U) = (-1)^i \sum_{k=j}^{i} s(i,k)S(k,j)$$

and $\chi(i = j) = 0$. So it suffices to define a sign-reversing involution I on X with no fixed points. Given $(w,U) \in X$, there must exist a block of U such that the cycles in this block collectively involve more than one point in $\{1,2,\ldots,i\}$. This follows from the fact that i (the number of points) exceeds j (the number of blocks). Among all such blocks in U, choose the block that contains the smallest possible element in $\{1,2,\ldots,i\}$. Let this smallest element be a, and let the second-smallest element in this block be b. To calculate $I(w,U)$, modify the cycles in this block of U as we did above, with a and b playing the roles of 1 and 2. More specifically, a cycle of the form

$$(a, x_1, \ldots, x_k, b, y_1, \ldots, y_j)$$

gets replaced (within its block) by

$$(a, x_1, \ldots, x_k)(b, y_1, \ldots, y_j)$$

and vice versa. It is routine to check that I is a sign-reversing involution on X with no fixed points. For example, suppose $i = 10$, $j = 3$, w has cycles $(1), (3, 5), (2, 6, 9), (4, 8), (7), (10)$, and

$$U = \{\{(1)\}, \{(3, 5), (10)\}, \{(2, 6, 9), (4, 8), (7)\}\}.$$

Here the block of U modified by the involution is $\{(2, 6, 9), (4, 8), (7)\}$, $a = 2$, and $b = 4$. We compute $I(w, U)$ by replacing the cycles $(2, 6, 9)$ and $(4, 8)$ in w by the single cycle $(2, 6, 9, 4, 8)$ and letting the new set partition be

$$\{\{(1)\}, \{(3, 5), (10)\}, \{(2, 6, 9, 4, 8), (7)\}\}.$$

4.2 The Inclusion-Exclusion Formula

Recall the sum rule: if $S_1, \ldots, S_n$ are *pairwise disjoint* finite sets, then $|S_1 \cup \cdots \cup S_n| = \sum_{i=1}^{n} |S_i|$. Can we find a formula for $|S_1 \cup \cdots \cup S_n|$ in the case where the given sets S_i are not necessarily disjoint? The answer is provided by the inclusion-exclusion formula, which we discuss now.

We have already seen the simplest case of the inclusion-exclusion formula. Specifically, if S and T are any two finite sets, the binary union rule 1.4 states that

$$|S \cup T| = |S| + |T| - |S \cap T|.$$

Intuitively, the sum $|S| + |T|$ overestimates the cardinality of $|S \cup T|$ because elements of $|S \cap T|$ are included twice in this sum. To correct this, we exclude one copy of each of the elements in $S \cap T$ by subtracting $|S \cap T|$.

Now consider three finite sets S, T, and U. The sum $|S| + |T| + |U|$ overcounts the size of $|S \cup T \cup U|$ since elements in the overlaps between these sets are counted twice (or three times, in the case of elements $z \in S \cap T \cap U$). We may try to account for this by subtracting $|S \cap T| + |S \cap U| + |T \cap U|$ from $|S| + |T| + |U|$. If x belongs to S and U but not T (say), this subtraction will cause x to be counted only once in the overall expression. A similar comment applies to elements in $(S \cap T) \sim U$ and $(T \cap U) \sim S$. However, an element $z \in S \cap T \cap U$ is counted three times in $|S| + |T| + |U|$ and subtracted three times in $|S \cap T| + |S \cap U| + |T \cap U|$. So we must include such elements once again by adding the term $|S \cap T \cap U|$. In summary, we have given an informal argument suggesting that the formula

$$|S \cup T \cup U| = |S| + |T| + |U| - |S \cap T| - |S \cap U| - |T \cap U| + |S \cap T \cap U|$$

should be true.

Generalizing the pattern in the preceding example, we arrive at the following formula, known as the *inclusion-exclusion formula.*

4.7. Inclusion-Exclusion Formula. Suppose $n > 0$ and $S_1, \ldots, S_n$ are any finite sets. Then

$$|S_1 \cup S_2 \cup \cdots \cup S_n| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} |S_{i_1} \cap S_{i_2} \cap \cdots \cap S_{i_k}|. \tag{4.1}$$

4.8. Example. If $n = 4$, the inclusion-exclusion formula for $|S_1 \cup S_2 \cup S_3 \cup S_4|$ is

$$\begin{array}{l} |S_1| + |S_2| + |S_3| + |S_4| \\ -|S_1 \cap S_2| - |S_1 \cap S_3| - |S_1 \cap S_4| - |S_2 \cap S_3| - |S_2 \cap S_4| - |S_3 \cap S_4| \\ +|S_1 \cap S_2 \cap S_3| + |S_1 \cap S_2 \cap S_4| + |S_1 \cap S_3 \cap S_4| + |S_2 \cap S_3 \cap S_4| \\ -|S_1 \cap S_2 \cap S_3 \cap S_4|. \end{array}$$

4.9. Remark. By setting $I = \{i_1, i_2, \ldots, i_k\}$, the inclusion-exclusion formula can also be written

$$|S_1 \cup \cdots \cup S_n| = \sum_{\emptyset \neq I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|-1} \left| \bigcap_{i \in I} S_i \right|.$$

We will give several proofs of the inclusion-exclusion formula. Each proof illustrates different techniques and can be generalized in different ways.

4.10. Proof of Inclusion-Exclusion by Induction. We prove that (4.1) holds for all $n > 0$ and all finite sets $S_1, \ldots, S_n$ by induction on n. The formula reduces to $|S_1| = |S_1|$ for $n = 1$, and this is certainly true. For $n = 2$, the formula becomes

$$|S_1 \cup S_2| = |S_1| + |S_2| - |S_1 \cap S_2|,$$

and this is the binary union rule 1.4 proved previously. Now assume $n > 2$ and that formula (4.1) is already known to hold for any union of $n - 1$ finite sets. Let $S_1, \ldots, S_n$ be fixed finite sets. The n-fold union $S_1 \cup \cdots \cup S_n$ can be regarded as the union of the two sets $S = S_1 \cup S_2 \cup \cdots \cup S_{n-1}$ and $T = S_n$. Hence, by the binary union rule 1.4,

$$|S_1 \cup \cdots \cup S_n| = |S_1 \cup \cdots \cup S_{n-1}| + |S_n| - |(S_1 \cup \cdots \cup S_{n-1}) \cap S_n|.$$

Since the set operations $\cap$ and $\cup$ obey the distributive law, we can write the subtracted term as

$$|(S_1 \cap S_n) \cup (S_2 \cap S_n) \cup \cdots \cup (S_{n-1} \cap S_n)|,$$

which is the union of the $n - 1$ finite sets $S_i \cap S_n$ $(1 \leq i \leq n - 1)$. So we can apply the induction hypothesis to this term, and to the first term $|S_1 \cup \cdots \cup S_{n-1}|$. We obtain

$$\begin{aligned} |S_1 \cup \cdots \cup S_n| &= |S_n| + \sum_{k=1}^{n-1} (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n-1} |S_{i_1} \cap \cdots \cap S_{i_k}| \\ &\quad - \sum_{j=1}^{n-1} (-1)^{j-1} \sum_{1 \leq i_1 < \cdots < i_j \leq n-1} |(S_{i_1} \cap S_n) \cap \cdots \cap (S_{i_j} \cap S_n)|. \end{aligned}$$

We modify the second line of this formula as follows. First, observe that

$$\bigcap_{r=1}^{j} (S_{i_r} \cap S_n) = S_{i_1} \cap S_{i_2} \cap \cdots \cap S_{i_j} \cap S_n.$$

Next, change the summation index by setting $k = j + 1$ and defining $i_k = n$. The formula now reads

$$\begin{aligned} |S_1 \cup \cdots \cup S_n| &= \sum_{k=1}^{n-1} (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k < n} |S_{i_1} \cap \cdots \cap S_{i_k}| \\ &\quad + |S_n| + \sum_{k=2}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_{k-1} < i_k = n} |S_{i_1} \cap \cdots \cap S_{i_k}|. \end{aligned}$$

We can absorb $|S_n|$ into the sum on the second line by allowing k to range from 1 to n there. Also, letting k range from 1 to n in the first summation does not introduce any new terms. After making these adjustments, the only difference between the formulas on the first and second lines is that $i_k < n$ in the first line while $i_k = n$ in the second line. We can now combine the two summations to obtain

$$|S_1 \cup \cdots \cup S_n| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \le i_1 < \cdots < i_k \le n} |S_{i_1} \cap \cdots \cap S_{i_k}|, \tag{4.2}$$

which is the desired formula (4.1). This completes the induction.

In some counting problems, the following versions of the inclusion-exclusion formula are needed.

4.11. Alternate Version of Inclusion-Exclusion Formula. Suppose $S_1, \ldots, S_n$ are subsets of a finite set X. The number of elements $x \in X$ that belong to *none* of the S_i is

$$|X \sim (S_1 \cup \cdots \cup S_n)| = |X| + \sum_{k=1}^{n} (-1)^k \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} |S_{i_1} \cap S_{i_2} \cap \cdots \cap S_{i_k}|.$$

This formula follows from the original inclusion-exclusion formula and the difference rule 1.3.

Intuitively, the preceding formula is applicable when we are trying to count objects in X that must simultaneously avoid a number of specified "bad" properties. Each set S_i consists of those objects in X that have the ith bad property (and possibly other bad properties too).

4.12. Simplified Version of the Inclusion-Exclusion Formula. Let $S_1, \ldots, S_n$ be finite sets. Suppose that for all $k \ge 1$, the intersection of any k distinct sets among the S_j's always has cardinality $N(k)$. In other words, $|S_{i_1} \cap S_{i_2} \cap \cdots \cap S_{i_k}| = N(k)$ for all choices of $i_1 < i_2 < \cdots < i_k$. Then

$$|S_1 \cup \cdots \cup S_n| = \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} N(k).$$

If all S_j's are subsets of a given finite set X, we also have

$$|X \sim (S_1 \cup \cdots \cup S_n)| = |X| + \sum_{k=1}^{n} (-1)^k \binom{n}{k} N(k).$$

These formulas follow by substituting $N(k)$ for each summand $|S_{i_1} \cap \cdots \cap S_{i_k}|$ in the previous inclusion-exclusion formulas and noting that there are $\binom{n}{k}$ such summands.

4.3 More Proofs of Inclusion-Exclusion

This section presents two proofs of the inclusion-exclusion formula that are more combinatorial than the inductive computation already given.

4.13. Involution Proof of Inclusion-Exclusion. If we move all terms in (4.1) to the left side, we obtain the formula

$$\sum_{k=0}^{n}(-1)^k \sum_{1\le i_1<\cdots<i_k\le n} |S_{i_1}\cap\cdots\cap S_{i_k}| = 0. \tag{4.3}$$

In this equation, the summand corresponding to $k=0$ is defined to be $|S_1\cup\cdots\cup S_n|$. We will prove this formula by introducing an involution on a suitable set of signed objects.

Let X be the set of all sequences $(x;i_1,i_2,\ldots,i_k)$ such that $1\le i_1<i_2<\cdots<i_k\le n$, $0\le k\le n$, and $x\in S_{i_1}\cap\cdots\cap S_{i_k}$. (If $k=0$, then the object looks like $(x;)$, and the last condition is interpreted to mean $x\in S_1\cup\cdots\cup S_n$.) Define $\operatorname{sgn}(x;i_1,i_2,\ldots,i_k)=(-1)^k$. It follows from the sum rule that $\sum_{z\in X}\operatorname{sgn}(z)$ is the left side of (4.3). So it suffices to define a sign-reversing involution on X with no fixed points.

Given $z=(x;i_1,\ldots,i_k)\in X$, we must have $x\in S_1\cup S_2\cup\cdots\cup S_n$ no matter what the value of k is. Let i be the minimum index in $\{1,2,\ldots,n\}$ such that $x\in S_i$. By definition of X, we either have $k=0$ or $i<i_1$ or $i=i_1$. If $k=0$ or $i<i_1$, define $I(z)=(x;i,i_1,i_2,\ldots,i_k)$. If instead $i=i_1$, define $I(z)=(x;i_2,\ldots,i_k)$. It is immediate that $I(I(z))=z$ and $\operatorname{sgn}(I(z))=-\operatorname{sgn}(z)$ for all $z\in X$.

The preceding proof is quite ingenious, since it establishes a rather complicated formula by a remarkably simple bookkeeping bijection. On the other hand, it would be nice to have a combinatorial proof of inclusion-exclusion that is tied more closely to the intuitive "including and excluding" arguments we used originally to guess the formula for $|S\cup T\cup U|$. We present such a proof next.

4.14. Counting Proof of Inclusion-Exclusion. Fix n finite sets $S_1,\ldots,S_n$, and put $X=S_1\cup\cdots\cup S_n$. We consider a large matrix A whose rows are indexed by the elements $x\in X$ and whose columns are indexed by all nonempty subsets T of $\{1,2,\ldots,n\}$. Define the entry in row x and column T of A to be $(-1)^{|T|-1}$ iff $x\in\bigcap_{i\in T}S_i$, and define this entry to be zero otherwise. The sum of the entries in the column of A indexed by $T=\{i_1<i_2<\cdots<i_k\}\subseteq\{1,2,\ldots,n\}$ is

$$(-1)^{k-1}|S_{i_1}\cap\cdots\cap S_{i_k}|.$$

Adding up all these column sums, we see that the sum s of all entries in A is

$$s=\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\le i_1<i_2<\cdots<i_k\le n}|S_{i_1}\cap\cdots\cap S_{i_k}|.$$

Now, let us compute s by adding up the row sums of A. Intuitively, the sum of the 1's and -1's in row x of A represents the net number of times x has been counted in the inclusion-exclusion sum written above. We claim that this number is 1 for all $x\in X$, so that the sum of the row sums is $s=|X|=|S_1\cup\cdots\cup S_n|$. This will complete the proof of the inclusion-exclusion formula.

Fix $x\in X$, and let $U=\{i_1<i_2<\cdots<i_m\}$ be the set of all indices i_j such that $x\in S_{i_j}$. We have $m>0$ since x lies in at least one S_i. The entry in row x and column T of A is $(-1)^{|T|-1}$ if $T\subseteq U$, and this entry is zero if T is not a subset of U. Note that there are $\binom{m}{k}$ subsets of U of size k, each of which contributes $(-1)^{k-1}$ to the row sum. Grouping all such terms together and invoking the binomial theorem, we conclude that the sum of the entries in row x of A is

$$\sum_{k=1}^{m}\binom{m}{k}(-1)^{k-1}=1-\sum_{k=0}^{m}\binom{m}{k}(-1)^k1^{m-k}=1-(-1+1)^m=1.$$

4.4 Applications of the Inclusion-Exclusion Formula

We can use the inclusion-exclusion formula to count complicated combinatorial collections that cannot be conveniently enumerated by the sum and product rules alone. Recall that, when using inclusion-exclusion, we often set up the problem so that each set S_i consists of those objects in some big set X that have a certain "bad" property. Our desired answer is then the cardinality of $X \sim (S_1 \cup \cdots \cup S_n)$, which is given by the inclusion-exclusion formula in 4.11.

4.15. Example: Bridge Hands. A *bridge hand* is a 13-element subset of a 52-card deck. A *face card* is a jack, queen, king, or ace. How many bridge hands have at least one of each kind of face card? To answer this question, let X be the set of all bridge hands; note $|X| = \binom{52}{13}$. Define S_1 (resp. S_2, S_3, S_4) to be the set of all hands in X that do *not* have a jack (resp. queen, king, ace). The card hands we want are the elements of the set $X \sim (S_1 \cup S_2 \cup S_3 \cup S_4)$. We must now compute the sizes of the various intersections $S_{i_1} \cap \cdots \cap S_{i_k}$. Note that $|S_1| = \binom{48}{13}$ since we can build all hands in S_1 by choosing 13 cards out of the 48 non-jacks in the deck. Similarly, $|S_2| = |S_3| = |S_4| = \binom{48}{13}$. Next, $|S_1 \cap S_3| = \binom{44}{13}$ since we can build hands in $S_1 \cap S_3$ by choosing 13 cards out of the 44 cards in the deck that are neither jacks nor kings. The same formula holds for all other twofold intersections. Similarly, each threefold intersection has size $\binom{40}{13}$, while $|S_1 \cap S_2 \cap S_3 \cap S_4| = \binom{36}{13}$. It follows from inclusion-exclusion that the answer to the original question is

$$\binom{52}{13} - 4\binom{48}{13} + 6\binom{44}{13} - 4\binom{40}{13} + \binom{36}{13} = 128,971,619,088.$$

Next, how many 13-card bridge hands have at least one jack, at least one queen, and at least one king, but do not contain any ace cards or spade cards? The last condition can be dealt with as follows: throw out the $13 + 4 - 1 = 16$ aces and spades at the outset, leaving $52 - 16 = 36$ cards. An inclusion-exclusion argument like the one in the last paragraph now leads to the answer

$$\binom{36}{13} - 3\binom{33}{13} + 3\binom{30}{13} - \binom{27}{13} = 930,511,530.$$

4.16. Example. How many words $w \in X = \mathcal{R}(1^2 2^2 3^2 \cdots n^2)$ never have two adjacent letters that are equal? Note first that $|X| = \binom{2n}{2,2,\ldots,2} = (2n)!/2^n$. For $1 \leq i \leq n$, let S_i be the set of words in X in which the two copies of letter i are adjacent to each other. We wish to count the words in $X \sim (S_1 \cup \cdots \cup S_n)$. To do so, fix $i_1 < i_2 < \cdots < i_k$ and consider a typical intersection $S_{i_1} \cap \cdots \cap S_{i_k}$. Given a word w in this intersection, form a new word by replacing the two consecutive copies of i_j by a single copy of i_j, for $1 \leq j \leq k$. This operation defines a bijection from $S_{i_1} \cap \cdots \cap S_{i_k}$ onto the set $\mathcal{R}(1^{a_1} 2^{a_2} \cdots n^{a_n})$, where $a_i = 1$ if $i = i_j$ for some j, and $a_i = 2$ otherwise. (The inverse bijection replaces each i_j by two consecutive copies of i_j.) It follows that

$$|S_{i_1} \cap \cdots \cap S_{i_k}| = \binom{1k + 2(n-k)}{\underbrace{1,\ldots,1}_{k},\underbrace{2,\ldots,2}_{n-k}} = (2n-k)!/2^{n-k}.$$

This expression does not depend on the indices $i_1, \ldots, i_k$. Also, when $k = 0$, this expression reduces to $|X|$. Using the simplified inclusion-exclusion formula 4.12, we conclude that

$$|X \sim (S_1 \cup \cdots \cup S_n)| = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{(2n-k)!}{2^{n-k}}.$$

For our next examples, we use inclusion-exclusion to enumerate certain combinatorial collections that have arisen in earlier chapters.

4.17. Theorem: Enumeration of Surjections. Let $\mathrm{Surj}(m,n)$ be the number of surjections from an m-element set onto an n-element set. If $m \geq n \geq 1$, then

$$\mathrm{Surj}(m,n) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)^m.$$

Proof. Let X be the set of all functions $f : \{1,2,\ldots,m\} \to \{1,2,\ldots,n\}$. Note that $|X| = n^m$. For $1 \leq i \leq n$, let S_i consist of all functions $f \in X$ such that i is *not* in the image of f. A function $f \in X$ is a surjection iff f belongs to none of the S_i. Thus, we must compute $|X \sim (S_1 \cup \cdots \cup S_n)|$. Consider a typical intersection $S_{i_1} \cap \cdots \cap S_{i_k}$, where $i_1 < i_2 < \cdots < i_k$. A function f belonging to this intersection is the same thing as an arbitrary function mapping $\{1,2,\ldots,m\}$ into the $(n-k)$-element set $\{1,2,\ldots,n\} \sim \{i_1,i_2,\ldots,i_k\}$. The number of such functions is $(n-k)^m$, independent of $i_1,\ldots,i_k$. Using the simplified inclusion-exclusion formula 4.12, we get

$$\mathrm{Surj}(m,n) = |X \sim (S_1 \cup \cdots \cup S_n)| = n^m + \sum_{k=1}^{n} (-1)^k \binom{n}{k} (n-k)^m,$$

which is equivalent to the formula of the theorem. □

Since $S(m,n) = \mathrm{Surj}(m,n)/n!$ by 2.58, we deduce the following formula for Stirling numbers of the second kind.

4.18. Theorem: Summation Formula for Stirling Numbers of the Second Kind.

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

Our next illustration of inclusion-exclusion comes from number theory.

4.19. Definition: Euler's ϕ Function. For each integer $m \geq 1$, let $\phi(m)$ be the number of integers $x \in \{1,2,\ldots,m\}$ such that $\gcd(x,m) = 1$.

For example, if $m = 12$, then the relevant integers x are 1, 5, 7, and 11, so $\phi(12) = 4$. The function ϕ is prominent in algebra and number theory and has applications to modern cryptography.

4.20. Theorem: Formula for $\phi(m)$. Suppose an integer $m > 1$ has prime factorization $m = p_1^{e_1} p_2^{e_2} \cdots p_n^{e_n}$. Then

$$\phi(m) = \prod_{i=1}^{n} p_i^{e_i - 1}(p_i - 1) = m \prod_{i=1}^{n} (1 - 1/p_i).$$

Proof. Let $X = \{1,2,\ldots,m\}$, and let $S_i = \{x \in X : p_i | x\}$. (The symbol $p_i|x$ means that p_i divides x.) By the fundamental theorem of arithmetic, $x \in X$ is not relatively prime to m iff x and m have a common factor greater than 1 iff x and m have a common *prime* factor. It follows that

$$\phi(m) = |X \sim (S_1 \cup S_2 \cup \cdots \cup S_n)|.$$

So we are in a position to use inclusion-exclusion. Here it is convenient to write the inclusion-exclusion formula as follows:

$$|X \sim (S_1 \cup \cdots \cup S_n)| = \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|} \left| \bigcap_{i \in I} S_i \right|,$$

where we interpret $\bigcap_{i\in\emptyset} S_i$ as the set X. Fix a subset $I = \{i_1 < \cdots < i_k\} \subseteq \{1,2,\ldots,n\}$, and consider the intersection $S_{i_1} \cap \cdots \cap S_{i_k}$. An integer $x \leq m$ lies in this intersection iff $p_{i_j}|x$ for $1 \leq j \leq k$ iff the product $q = p_{i_1}p_{i_2}\cdots p_{i_k}$ divides x iff x is a multiple of q. Now, the number of multiples of q between 1 and m is $m/q = m/\prod_{i\in I} p_i$. If $I = \emptyset$ and the empty product is interpreted as 1, this expression becomes $m = |X|$. Hence, the inclusion-exclusion formula can be written

$$\phi(m) = |X \sim (S_1 \cup \cdots \cup S_n)| = m \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|} \frac{1}{\prod_{i\in I} p_i}.$$

On the other hand, consider what happens when we expand the product

$$m \prod_{i=1}^{n} \left(1 - \frac{1}{p_i}\right)$$

using the generalized distributive law (cf. 2.7). We will obtain a sum of 2^n terms, each of which is obtained by choosing either 1 or $-\frac{1}{p_i}$ from the ith factor of the product and multiplying these choices together. We can index these 2^n terms by subsets $I \subseteq \{1,2,\ldots,n\}$, where $i \in I$ iff we chose $-\frac{1}{p_i}$ from the ith factor. It follows that

$$m \prod_{i=1}^{n} \left(1 - \frac{1}{p_i}\right) = m \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|} \frac{1}{\prod_{i\in I} p_i} = \phi(m). \qquad \square$$

4.21. Remark. We sketch an alternative proof of the formula for $\phi(m)$ that avoids inclusion-exclusion. This proof sketch will use some facts from algebra and number theory without proof. For any commutative ring R, we let $R^\times$ be the set of *units* in R; i.e., the set of $x \in R$ such that there exists $y \in R$ with $xy = yx = 1_R$. The following facts are routinely verified. First, if R and S are isomorphic rings, then $|R^\times| = |S^\times|$. Second, given a product ring $R \times S$, we have $(R \times S)^\times = R^\times \times S^\times$ and hence (by the product rule) $|(R\times S)^\times| = |R^\times| \cdot |S^\times|$. Third, $\gcd(x,n) = 1$ iff there exist integers y, z with $xy + nz = 1$ iff x has a multiplicative inverse in the ring of integers modulo n. So $\phi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times|$. Fourth, by the Chinese Remainder Theorem, the rings $\mathbb{Z}/mn\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ are isomorphic whenever $\gcd(m,n) = 1$. Combining these four facts, we see that $\gcd(m,n) = 1$ implies

$$\phi(mn) = |(\mathbb{Z}/mn\mathbb{Z})^\times| = |(\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z})^\times| = |(\mathbb{Z}/m\mathbb{Z})^\times| \cdot |(\mathbb{Z}/n\mathbb{Z})^\times| = \phi(m)\phi(n).$$

Iteration of this result gives

$$\phi(p_1^{e_1} \cdots p_n^{e_n}) = \prod_{i=1}^{n} \phi(p_i^{e_i})$$

whenever $p_1, \ldots, p_n$ are distinct primes. Thus, it suffices to evaluate ϕ at prime powers. But a direct counting argument using the difference rule and the definition of ϕ shows that $\phi(p^e) = p^e - p^{e-1} = p^{e-1}(p-1)$ when p is prime and $e \geq 1$. So we obtain the first formula for $\phi(n)$ given in 4.20.

4.5 Derangements

The inclusion-exclusion formula allows us to enumerate a special class of permutations called derangements. Intuitively, a derangement of $1, 2, \ldots, n$ is a rearrangement of these n symbols such that no symbol remains in its original position. The formal definition is as follows.

4.22. Definition: Derangements. A *derangement* of a set S is a bijection $f : S \to S$ such that $f(x) \neq x$ for all $x \in S$. For $n \geq 0$, let D_n be the set of derangements of $\{1, 2, \ldots, n\}$, and let $d_n = |D_n|$.

Note that $d_0 = 1$ (since the function with empty graph satisfies the definition of derangement), while $d_1 = 0$. To give more examples of derangements, let us identify an element $f \in D_n$ with the word $f(1)f(2)\cdots f(n)$. Then $d_2 = 1$ since 21 is the unique derangement of two letters. The derangements of three letters are 312 and 231, so that $d_3 = 2$. The permutation 5317426 is a derangement of seven letters.

4.23. Summation Formula for Derangements. For $n \geq 1$, the number of derangements of n letters is

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

Consequently, d_n is the closest integer to $n!/e$ for $n \geq 1$.

Proof. Let X be the set of all permutations of n letters; note that $|X| = n!$. For $1 \leq i \leq n$, let $S_i = \{f \in X : f(i) = i\}$. The set D_n consists of precisely those elements in X that belong to none of the S_i, so $D_n = X \sim (S_1 \cup \cdots \cup S_n)$. To apply the inclusion-exclusion formula, we must consider a typical intersection $S_{i_1} \cap S_{i_2} \cap \cdots \cap S_{i_k}$, where $i_1 < i_2 < \cdots < i_k$. A permutation $f \in X$ belongs to this intersection iff f fixes $i_1, \ldots, i_k$ and permutes the remaining $n - k$ letters among themselves. The number of such permutations is $(n-k)!$. This number depends only on k and not on the indices $i_1, \ldots, i_k$. Applying the simplified inclusion-exclusion formula 4.12, we obtain

$$d_n = n! + \sum_{k=1}^{n} (-1)^k \binom{n}{k} (n-k)! = n! + \sum_{k=1}^{n} (-1)^k \frac{n!}{k!} = n! \sum_{k=0}^{n} (-1)^k \frac{1}{k!}.$$

To relate this formula to the expression $n!/e$, recall from calculus that

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \qquad (x \in \mathbb{R}).$$

Setting $x = -1$, we see that

$$1/e = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}.$$

Multiplying by $n!$ and comparing to our formula for d_n, we see that

$$n!/e - d_n = n! \sum_{k=n+1}^{\infty} \frac{(-1)^k}{k!}.$$

It now suffices to show that the right side of this formula is less than 1/2 in absolute value. Factoring out $\frac{1}{(n+1)!}$ from each term in the series, we obtain

$$|n!/e - d_n| = \frac{1}{n+1}\left|1 - \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} - \frac{1}{(n+2)(n+3)(n+4)} + \cdots\right|.$$

The series within the absolute values on the right side is an alternating series that converges to a sum strictly less than 1. Since $n \geq 1$, it follows that

$$|n!/e - d_n| < \frac{1}{n+1} \cdot 1 \leq 1/2. \qquad \square$$

The following table lists the first few values of d_n.

n	0	1	2	3	4	5	6	7	8	9
d_n	1	0	1	2	9	44	265	1854	14,833	133,496

Like any permutation, a derangement has a functional digraph consisting of the disjoint union of one or more cycles. A permutation is a derangement iff there are no 1-cycles in its functional digraph. This observation leads to the following recursion for derangements.

4.24. Theorem: Recursion for Derangements. We have $d_0 = 1$, $d_1 = 0$, and

$$d_n = (n-1)d_{n-1} + (n-1)d_{n-2} \qquad (n \geq 2).$$

Proof. Fix $n \geq 2$. Write the set of derangements D_n as the disjoint union of sets A and B, where A consists of those derangements in which n is involved in a cycle of length 2, and B consists of the derangements where n is in a cycle of length greater than 2. To build an object in A, choose the partner of n in its 2-cycle ($n-1$ ways), and then choose a derangement of the remaining objects (D_{n-2} ways). To build an object in B, choose a derangement of the first $n-1$ objects (D_{n-1} ways), consider the functional digraph of this derangement, and splice n into a cycle just before any of the $n-1$ available elements (which is guaranteed to create a cycle of length 3 or more). The recursion now follows from the sum and product rules. □

4.25. Theorem: Second Recursion for Derangements. We have $d_0 = 1$ and

$$d_n = nd_{n-1} + (-1)^n \qquad (n \geq 1).$$

Proof. We argue by induction on n. If $n = 1$, then

$$d_n = d_1 = 0 = 1 \cdot 1 + (-1)^1 = nd_{n-1} + (-1)^n.$$

Now assume $n > 1$ and that $d_{n-1} = (n-1)d_{n-2} + (-1)^{n-1}$. We can use this assumption to eliminate $(n-1)d_{n-2}$ in the first recursion 4.24 for d_n (which is already known to hold for all n). We thereby obtain

$$d_n = (n-1)d_{n-1} + (n-1)d_{n-2} = (n-1)d_{n-1} + (d_{n-1} - (-1)^{n-1}) = nd_{n-1} + (-1)^n.$$

This completes the induction. □

4.6 Coefficients of Chromatic Polynomials

Let G be a simple graph. Recall that $\chi_G(x)$ denotes the number of proper colorings of the vertices of G using x available colors. We have seen in 3.100 that $\chi_G(x)$ is always a *polynomial* in x. In this section, we use inclusion-exclusion to analyze the chromatic polynomial of G. This analysis will lead to a combinatorial interpretation for the coefficients of the chromatic polynomial $\chi_G(x)$.

4.26. Definition: Vertex-spanning Subgraph. Let $G = (V(G), E(G))$ be a simple graph. A *vertex-spanning subgraph* of G is a subgraph H of G such that $V(H) = V(G)$.

The map $H \mapsto E(H)$ is a bijection between the set of vertex-spanning subgraphs of G and the set of all subsets of $E(G)$.

4.27. Theorem: Coefficients of Chromatic Polynomials. Let G be a simple graph. For each $e, c \geq 0$, let $n(e, c)$ be the number of vertex-spanning subgraphs of G with e edges and c connected components. Then

$$\chi_G(x) = \sum_{e,c\geq 0} (-1)^e n(e, c) x^c.$$

Proof. Let $e_1, \ldots, e_n$ be the edges of G. Let X be the set of all colorings of G (proper or not) using x available colors, and let S_i be the set of colorings in X such that both endpoints of the edge e_i receive the same color. We wish to compute $|X \sim (S_1 \cup \cdots \cup S_n)|$. Consider a typical intersection $\bigcap_{i\in T} S_i$, where $T \subseteq \{1, 2, \ldots, n\}$. The edge subset $\{e_i : i \in T\}$ determines a vertex-spanning subgraph H of G with $|T|$ edges and some number $cc(H)$ of connected components. One may check that a coloring f belongs to $\bigcap_{i\in T} S_i$ iff f is constant on each connected component of H. It follows from the product rule that $|\bigcap_{i\in T} S_i| = x^{cc(H)}$, since we can choose one of x colors for each connected component of H. Note also that $|X| = x^{|V(G)|} = x^{cc(H_0)}$ where $H_0 = (V(G), \emptyset)$. By inclusion-exclusion,

$$\begin{aligned}
\chi_G(x) &= |X| + \sum_{\emptyset\neq T\subseteq\{1,2,\ldots,n\}} (-1)^{|T|} \left|\bigcap_{i\in T} S_i\right| \\
&= \sum_{\text{vertex-spanning subgraphs } H} (-1)^{|E(H)|} x^{cc(H)} = \sum_{e,c\geq 0} (-1)^e n(e, c) x^c. \quad \square
\end{aligned}$$

4.7 Classical Möbius Inversion

We conclude this chapter with a brief introduction to the theory of Möbius inversion. We begin in this section by studying the number-theoretic Möbius function and the classical Möbius inversion formula. Later sections discuss the generalization of the Möbius function and inversion formula to posets.

4.28. Definition: Classical Möbius Function. Suppose $m \geq 1$ is an integer with prime factorization $m = p_1^{e_1} p_2^{e_2} \cdots p_n^{e_n}$, where $n \geq 0$, $e_i > 0$, and the p_i's are distinct primes. (We take $n = 0$ when $m = 1$.) The *Möbius function* $\mu : \mathbb{N}^+ \to \{-1, 0, 1\}$ is defined by $\mu(m) = 0$ if $e_i > 1$ for some i, whereas $\mu(m) = (-1)^n$ if $e_i = 1$ for all i.

In other words, $\mu(m)$ is zero if m is divisible by the square of a prime; $\mu(m) = +1$ if m is the product of an even number of distinct primes; and $\mu(m) = -1$ if m is the product of an odd number of distinct primes. For example,

$$\mu(1) = 1,\ \mu(7) = -1,\ \mu(10) = 1,\ \mu(12) = 0,\ \mu(30) = -1.$$

The following theorem is the key to proving the Möbius inversion formula.

4.29. Theorem. For all integers $m \geq 1$, $\sum_{d|m} \mu(d) = \chi(m = 1)$. (Here and below, the symbol $\sum_{d|m}$ means that we sum over all *positive* divisors d of the integer m.)

Proof. When $m = 1$, we have $\sum_{d|1} \mu(d) = \mu(1) = 1 = \chi(m = 1)$. Suppose next that $m > 1$ and m has prime factorization $p_1^{e_1} \cdots p_n^{e_n}$. Instead of summing $\mu(d)$ over *all* divisors d of m, we may equally well sum over just the *square-free* divisors d of m, which give the only nonzero contributions to the sum. Examining prime factorizations, we see that there are 2^n such square-free divisors, which have the form $\prod_{i\in T} p_i$ as T ranges over all subsets of $\{1, 2, \ldots, n\}$. Therefore,

$$\sum_{d|m} \mu(d) = \sum_{T\subseteq\{1,2,\ldots,n\}} \mu\left(\prod_{i\in T} p_i\right) = \sum_{T\subseteq\{1,2,\ldots,n\}} (-1)^{|T|}.$$

Collecting together summands indexed by subsets T of the same size k, we conclude that

$$\sum_{d|m} \mu(d) = \sum_{k=0}^{n} \sum_{\substack{T\subseteq\{1,\ldots,n\}\\ |T|=k}} (-1)^{|T|} = \sum_{k=0}^{n} \binom{n}{k} (-1)^k 1^{n-k} = (-1+1)^n = 0. \qquad \square$$

4.30. Classical Möbius Inversion Formula. Suppose f and g are functions with domain $\mathbb{N}^+$ such that

$$f(m) = \sum_{d|m} g(d) \qquad (m \geq 1).$$

Then

$$g(m) = \sum_{d|m} f(m/d)\mu(d) = \sum_{d|m} f(d)\mu(m/d) \qquad (m \geq 1).$$

Proof. We use the definition of f to expand the first claimed formula for $g(m)$:

$$\sum_{d|m} f(m/d)\mu(d) = \sum_{d|m} \left(\sum_{c|(m/d)} g(c) \right) \mu(d) = \sum_{(c,d)\in S} g(c)\mu(d),$$

where $S = \{(c, d) \in \mathbb{N}^+ \times \mathbb{N}^+ : d|m, c|(m/d)\}$. It follows routinely from the definition of divisibility that

$$S = \{(c, d) : d|m, cd|m\} = \{(c, d) : c|m, cd|m\} = \{(c, d) : c|m, d|(m/c)\}.$$

Therefore, the calculation continues as follows:

$$\begin{aligned}\sum_{(c,d)\in S} g(c)\mu(d) &= \sum_{c|m} \sum_{d|(m/c)} g(c)\mu(d) = \sum_{c|m} g(c) \left(\sum_{d|(m/c)} \mu(d) \right) \\ &= \sum_{c|m} g(c)\chi(m/c = 1) = g(m).\end{aligned}$$

The next-to-last step used 4.29 to simplify the inner sum. We conclude that

$$g(m) = \sum_{d|m} f(m/d)\mu(d) = \sum_{d|m} f(d)\mu(m/d),$$

where the final equality results by replacing d by m/d in the summation. □

To give examples of the Möbius inversion formula, we first introduce some functions that are studied in number theory.

4.31. Definition: Number-Theoretic Functions τ, σ, and σ_2. Let $m \geq 1$ be an integer. Define

$$\tau(m) = \sum_{d|m} 1; \qquad \sigma(m) = \sum_{d|m} d; \qquad \sigma_2(m) = \sum_{d|m} d^2.$$

Thus, $\tau(m)$ is the number of positive divisors of m; $\sigma(m)$ is the sum of these divisors; and $\sigma_2(m)$ is the sum of the squares of these divisors.

4.32. Example. Taking $m = 1, 4, 7, 12, 30$, we calculate:

$$\begin{array}{lllll}
\tau(1) = 1, & \tau(4) = 3, & \tau(7) = 2, & \tau(12) = 6, & \tau(30) = 8; \\
\sigma(1) = 1, & \sigma(4) = 7, & \sigma(7) = 8, & \sigma(12) = 28, & \sigma(30) = 72; \\
\sigma_2(1) = 1, & \sigma_2(4) = 21, & \sigma_2(7) = 50, & \sigma_2(12) = 210, & \sigma_2(30) = 1300.
\end{array}$$

If m has prime factorization $p_1^{e_1} \cdots p_n^{e_n}$, then the divisors of m have the form $p_1^{f_1} \cdots p_n^{f_n}$ where $0 \leq f_i \leq e_i$ for all i. The product rule therefore gives $\tau(m) = \prod_{i=1}^n (e_i + 1)$ (build a divisor by choosing $f_1, \ldots, f_n$). Using the generalized distributive law and the geometric series formula, one may also check that

$$\sigma(m) = \prod_{i=1}^n \left(\sum_{f_i=0}^{e_i} p_i^{f_i} \right) = \prod_{i=1}^n \frac{p_i^{e_i+1} - 1}{p_i - 1}.$$

Applying the Möbius inversion formula to the definitions of τ, σ, and σ_2, we obtain the following identities.

4.33. Theorem. For $m \geq 1$, we have

$$1 = \sum_{d|m} \tau(m/d)\mu(d); \qquad m = \sum_{d|m} \sigma(m/d)\mu(d); \qquad m^2 = \sum_{d|m} \sigma_2(m/d)\mu(d).$$

The next result uses Möbius inversion to deduce information about Euler's ϕ function.

4.34. Theorem: ϕ versus μ. For all $m \geq 1$,

$$m = \sum_{d|m} \phi(d) \qquad \text{and so} \qquad \phi(m) = \sum_{d|m} \mu(d)(m/d).$$

Proof. To prove the first formula, fix $m \geq 1$. For each divisor d of m, let

$$S_d = \{x \in \mathbb{N}^+ : 1 \leq x \leq m \text{ and } \gcd(x, m) = d\}.$$

It is immediate that the m-element set $\{1, 2, \ldots, m\}$ is the disjoint union of the sets S_d as d ranges over the positive divisors of m. Whenever d divides m, we have $\gcd(x, m) = d$ iff d divides x and $\gcd(x/d, m/d) = 1$. It follows that division by d gives a bijection from the set S_d onto the set of numbers counted by $\phi(m/d)$. Therefore, $|S_d| = \phi(m/d)$. By the sum rule,

$$m = \sum_{d|m} |S_d| = \sum_{d|m} \phi(m/d) = \sum_{d|m} \phi(d).$$

The last equality follows by noting that the number m/d ranges over all positive divisors of m as d ranges over all positive divisors of m. Applying Möbius inversion (with $f(m) = m$ and $g(m) = \phi(m)$), we obtain the second formula in the theorem. □

Some applications of these results to field theory are presented in §12.6.

4.8 Partially Ordered Sets

We will see that the inclusion-exclusion formula 4.7 and the classical Möbius inversion formula 4.30 are special cases of the general Möbius inversion formula for partially ordered sets (posets). First we must review some definitions and examples concerning posets.

Recall from 2.54 the definition of relations and the notions of reflexive, irreflexive, symmetric, antisymmetric, and transitive relations. Given a relation R on a finite set X, the pair (X, R) is a digraph G with vertex set X and directed edge set R. Reflexivity means that *every* vertex of G has a loop edge; irreflexivity means that *no* vertex of G has a loop edge. Symmetry means that the reversal of every edge is also an edge (so we can think of G as undirected); antisymmetry means that it is never true that a non-loop edge and its reversal are both in G. Finally, transitivity means that whenever there is a walk (x, y, z) of length 2 in G, the edge (x, z) is also present in G. More generally, we see by induction that when R is transitive, there exists a walk from x to z in G of positive length iff the edge (x, z) is present in G.

4.35. Poset Definitions. A *partial order relation* on X is a relation that is antisymmetric, transitive, and reflexive on X. A *strict order relation* on X is a relation that is transitive and irreflexive on X. A *partially ordered set (poset)* is a pair $(X, \leq)$ where $\leq$ is a partial order relation on X. A *totally ordered set* is a poset $(X, \leq)$ such that for all $x, y \in X$, either $x \leq y$ or $y \leq x$.

4.36. Example. Let $X = \{1, 2, \ldots, n\}$ and take $\leq$ to be the usual ordering of integers. Then $(X, \leq)$ is an n-element totally ordered poset. More generally, for any $S \subseteq \mathbb{R}$, $(S, \leq)$ is a totally ordered poset.

4.37. Example: Boolean Posets. Let S be any set, and let $X = \mathcal{P}(S)$ be the set of all subsets of S. Then $(X, \subseteq)$ is a poset, where $A \subseteq B$ means that A is a subset of B. In particular, $(\mathcal{P}(\{1, 2, \ldots, n\}), \subseteq)$ is a poset of size 2^n. This poset is not totally ordered when $n > 1$.

4.38. Example: Divisibility Posets. Consider the divisibility relation $|$ on $\mathbb{N}^+$ defined by $a|b$ iff $b = ac$ for some $c \in \mathbb{N}^+$. Then $(\mathbb{N}^+, |)$ is an infinite poset. Given a fixed positive integer n, let X be the set of all divisors of n. Restricting $|$ to X gives a finite poset $(X, |)$. This poset is a totally ordered set iff n is a prime power.

The next result shows that partial order relations and strict order relations are essentially equivalent concepts.

4.39. Theorem: Partial Orders vs. Strict Orders. Let X be a set, let P be the set of all partial order relations on X, and let S be the set of all strict order relations on X. There is a canonical bijection between P and S.

Proof. Let $\Delta = \{(x,x) : x \in X\}$ be the "diagonal" of $X \times X$. Define $f : P \to S$ by setting $f(\leq) = \leq \sim \Delta$ for each partial ordering $\leq$ on X. Define $g : S \to P$ by setting $g(<) = < \cup \Delta$ for each strict ordering $<$ on X. In terms of the digraphs, f removes self-loops from all vertices and g restores the self-loops. It is an exercise for the reader to show that f does map P into S, g does map S into P, and $f \circ g$ and $g \circ f$ are both identity maps. □

4.9 Möbius Inversion for Posets

4.40. Definition: Matrix of a Relation. Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite set, and let R be a relation on X. Define the *matrix of* R to be the $n \times n$ matrix $A = A(R)$ such that $A_{i,j} = \chi(x_i R x_j)$. $A(R)$ is the adjacency matrix of the digraph (X, R).

4.41. Theorem. Let $\leq$ be a partial ordering of $X = \{x_1, \ldots, x_n\}$, and let $<$ be the associated strict ordering of X (see 4.39). Consider the matrices $Z = A(\leq)$ and $N = A(<)$. Then $Z = I + N$; N is nilpotent; Z is invertible; and

$$Z^{-1} = I - N + N^2 - N^3 + \cdots + (-1)^{n-1} N^{n-1}. \tag{4.4}$$

Proof. The matrix identity $Z = I + N$ holds since $(X, \leq)$ is obtained from $(X, <)$ by adding self-loops at each $x \in X$. Next, we claim that the digraph $(X, <)$ is acyclic. For if $(z_1, z_2, \ldots, z_k, z_1)$ were a directed cycle in this digraph, we must have $z_1 < z_2 < \cdots < z_k < z_1$. Then transitivity gives $z_1 < z_1$, which contradicts irreflexivity. By 3.24, N is nilpotent. The statements about the inverse of Z now follow from 3.25, taking A there to be $-N$. □

4.42. Definition: Möbius Function of a Finite Poset. Keeping the notation of the preceding theorem, define $\mu = \mu_{(X,\leq)} : X \times X \to \mathbb{Z}$ by setting $\mu(x_i, x_j)$ to be the i,j-entry of Z^{-1}. The function μ is called the *Möbius function of the poset* $(X, \leq)$.

4.43. Example. Let $X = \{1, 2, 3, 4\}$ with the usual ordering. For this poset, we have

$$Z = A(\leq) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad N = A(<) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The powers of N are

$$N^2 = \begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad N^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad N^4 = 0.$$

So the inverse of Z is

$$Z^{-1} = I - N + N^2 - N_3 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

So $\mu(i,i) = 1$, $\mu(i, i+1) = -1$, and $\mu(i,j) = 0$ for all $j \neq i, i+1$.

4.44. Example: Möbius Function of a Totally Ordered Poset. The preceding example generalizes as follows. Let $X = \{1, 2, \ldots, n\}$ with the usual ordering. We have $Z_{i,j} = 1$ for $i \leq j$ and $Z_{i,j} = 0$ for $i > j$. For all i, let $M_{i,i} = 1$, $M_{i,i+1} = -1$, and $M_{i,j} = 0$ for $j \neq i, i+1$. A routine matrix calculation shows that $ZM = MZ = I$. So for this poset,

$$\mu(i,i) = 1, \qquad \mu(i,i+1) = -1, \qquad \mu(i,j) = 0 \text{ for } j \neq i, i+1.$$

4.45. Example: Möbius Function for Boolean Posets. Consider the poset $(X, \subseteq)$, where X consists of all subsets of $[n] = \{1, 2, \ldots, n\}$. In this example, we will index the rows and columns of matrices by subsets of $[n]$. For $S, T \subseteq [n]$, the S,T-entry of Z is 1 if $S \subseteq T$, and 0 otherwise. We claim that the inverse matrix $M = Z^{-1}$ has S,T-entry $\mu(S,T) = (-1)^{|T \sim S|}$ if $S \subseteq T$, and zero otherwise. To verify this, let us show that $ZM = I$. The S,T-entry of ZM is

$$\sum_{U \subseteq [n]} Z(S,U)M(U,T) = \sum_{U : S \subseteq U \subseteq T} (-1)^{|T \sim U|}.$$

If $S = T$, this sum is 1; while if $S \not\subseteq T$, this sum is 0. Now consider the case where $S \subsetneq T$. Let S have a elements and T have $a + b$ elements, where $b > 0$. For $0 \leq c \leq b$, the number of sets U with $S \subseteq U \subseteq T$ and $|T \sim U| = c$ is $\binom{b}{b-c} = \binom{b}{c}$. Grouping terms in the sum based on the size of $|T \sim U|$, we see that

$$(ZM)(S,T) = \sum_{c=0}^{b} (-1)^c \binom{b}{c} = (-1+1)^b = 0.$$

So the Möbius function for this poset is

$$\mu(S,T) = (-1)^{|T \sim S|} \chi(S \subseteq T) \qquad (S, T \subseteq [n]).$$

An alternate proof of this formula will be given in 4.58 below.

4.46. Example: Möbius Function for Divisibility Posets. Let n be a fixed positive integer, let X be the set of positive divisors of n, and consider the divisibility poset $(X, |)$. There is a close relation between the classical Möbius function μ and the Möbius function μ_X for this poset. More precisely, we claim that

$$\mu(d) = \mu_X(1, d) \qquad \text{for all } d \text{ dividing } n.$$

To verify this, let us work with matrices whose rows and columns are indexed by the positive divisors of n, considered in increasing order. As above, let Z be the matrix such that $Z_{d,e} = \chi(d|e)$; let M be the inverse matrix, which is uniquely determined by Z; and let $\vec{v}$ be the row vector $(\mu(d) : d|n)$. The identity $\sum_{d|m} \mu(d) = \chi(m = 1)$, which is valid for all m dividing n, can now be rewritten as the vector identity $\vec{v}Z = (1, 0, \ldots, 0)$. This shows that $\vec{v}$ must be the first row of M. It will be shown in 4.59 that $\mu_X(d,e) = \mu(e/d)$ whenever $d|e$ and $e|n$, whereas $\mu_X(d,e) = 0$ if d does not divide e.

The next definition will be used to give a combinatorial interpretation for the values of the Möbius function.

4.47. Definition: Chains in a Poset. Let $(X, \leq)$ be a poset. A *chain of length* k in X is a sequence $C = (z_0, z_1, \ldots, z_k)$ of elements of X such that

$$z_0 < z_1 < \cdots < z_k.$$

We say that C is a chain *from* z_0 *to* z_k and write $\text{len}(C) = k$. The *sign* of the chain C is $\text{sgn}(C) = (-1)^k$.

4.48. Theorem: Möbius Functions and Signed Chains. Let $(X,\leq)$ be a finite poset. Given $y,z\in X$, let S be the set of all chains in X from y to z. Then

$$\mu_{(X,\leq)}(y,z)=\sum_{C\in S}\operatorname{sgn}(C).$$

In particular, if $y\not\leq z$, then $\mu_{(X,\leq)}(y,z)=0$.

Proof. We know from (4.4) that

$$\mu_{(X,\leq)}(y,z)=\sum_{k\geq 0}(-1)^k N^k(y,z),$$

where N is the adjacency matrix of the digraph $G=(X,<)$. A chain of length k from y to z is the same as a walk (or path) of length k from y to z in G. By 3.18, the number of such walks is $N^k(y,z)$. The theorem now follows from the sum rule. □

4.49. Theorem: Möbius Inversion Formula on Posets. Let $(X,\leq)$ be a finite poset with Möbius function μ. Suppose R is a commutative ring and $f,g:X\to R$ are two functions. Then

$$\left(\forall x\in X, g(x)=\sum_{y\leq x}f(y)\right)\qquad\text{iff}\qquad\left(\forall x\in X, f(x)=\sum_{y\leq x}g(y)\mu(y,x)\right).$$

Proof. Let $X=\{x_1,\ldots,x_n\}$, and define $Z=A(\leq)$ and $M=Z^{-1}$ as in 4.41. Also define row vectors $F=[f(x_1),\ldots,f(x_n)]$ and $G=[g(x_1),\ldots,g(x_n)]$. The left-hand formula in the theorem is equivalent to the matrix identity $G=FZ$, since $G_j=g(x_j)$ and

$$(FZ)_j=\sum_{k=1}^n F_k Z_{k,j}=\sum_{k=1}^n f(x_k)\chi(x_k\leq x_j)=\sum_{y\leq x_j}f(y).$$

Similarly, keeping in mind that $\mu(y,x)=0$ unless $y\leq x$, the right-hand formula in the theorem is equivalent to the matrix identity $F=GM$. Since M and Z are inverse matrices, $G=FZ$ is equivalent to $GM=F$. □

4.50. Example. In the special case where $X=\{1,2,\ldots,n\}$ with the usual ordering, 4.49 reduces to the following statement: given $f_1,\ldots,f_n\in R$ and $g_1,\ldots,g_n\in R$, we have ($g_i=f_1+f_2+\cdots+f_i$ for all i) iff ($f_1=g_1$ and $f_i=g_i-g_{i-1}$ for $1<i\leq n$).

4.51. Example. In the special case where X is the set of divisors of n ordered by divisibility, 4.49 reduces to the classical inversion formula 4.30, using the fact that $\mu_X(d,e)=\mu(e/d)$ when $d|e$, and $\mu_X(d,e)=0$ otherwise.

4.52. Example. In the special case where $X=\mathcal{P}([n])$ ordered by containment of subsets, 4.49 reduces to the following statement:

$$\left(\forall T\subseteq[n], g(T)=\sum_{S\subseteq T}f(S)\right)\qquad\text{iff}\qquad\left(\forall T\subseteq[n], f(T)=\sum_{S\subseteq T}(-1)^{|T\sim S|}g(S)\right).$$

If instead we use the "opposite" poset $(X,\supseteq)$, one obtains:

$$\left(\forall T\subseteq[n], g(T)=\sum_{S\supseteq T}f(S)\right)\qquad\text{iff}\qquad\left(\forall T\subseteq[n], f(T)=\sum_{S\supseteq T}(-1)^{|S\sim T|}g(S)\right).$$

We now use this result to rederive a version of the original inclusion-exclusion formula. Let $Z_1, \ldots, Z_n$ be given subsets of a finite set Z. For $S \subseteq [n]$, let $f(S)$ be the number of objects $z \in Z$ such that $z \in Z_i$ if and only if $i \in S$. For $S \subseteq [n]$, let $g(S)$ be the number of objects $z \in Z$ such that $z \in Z_i$ if $i \in S$. Regarding Z_i as the set of objects in Z with a certain property i, we can say that $f(S)$ counts objects that have *exactly* the properties in S, whereas $g(S)$ counts the objects that have *at least* the properties in S. It follows from this that $g(T) = \sum_{S \supseteq T} f(S)$ for all T, so 4.49 tells us that $f(T) = \sum_{S \supseteq T} (-1)^{|S \sim T|} g(S)$ for all T. Now, $f(\emptyset) = |Z \sim (Z_1 \cup \cdots \cup Z_n)|$ and $g(\{i_1, \ldots, i_k\}) = |Z_{i_1} \cap \cdots \cap Z_{i_k}|$. The formula in 4.11 follows from these observations.

4.10 Product Posets

This section introduces a construction on posets that leads to alternative derivations of the Möbius functions for the posets $(\mathcal{P}([n]), \subseteq)$ and $(\{d : d|n\}, |)$.

4.53. Definition: Product Posets. Let $(X_1, \leq_1), \ldots, (X_n, \leq_n)$ be posets. Consider the Cartesian product $X = X_1 \times \cdots \times X_n$, which consists of all n-tuples $x = (x_1, \ldots, x_n)$ with $x_i \in X_i$. For $x = (x_i)$ and $y = (y_i)$ in X, define $x \leq y$ iff $x_i \leq_i y_i$ for $1 \leq i \leq n$. One immediately verifies that $\leq$ is a partial ordering on X. The poset $(X, \leq)$ is called the *product* of the posets $(X_i, \leq_i)$.

4.54. Example. Let $X_1 = X_2 = \{1, 2\}$ with the usual ordering. Both X_1 and X_2 are totally ordered posets, but $X = X_1 \times X_2$ is not totally ordered. For example, $(1, 2)$ and $(2, 1)$ are two incomparable elements of X.

4.55. Theorem: Möbius Function for a Product Poset. Let $(X, \leq)$ be the product of posets $(X_i, \leq_i)$ for $1 \leq i \leq k$. Given $x = (x_i)$ and $y = (y_i)$ in X, we have

$$\mu_{(X,\leq)}(x, y) = \prod_{i=1}^{k} \mu_{(X_i,\leq_i)}(x_i, y_i).$$

Proof. For brevity, write $\mu = \mu_{(X,\leq)}$ and $\mu_i = \mu_{(X_i,\leq_i)}$. By induction, we can reduce to the case $k = 2$. We have the matrices

$$\begin{array}{ll} Z_1 = [\chi(u_1 \leq_1 v_1) : u_1, v_1 \in X_1], & M_1 = [\mu_1(u_1, v_1) : u_1, v_1 \in X_1], \\ Z_2 = [\chi(u_2 \leq_2 v_2) : u_2, v_2 \in X_2], & M_2 = [\mu_2(u_2, v_2) : u_2, v_2 \in X_2], \\ Z = [\chi(u \leq v) : u, v \in X], & M = [\mu(u, v) : u, v \in X], \end{array}$$

which satisfy $Z_1 M_1 = I$, $Z_2 M_2 = I$ and $ZM = I$. Define a matrix M', with rows and columns indexed by elements of X, such that for $u = (u_1, u_2)$ and $v = (v_1, v_2)$ in X, the u, v-entry of M' is $\mu_1(u_1, v_1)\mu_2(u_2, v_2)$. Note that the u, v-entry of Z is $\chi((u_1, u_2) \leq (v_1, v_2)) = \chi(u_1 \leq_1 v_1)\chi(u_2 \leq_2 v_2)$. The following computation verifies that $ZM' = I$, and hence $M' = M$:

$$\begin{aligned} (ZM')(u, w) &= \sum_{v \in X} Z(u, v) M'(v, w) \\ &= \sum_{v_1 \in X_1} \sum_{v_2 \in X_2} \chi(u_1 \leq_1 v_1)\chi(u_2 \leq_2 v_2)\mu_1(v_1, w_1)\mu_2(v_2, w_2) \\ &= \left(\sum_{v_1 \in X_1} \chi(u_1 \leq_1 v_1)\mu_1(v_1, w_1) \right) \cdot \left(\sum_{v_2 \in X_2} \chi(u_2 \leq_2 v_2)\mu_2(v_2, w_2) \right) \\ &= \chi(u_1 = w_1)\chi(u_2 = w_2) = \chi(u = w). \quad \square \end{aligned}$$

4.56. Definition: Poset Isomorphisms. Given posets $(X, \leq)$ and $(X', \leq')$, a *poset isomorphism* is a bijection $f : X \to X'$ such that

$$u \leq v \text{ iff } f(u) \leq' f(v) \qquad (u, v \in X).$$

4.57. Theorem. If $f : (X, \leq) \to (X', \leq')$ is a poset isomorphism, then

$$\mu_{(X',\leq')}(f(u), f(v)) = \mu_{(X,\leq)}(u, v) \qquad (u, v \in X).$$

Proof. This follows, for instance, from 4.48. For, the chains of a given length from u to v in $(X, \leq)$ correspond bijectively to the chains of that length from $f(u)$ to $f(v)$ in $(X', \leq')$; we merely apply f to each element in the chain. □

4.58. Example: Möbius Function of a Boolean Poset. Consider again the poset $X = (\mathcal{P}([n]), \subseteq)$. For $1 \leq i \leq n$, take $Y_i = \{0, 1\}$ with the usual ordering, and let $Y = Y_1 \times \cdots \times Y_n$ be the product poset. There is a bijection f from $\mathcal{P}([n])$ to $\{0, 1\}^n$ that sends a subset S to the word $f(S) = w = w_1 w_2 \cdots w_n$ with $w_i = 1$ for $i \in S$ and $w_i = 0$ for $i \notin S$. One readily sees that f is a poset isomorphism, so $\mu_X(S, T) = \mu_Y(f(S), f(T))$. Writing $f(T) = z = z_1 z_2 \cdots z_n$, 4.55 shows that $\mu_Y(w, z) = \prod_{i=1}^{n} \mu_{Y_i}(w_i, z_i)$. As in 4.44, we see that

$$\mu_{Y_i}(0, 0) = \mu_{Y_i}(1, 1) = 1; \quad \mu_{Y_i}(0, 1) = -1; \quad \mu_{Y_i}(1, 0) = 0.$$

So $\mu_Y(w, z) = 0$ unless $w \leq z$. If $w \leq z$ and z has k more ones than w does, we see that $\mu_Y(w, z) = (-1)^k$. Translating back to subsets via f^{-1}, this says that $\mu_X(S, T) = 0$ when $S \not\subseteq T$, and $\mu_X(S, T) = (-1)^{|T \sim S|}$ when $S \subseteq T$.

4.59. Example: Möbius Function of a Divisibility Poset. Let n be a fixed positive integer with prime factorization $n = p_1^{n_1} \cdots p_k^{n_k}$, and consider the divisibility poset $(X, |)$, where $X = \{d : d|n\}$. For $1 \leq i \leq k$, let $Y_i = \{0, 1, \ldots, n_i\}$ with the usual ordering, and take Y to be the product poset $Y_1 \times \cdots \times Y_k$. Any $d \in X$ has prime factorization $d = p_1^{d_1} \cdots p_k^{d_k}$ for some $d_k \leq n_k$. The map $d \mapsto (d_1, \ldots, d_k)$ is readily seen to be a poset isomorphism from X to Y. So

$$\mu_X(d, e) = \mu_Y((d_1, \ldots, d_k), (e_1, \ldots, e_k)) = \prod_{i=1}^{k} \mu_{Y_i}(d_i, e_i).$$

As in 4.44, we see that $\mu_{Y_i}(d_i, e_i) = \chi(e_i = d_i) - \chi(e_i = d_i + 1)$. It follows that $\mu_X(d, e) = 0$ unless e is obtained from d by multiplying by a set of s *distinct* prime factors chosen from $\{p_1, \ldots, p_k\}$, in which case $\mu_X(d, e) = (-1)^s$. It is now routine to check that whenever $d|e$, $\mu_X(d, e) = \mu(e/d)$, where μ is the number-theoretic Möbius function.

Summary

- *Involutions.* An involution is a function $I : X \to X$ with $I \circ I = \mathrm{id}_X$. The fixed point set of I is $\mathrm{Fix}(I) = \{x \in X : I(x) = x\}$. If X consists of signed objects, I is sign-reversing iff $\mathrm{sgn}(I(x)) = -\mathrm{sgn}(x)$ for all $x \in X \sim \mathrm{Fix}(I)$. For a sign-reversing involution I with domain X,

$$\sum_{x \in X} \mathrm{sgn}(x) = \sum_{x \in \mathrm{Fix}(I)} \mathrm{sgn}(x).$$

  Involutions provide combinatorial proofs of identities that involve signed terms.

- *Inclusion-Exclusion Formulas.* For arbitrary finite sets $S_1, \ldots, S_n$,

$$|S_1 \cup S_2 \cup \cdots \cup S_n| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} |S_{i_1} \cap S_{i_2} \cap \cdots \cap S_{i_k}|.$$

  If each S_i is a subset of a finite set X, then

$$|X \sim (S_1 \cup \cdots \cup S_n)| = \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|} \left| \bigcap_{i \in I} S_i \right|,$$

  where the summand indexed by $I = \emptyset$ is interpreted as $|X|$. In the special case where $\left|\bigcap_{i \in I} S_i\right| = N(k)$ for all k-element subsets I, the formula simplifies to

$$|X \sim (S_1 \cup \cdots \cup S_n)| = |X| + \sum_{k=1}^{n} (-1)^k \binom{n}{k} N(k).$$

- *Surjections and Stirling Numbers.* For $m \ge n \ge 1$, there are $\sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)^m$ surjections from an m-element set onto an n-element set. A summation formula for the Stirling number of the second kind is

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

- *Euler's ϕ Function.* For $m \ge 1$, $\phi(m)$ is the number of positive integers $x \le m$ with $\gcd(x,m) = 1$. We have $\phi(m) = m \prod_{p|m} (1 - p^{-1})$ where the product ranges over all prime divisors p of m. For $m = q^e$ with q prime, $\phi(q^e) = q^e - q^{e-1}$. If $\gcd(m,n) = 1$, then $\phi(mn) = \phi(m)\phi(n)$. For $m \ge 1$, $\sum_{d|m} \phi(d) = m$.

- *Derangements.* A *derangement of* S is a bijection $f : S \to S$ with $f(x) \ne x$ for all $x \in S$. Let d_n be the number of derangements of an n-element set. Then $d_n = n! \sum_{k=0}^{n} (-1)^k / k!$ is the closest integer to $n!/e$. Moreover, the numbers d_n satisfy the recursions

$$d_n = (n-1)d_{n-1} + (n-1)d_{n-2} \qquad (n \ge 2);$$

$$d_n = nd_{n-1} + (-1)^n \qquad (n \ge 1).$$

- *Coefficients of Chromatic Polynomials.* For a simple graph G, the coefficient of x^c in the chromatic polynomial $\chi_G(x)$ is $\sum_{e \ge 0} (-1)^e n(e,c)$, where $n(e,c)$ is the number of subgraphs H of G such that $V(H) = V(G)$, $|E(H)| = e$, and H has c connected components.

- *Number-theoretic Möbius Function.* Define $\mu : \mathbb{N}^+ \to \{-1, 0, 1\}$ by $\mu(n) = (-1)^s$ if n is the product of $s \ge 0$ distinct primes, and $\mu(n) = 0$ otherwise. Then $\sum_{d|m} \mu(d) = \chi(m = 1)$. Given functions f and g such that $f(m) = \sum_{d|m} g(d)$ for $m \ge 1$, the classical Möbius inversion formula states that

$$g(m) = \sum_{d|m} f(m/d)\mu(d) = \sum_{d|m} f(d)\mu(m/d) \qquad (m \ge 1).$$

  It follows that $\phi(m) = \sum_{d|m} \mu(d) m/d$.

- *Posets.* A *partial ordering* of X is a relation $\leq$ on X that is reflexive, antisymmetric, and transitive; the pair $(X, \leq)$ is called a *poset*. A *strict ordering* of X is a relation $<$ on X that is irreflexive and transitive. There is a bijection between partial orders on X and strict orders on X defined by removing the diagonal $\{(x,x) : x \in X\}$. A *chain of length k* in a poset $(X, \leq)$ is a sequence $(z_0, z_1, \ldots, z_k)$ with $z_0 < z_1 < \cdots < z_k$. Such a chain *goes from z_0 to z_k* and has *sign* $(-1)^k$.

- *Möbius Functions for Posets.* Given a poset $(X = \{x_1, \ldots, x_n\}, \leq)$, define $n \times n$ matrices Z, N, and M by $Z_{ij} = \chi(x_i \leq x_j)$, $N_{ij} = \chi(x_i < x_j)$, and $M_{ij} =$ the signed sum of all chains in the poset from x_i to x_j. Then $Z = I + N$; N is nilpotent; and M is the matrix inverse of Z. We write $\mu_X(x_i, x_j) = M_{ij}$ and call μ the *Möbius function* of the poset $(X, \leq)$. Suppose f and g are functions with domain X. The *Möbius inversion formula for posets* states that

$$g(x) = \sum_{y \leq x} f(y) \text{ for all } x \in X \text{ iff } f(x) = \sum_{y \leq x} g(y)\mu_X(y, x) \text{ for all } x \in X.$$

- *Product Posets.* Given posets $(X_i, \leq_i)$ for $1 \leq i \leq n$, the product set $X = X_1 \times \cdots \times X_n$ becomes a poset by defining $(x_1, \ldots, x_n) \leq (y_1, \ldots, y_n)$ iff $x_i \leq_i y_i$ for all i. The Möbius function for the product poset satisfies

$$\mu_X((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \prod_{i=1}^{n} \mu_{X_i}(x_i, y_i).$$

- *Examples of Möbius Functions.* The poset $X = \{1, 2, \ldots, n\}$ with the usual total ordering has Möbius function

$$\mu_X(i,i) = 1, \quad \mu_X(i, i+1) = -1, \quad \mu_X(i,j) = 0 \text{ for } j \neq i, i+1.$$

The Boolean poset $(\mathcal{P}(X), \subseteq)$ of subsets of $\{1, 2, \ldots, n\}$ ordered by inclusion has Möbius function

$$\mu(S,T) = (-1)^{|T \sim S|}\chi(S \subseteq T) \qquad (S, T \subseteq [n]).$$

If N has prime factorization $p_1^{n_1} \cdots p_k^{n_k}$, then the poset of divisors of N under the divisibility ordering has Möbius function

$$\mu(d,e) = \begin{cases} (-1)^s & \text{if } e/d \text{ is a product of } s \text{ distinct primes;} \\ 0 & \text{otherwise.} \end{cases}$$

These results follow since the Boolean poset is isomorphic to the product of n copies of the totally ordered set $\{0, 1\}$, whereas the divisiblity poset is isomorphic to the product poset $\{0, 1, \ldots, n_1\} \times \cdots \times \{0, 1, \ldots, n_k\}$.

Exercises

4.60. Given that $|S| = 15$, $|T| = 13$, $|U| = 12$, $|S \cap T| = 6$, $|S \cap U| = 3$, $|T \cap U| = 4$, and $|S \cap T \cap U| = 1$, find: (a) $|S \cup T|$; (b) $|S \cup T \cup U|$; (c) the number of objects in exactly one of the sets S, T, U.

4.61. Given that S, T, U are subsets of X with $|X| = 35$, $|S| = 12$, $|T| = 14$, $|U| = 15$, $|S \cap T| = 5 = |S \cap U|$, $|T \cap U| = 6$, and $|(S \cup T) \cap U| = 9$, find: (a) $|S \cap T \cap U|$; (b) $|X \sim (S \cup T \cup U)|$; (c) the number of objects in exactly two of the sets S, T, U.

4.62. List all the derangements in D_4.

4.63. Compute d_{10} in four ways: (a) by rounding $10!/e$ to the nearest integer; (b) by using the summation formula 4.23; (c) by using the recursion in 4.24; (d) by using the recursion in 4.25.

4.64. Compute $\phi(n)$, $\mu(n)$, $\tau(n)$, and $\sigma(n)$ for the following choices of n: (a) 6; (b) 11; (c) 28; (d) 60; (e) 1001; (f) 121.

4.65. Verify 4.34 by direct calculation for (a) $m = 24$; (b) $m = 30$.

4.66. Given n married couples, how many ways can the n men and n women be paired up so that no pair consists of a man and his wife?

4.67. How many five-card poker hands have at least one card of every suit?

4.68. How many five-card poker hands have at least one face card, at least one diamond, and do not contain both a 2 and a 3?

4.69. How many ten-digit numbers contain at least one 4, one 5, and one 7?

4.70. How many bridge hands are void in clubs and have at least one card of value p for each prime $p < 10$?

4.71. How many surjections $f : \{1, 2, \ldots, m\} \to \{1, 2, \ldots, n\}$ have the property that $f(x) = 1$ for exactly one $x \leq m$?

4.72. (a) What is the chromatic polynomial for the 4-cycle C_4? (b) For each coefficient of this chromatic polynomial, draw the vertex-spanning subgraphs of C_4 counted by that coefficient.

4.73. For even $n \geq 2$, determine the number of integers $x \leq n$ with $\gcd(x, n) = 2$.

4.74. For $k \geq 0$ and $m \geq 1$, let $\sigma_k(m) = \sum_{d|m} d^k$. (a) Find a formula for $\sigma_k(m)$ in terms of the prime factorization of m. (b) Find a formula for m^k involving σ_k and μ.

4.75. Use 4.20 to show that $\phi(mn) = \phi(m)\phi(n)$ iff $\gcd(m, n) = 1$.

4.76. Explicitly compute how the first involution discussed in 4.6 matches up the 24 objects counted by $\sum_{k=1}^{4} s(4, k)$ into pairs of objects of opposite sign.

4.77. Suppose w has cycles (1), (2), $(3, 8, 7)$, $(5, 6, 9)$, (4), and

$$U = \{\{(1)\}, \{(2)\}, \{(4), (5, 6, 9)\}, \{(3, 8, 7)\}\}.$$

Compute $I(w, U)$, where I is the involution defined at the end of 4.6.

4.78. Consider the derangement $w = 436215 \in D_6$. Find the six derangements in D_7 and the seven derangements in D_8 that can be built from w by the construction in the proof of 4.24.

4.79. Use the recursion for derangements in 4.25 to give a proof by induction of the summation formula for derangements in 4.23.

4.80. Give the details of the proof of 4.39.

4.81. Show that if G is a simple graph with c connected components, then the chromatic polynomial $\chi_G(x)$ must be divisible by x^c.

4.82. (a) Give an algebraic proof that $\sum_{k=0}^{n}\binom{n}{k}2^k(-1)^{n-k} = 1$ for $n \geq 0$. (b) Prove the identity in (a) using an involution.

4.83. For integers $a \geq b > 0$, evaluate $\sum_{k=0}^{n}\binom{n}{k}a^{n-k}(-b)^k$ by using an involution.

4.84. For integers $0 \leq a \leq b \leq c$, evaluate $\sum_{k=a}^{b}(-1)^k\binom{c}{k}$ by using an involution.

4.85. Let $S \subseteq T$ be given finite sets. (a) Use an involution to prove $\sum_{U:\ S\subseteq U\subseteq T}(-1)^{|T\sim U|} = \chi(S = T)$ (cf. 4.45). (b) In a similar manner, evaluate $\sum_{U:\ S\subseteq U\subseteq T}(-1)^{|U\sim S|}$.

4.86. Let $d, e \in \mathbb{N}^+$ with $d|e$. Use an involution to prove $\sum_{k:\ d|k|e}\mu(e/k) = \chi(d = e)$. Interpret this result in terms of the Möbius function of a poset.

4.87. Count the $n \times n$ matrices A with entries in $\{0, 1, 2\}$ such that: (a) no row of A contains all zeroes; (b) every column of A contains at least one zero; (c) there is no index j with $A(i, j) > 0$ and $A(j, i) > 0$ for all i.

4.88. An *arrowless vertex* in a simple digraph D is a vertex with indegree and outdegree zero. How many simple digraphs with vertex set $\{1, 2, \ldots, n\}$ have no arrowless vertices?

4.89. An *isolated vertex* in a simple digraph D is a vertex v such that there is no edge (u, v) or (v, u) in D with $u \neq v$. How many simple digraphs with vertex set $\{1, 2, \ldots, n\}$ have no isolated vertices?

4.90. How many simple graphs with vertex set $\{1, 2, \ldots, n\}$ have no isolated vertices?

4.91. Use 4.11 to compute the chromatic polynomial of the paw graph (see 3.124).

4.92. (a) How many anagrams in $\mathcal{R}(1^3 2^3 \cdots n^3)$ never have three equal letters in a row? (b) How many anagrams in $\mathcal{R}(1^k 2^k \cdots n^k)$ never have k equal letters in a row?

4.93. (a) Count the permutations w of $\{1, 2, \ldots, n\}$ such that $w_{i+1} \neq w_i + 1$ for all $i < n$. (b) Express your answer to (a) in terms of the derangement numbers d_k.

4.94. Given sequences $0 \leq a_1 \leq a_2 \leq \cdots \leq a_k \leq A$ and $0 \leq b_1 \leq b_2 \leq \cdots \leq b_k \leq B$, use inclusion-exclusion to derive a formula for the number of lattice paths from $(0, 0)$ to (A, B) that avoid all of the points (a_i, b_i) for $1 \leq i \leq k$.

4.95. Recursion for Möbius functions. (a) Show that the Möbius function of a poset $(X, \leq)$ can be computed recursively via $\mu(x, z) = -\sum_{y:\, x\leq y<z}\mu(x, y)$ for $x < z$, with initial conditions $\mu(x, x) = 1$ and $\mu(x, z) = 0$ whenever $x \not\leq z$. (b) Show that the Möbius function also satisfies the recursion $\mu(x, z) = -\sum_{y:\, x<y\leq z}\mu(y, z)$ for $x < z$.

4.96. Poset Associated to a DAG. Suppose $G = (X, R)$ is a DAG. Prove that there exists a unique smallest irreflexive, transitive relation $<$ that contains R. The corresponding poset $(X, \leq)$ is called the *poset associated to the DAG G*.

4.97. Let $(X, \leq)$ be the poset associated to the DAG

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

Compute the Möbius function μ_X in two ways, by: (a) inverting the matrix Z; (b) enumerating signed chains in $(X, \leq)$.

4.98. Let $(X, \leq)$ be the poset associated to the DAG

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

Compute the Möbius function μ_X in two ways, by: (a) inverting the matrix Z; (b) enumerating signed chains in $(X, \leq)$.

4.99. A *subposet* of a poset $(X, \leq)$ is a poset $(Y, \leq')$, where Y is a subset of X, and for $a, b \in Y$, $a \leq' b$ iff $a \leq b$. An *interval* in X is a subposet of the form $[x, z] = \{y \in X : x \leq y \leq z\}$. Show that for all $a, b, c, d \in X$, if the intervals $[a, b]$ and $[c, d]$ are isomorphic posets, then $\mu_X(a, b) = \mu_X(c, d)$.

4.100. Assume that X_1 and X_2 are finite disjoint sets. The *disjoint union* of the posets $(X_1, \leq_1)$ and $(X_2, \leq_2)$ is $(X, \leq)$ where $X = X_1 \cup X_2$ and for $a, b \in X$, $a \leq b$ iff $a, b \in X_1$ and $a \leq_1 b$, or $a, b \in X_2$ and $a \leq_2 b$. Determine μ_X in terms of μ_{X_1} and μ_{X_2}.

4.101. Given a poset $(X, \leq)$, define a new poset $(Y, \leq')$ by setting $Y = X \cup \{0\}$ (where 0 is a new symbol not in X), and letting $\leq'$ be the extension of $\leq$ such that $0 \leq' y$ for all $y \in Y$. Informally, $(Y, \leq')$ is obtained from $(X, \leq)$ by adjoining a new least element. Determine μ_Y in terms of μ_X.

4.102. Given posets $(X_1, \leq_1)$ and $(X_2, \leq_2)$ where X_1 and X_2 are finite disjoint sets, define a new poset $(X, \leq)$ by setting $X = X_1 \cup X_2$ and, for $a, b \in X$, $a \leq b$ iff $a, b \in X_i$ and $a \leq_i b$ $(i = 1, 2)$, or $a \in X_1$ and $b \in X_2$. Informally, $(X, \leq)$ is obtained from X_1 and X_2 by making everything in X_1 less than everything in X_2. Determine μ_X in terms of μ_{X_1} and μ_{X_2}.

4.103. Let $S_1, \ldots, S_n$ be any events in a sample space X with probability measure P. State and prove an analogue of 4.7 that can be used to compute $P(S_1 \cup \cdots \cup S_n)$.

4.104. Let $S_1, \ldots, S_n$ be independent events in a sample space X (see 1.84). Prove that for $1 \leq i \leq n$, the events $S_1, S_2, \ldots, X \sim S_i, \ldots, S_n$ are independent.

4.105. Let $S_1, \ldots, S_n$ be independent events in a sample space X, with $P(S_i) = p_i$ for each i. Find the probability that none of the events S_i occurs: (a) using inclusion-exclusion and the generalized distributive law; (b) using 4.104.

4.106. Use an involution to prove that for all $i, n \in \mathbb{N}$, $\sum_{k=0}^{i}(-1)^k\binom{n}{k}\binom{n-k}{i-k} = \chi(i = 0)$.

4.107. Use an involution to prove that for $0 \leq k \leq n$, $\sum_{i=k}^{n}(-1)^{i-k}\binom{n}{i}\binom{i}{k}2^{n-i} = \binom{n}{k}$.

4.108. Prove that for all $n, j > 0$, $n^j = \sum_{k=0}^{j}(-1)^{j-k}k!S(j,k)\binom{n+k-1}{k}$.

4.109. For $n > 0$, evaluate $\sum_{k=0}^{n-1}(-1)^k\binom{n}{k}(n-k)^n$.

4.110. Use an involution to prove the following identity satisfied by Catalan numbers: $C_n = \sum_{1 \leq k \leq (n+1)/2}(-1)^{k-1}C_{n-k}\binom{n+1-k}{k}$.

4.111. Let A be an $n \times n$ matrix with $A(i,j) = \binom{i-1}{j-1}$ for $1 \leq i, j \leq n$. (a) Look at small examples to guess a formula for $A^{-1}(i,j)$. (b) Prove your guess using an involution.

4.112. How many bijections $f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ are such that the functional digraph of f contains no cycle of length k?

4.113. How many anagrams in $\mathcal{R}(a^3b^3c^3d^3)$ never have two consecutive equal letters?

4.114. Prove or disprove: for every integer $y \geq 1$, there exist only finitely many integers $x \geq 1$ with $\phi(x) = y$.

4.115. How many compositions of n have k parts each of size at most m?

4.116. Call a function $f : X \to Y$ *doubly surjective* iff for all $y \in Y$, there exist at least two $x \in X$ with $f(x) = y$. Count the number of doubly surjective functions from an m-element set to an n-element set, where $m \geq 2n$. What is the answer when $m = 11$ and $n = 4$?

4.117. (a) Let $S_1, \ldots, S_n$ be subsets of a finite set X. Prove that the number of elements of X that lie in exactly k of the sets S_i is

$$\sum_{i=0}^{n-k} (-1)^i \binom{k+i}{i} \sum_{1 \leq j_1 < j_2 < \cdots < j_{k+i} \leq n} |S_{j_1} \cap \cdots \cap S_{j_{k+i}}|.$$

(b) Find and prove a similar formula for the number of elements of X that lie in at least k of the sets S_i.

4.118. For $0 \leq k \leq n$, let $d_{n,k}$ be the number of permutations of n objects that have exactly k fixed points. (a) Use 4.117 to find a formula for $d_{n,k}$. (b) Give algebraic and combinatorial proofs that $d_{n,k} = \binom{n}{k} d_{n-k}$.

4.119. How many integers between 1 and 2311 are divisible by exactly two of the primes in $\{2, 3, 5, 7\}$? (Use 4.117.)

4.120. Let (F_n) be the Fibonacci sequence ($F_0 = 0$, $F_1 = 1$, $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$). Find a formula for $\sum_{k=0}^{n} (-1)^k F_k$ and prove it, either algebraically or using an involution.

4.121. Find and prove a formula for $\sum_{k=0}^{n} (-1)^k F_k F_{n-k}$.

4.122. For each integer $x \geq 1$, evaluate $\sum_{k=1}^{x} \mu(k) \lfloor x/k \rfloor$.

4.123. For $n > 0$, evaluate $\sum_{k=1}^{n} (-1)^k \operatorname{Surj}(n, k)$.

4.124. For $n > 0$, evaluate $\sum_{k=1}^{n-1} (-1)^k (k-1)! S(n, k)$.

4.125. Consider an $n \times n$ lower-triangular matrix A such that $A(n, k)$ is the number of Dyck paths ending with exactly k east steps, for $1 \leq k \leq n$. Find a combinatorial description of A^{-1}, and prove that this is the inverse of A using an involution.

4.126. Garsia-Milne Involution Principle. Suppose I and J are involutions defined on finite signed sets X and Y, respectively. Suppose $f : X \to Y$ is a *sign-preserving* bijection, i.e., $\operatorname{sgn}(f(x)) = \operatorname{sgn}(x)$ for all $x \in X$. Suppose also that every object in $\operatorname{Fix}(I)$ and $\operatorname{Fix}(J)$ has positive sign. Construct an explicit bijection $g : \operatorname{Fix}(I) \to \operatorname{Fix}(J)$.

4.127. Bijective Subtraction. Suppose A, B, and C are finite, pairwise disjoint sets and $f : A \cup B \to A \cup C$ is a given bijection. Construct an explicit bijection $g : B \to C$.

4.128. Bijective Division by Two. Suppose A and B are finite sets. Given a bijection $f : \{0, 1\} \times A \to \{0, 1\} \times B$, can you use f to construct an explicit bijection $g : A \to B$?

4.129. In §4.1 we proved combinatorially that $\sum_k s(i, k) S(k, j) = \chi(i = j)$. Can you find a combinatorial proof that $\sum_k S(i, k) s(k, j) = \chi(i = j)$? (Compare with 2.77(d).)

4.130. Find a bijective proof of the derangement recursion $d_n = n d_{n-1} + (-1)^n$.

4.131. Let X_n be the set of set partitions of $\{1, 2, \ldots, n\}$. Define the *refinement ordering* on X_n by setting, for $P, Q \in X_n$, $P \preceq Q$ iff every block $S \in P$ is contained in some block $T \in Q$. (a) Show that $(X_n, \preceq)$ is a poset. (b) Compute the Möbius function of this poset for $1 \leq n \leq 4$. (c) Show that any interval $[P, Q]$ in X_n (see 4.99) is isomorphic to a poset $(X_k, \preceq)$ for some k. (d) Compute μ_{X_n} for all n.

Notes

A thorough treatment of posets from the combinatorial viewpoint appears in Chapter 3 of Stanley [127]. See Rota [118] for one of the seminal papers on Möbius inversion in combinatorics. A classic text on posets is the book by Birkhoff [12]. The Garsia-Milne involution principle in 4.126 was introduced in [49, 50]. For applications and extensions of this principle, the reader may consult the following sources [57, 73, 87, 88, 108, 109, 140]. An application of bijective subtraction (see 4.127) is presented in Loehr [85].