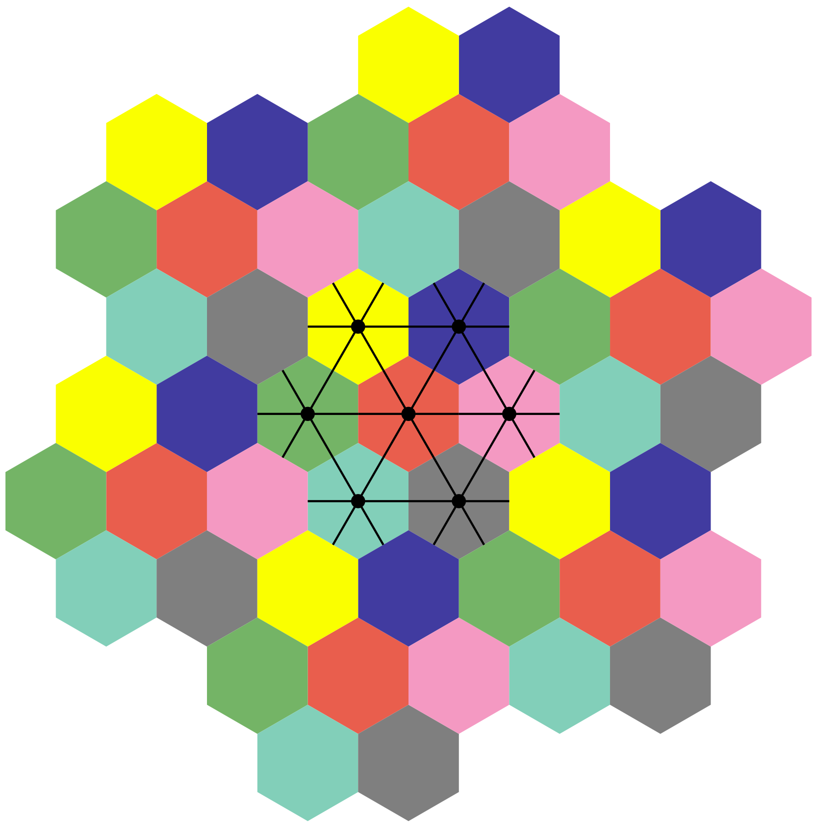
MATH 239/249

Introduction to Combinatorics



UNIVERSITY OF WATERLOO | FACULTY OF MATHEMATICS

Hello World, and Thanks!

The following parts of these notes are still UNDER CONSTRUCTION. (Estimated completion date: Fall 2021.) Most of these are additional topics not covered during the semester. The old notes are also available on the course website.

- Proof of Theorem 4.18 (inhomogeneous linear recurrence relations)
- Section 8.4 (planar duality)
- Part III: Chapters 12, 13, and 14 will appear in three installments over the course of the next year.

Department of Combinatorics and Optimization
Faculty of Mathematics
University of Waterloo

Contents

Preliminaries.

MATH 239 is an introduction to two of the main areas in combinatorics – enumeration and graph theory. MATH 249 is an advanced version of MATH 239 intended for very strong students. These courses are designed for students in the second year of an undergraduate program in mathematics or computer science.

The prerequisites required from first-year mathematics are as follows.

- From MATH 135. Abstract algebra I: sets and propositional logic, proofs, mathematical induction, modular arithmetic, complex numbers, the Fundamental Theorem of Algebra.
- From MATH 136. Linear algebra I: systems of linear equations, Gaussian elimination, matrix algebra, vector spaces.
- From MATH 137. Calculus I: algebra with power series, open/closed sets, continuous functions, differentiation (but not integration).

We use the following standard notation for various number systems.

natural numbers	$\mathbb{N} = \{0, 1, 2, 3, ...\}$ including zero 0
integers	$\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$
rational numbers	$\mathbb{Q}$
real numbers	$\mathbb{R}$
complex numbers	$\mathbb{C}$
integers (modulo n)	$\mathbb{Z}_n = \{[0], [1], [2], ..., [n-1]\}$
finite field of prime size	$\mathbb{F}_p = \mathbb{Z}_p$

The cardinality (size) of a set S is denoted by $|S|$.

For convenience, we use LHS and RHS as shorthand for "left-hand side" and "right-hand side", respectively.

Part I

Introduction to Enumeration

Overview.

Suppose I pay $5 for a lottery ticket – what is the chance that I win a share of the top prize? It depends on the details, of course. There are a certain number of ways to win, and a certain number of ways to lose. Enumeration is the art and science of figuring out this kind of thing. This is the subject of the first part of these course notes.

There are two broad principles of the subject. The combinatorial approach is to construct explicit one-to-one correspondences between sets to show that they have the same size. The algebraic approach is to translate the information about the problem from combinatorics into algebra, and then to use algebraic techniques to determine the sizes of the sets. We will see many examples of both approaches.

In Chapter 1 we begin by introducing the basic building blocks of the theory: subsets, lists and permutations, multisets, binomial coefficients, and so on. In Section 1.2 the use of these objects is illustrated by analyzing various applications and examples.

In Chapter 2 we introduce the idea of generating series. This begins with the Binomial Theorem and Binomial Series, which are of fundamental importance for later results. The general theory of generating series is developed in Section 2.2, and its use is illustrated by analyzing "compositions" in Section 2.3.

In Chapter 3 we consider the enumeration of various sets of binary strings, namely those which can be described by regular expressions – the "rational languages". This provides an interesting and varied class of examples to which the results of Chapters 2 and 4 apply.

In Chapter 4 we consider sequences which satisfy a homogeneous linear recurrence relation with initial conditions, the sequences arising in Chapters

2 and 3 being examples. This technique allows us to calculate the numbers which answer the various counting problems we have been considering. By using Partial Fractions we can derive an even better solution to such problems, although the calculations involved are also more complicated. (We include a proof of Partial Fractions for completeness.) In Section 4.4 we briefly discuss recurrence relations which are quadratic rather than linear.

Two additional topics are discussed in Chapters 11 and 12.

Chapter 1

Basic Principles of Enumeration.

1.1 The Essential Ideas.

1.1.1 Choices – "AND" versus "OR".

In the next few pages we will often be constructing an object of some kind by repeatedly making a sequence of choices. In order to count the total number of objects we could construct we must know how many choices are available at each step, *but we must know more*: we also need to know how to combine these numbers correctly. A generally good guideline is to look for the words "AND" and "OR" in the description of the sequence of choices available. Here are a few simple examples.

Example 1.1. On a table before you are 7 apples, 8 oranges, and 5 bananas.

- *Choose an apple and a banana.*
  There are 7 choices for an apple AND 5 choices for a banana: $7\times 5 = 35$ choices in all.
- *Choose an apple or an orange.*
  There are 7 choices for an apple OR 8 choices for an orange: $7+8 = 15$ choices in all.
- *Choose an apple and either an orange or a banana.*
  There are $7 \times (8+5) = 91$ possible choices.

- *Choose either an apple and an orange, or a banana.*
  There are $(7 \times 8) + 5 = 61$ possible choices.

Generally, "AND" corresponds to multiplication and "OR" corresponds to addition. The last two of the above examples show that it is important to determine exactly how the words "AND" and "OR" combine in the description of the problem.

From a mathematical point of view, "AND" corresponds to the Cartesian product of sets. If you choose one element of the set A AND you choose one element of the set B, then this is equivalent to choosing one element of the *Cartesian product* of A and B:

$$A \times B = \{(a, b) : \ a \in A \text{ and } b \in B\},$$

which is the set of all ordered pairs of elements (a, b) with $a \in A$ and $b \in B$. In general, the cardinalities of these sets are related by the formula

$$|A \times B| = |A| \cdot |B|.$$

Similarly, from a mathematical point of view, "OR" corresponds to the union of sets. If you choose one element of the set A OR you choose one element of the set B, then this is equivalent to choosing one element of the *union* of A and B:

$$A \cup B = \{c : \ c \in A \text{ or } c \in B\},$$

which is the set of all elements c which are either in A or in B.

It is not always true that $|A \cup B| = |A| + |B|$, because any elements in both A and B would be counted twice by $|A| + |B|$. The *intersection* of A and B is the set

$$A \cap B = \{c : \ c \in A \text{ and } c \in B\},$$

which is the set of all elements c which are both in A and in B. What is generally true is that

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

(This is the first instance of the Principle of Inclusion/Exclusion, which will be discussed in general in Subsection 1.1.6.) In particular, if $A \cap B = \varnothing$ then $|A \cup B| = |A| + |B|$. Thus, in order to interpret "OR" as addition, it is important to check that the sets of choices A and B have no elements in common. Such a union of sets A and B for which $A \cap B = \varnothing$ is called a *disjoint union* of sets.

When solving enumeration problems it is usually very useful to describe a choice sequence for constructing the set of objects of interest, paying close attention to the words "AND" and "OR".

1.1.2 Lists, permutations, and subsets.

A *list* of a set S is a list of the elements of S exactly once each, in some order. For example, the lists of the set $\{1, a, X, g\}$ are:

$1aXg$	$a1Xg$	$X1ag$	$g1aX$
$1agX$	$a1gX$	$X1ga$	$g1Xa$
$1Xag$	$aX1g$	$Xa1g$	$ga1X$
$1Xga$	$aXg1$	$Xag1$	$gaX1$
$1gaX$	$ag1X$	$Xg1a$	$gX1a$
$1gXa$	$agX1$	$Xga1$	$gXa1$

A *permutation* is a list of the set $\{1, 2, ..., n\}$ for some $n \in \mathbb{N}$. A permutation $\sigma : \ a_1 a_2 ... a_n$ can be interpreted as a function $\sigma : \{1, 2, ..., n\} \to \{1, 2, ..., n\}$ by putting $\sigma(i) = a_i$ for all $1 \leq i \leq n$.

To construct a list of S we can choose any element v of S to be the first element in the list and follow this with any list of the set $S \smallsetminus \{v\}$. That is how the table above is arranged – each of the four columns corresponds to one choice of an element of $\{1, a, X, g\}$ to be the first element of the list. Within each column, all the lists of the remaining elements appear after the first element.

Let p_n denote the number of lists of an n-element set S. The first sentence of the previous paragraph is translated into the equation

$$p_n = n \cdot p_{n-1},$$

provided that n is positive. (In this equation there are n choices for the first element v of the list, AND p_{n-1} choices for the list of $S \smallsetminus \{v\}$ which follows it.) It is important to note here that each list of S will be produced exactly once by this construction.

Since it is easy to see that $p_1 = 1$ (and $p_2 = 2$), a simple proof by induction on n shows the following:

Theorem 1.2. *For every $n \geq 1$, the number of lists of an n-element set S is*

$$n(n-1)(n-2)\cdots 3\cdot 2\cdot 1.$$

In particular, taking $S = \{1, 2, ..., n\}$, this is the number of permutations of size n. The term *n factorial* is used for the number $n(n-1)\cdots 3\cdot 2\cdot 1$, and it is denoted by $n!$ for convenience. We also define $0!$ to be the number of lists of the 0-element (empty) set $\varnothing$. Since we want the equation $p_n = n\cdot p_{n-1}$ to hold when $n = 1$, and since $p_1 = 1! = 1$, we conclude that $0! = p_0 = 1$ as well.

A *subset* of a set S is a collection of some (perhaps none or all) of the elements of S, at most once each and in no particular order.

To specify a particular subset A of S, one has to decide for each element v of S whether v is in A or v is not in A. Thus we have two choices – $v \in A$ OR $v \notin A$ – for each element v of S. If $S = \{v_1, v_2, ..., v_n\}$ has n elements then the total number of choices is 2^n since we have 2 choices for v_1 AND 2 choices for v_2 AND . . . AND 2 choices for v_n.

Theorem 1.3. *For every $n \geq 0$, the number of subsets of an n-element set is 2^n.*

A *partial list* of a set S is a list of a subset of S. That is, it is a list of some (perhaps none or all) of the elements of S, at most once each and listed in some particular order. We are going to count partial lists of length k of an n-element set.

First think about the particular case $n = 6$ and $k = 3$, and the set $S = \{a, b, c, d, e, f\}$. A partial list of S of length 3 is a list xyz of elements of S, which must all be different. There are:
6 choices for x (since x is in S), AND
5 choices for y (since $y \in S$ but $y \neq x$), AND
4 choices for z (since $z \in S$ but $z \neq x$ and $z \neq y$).
Altogether there are $6\cdot 5\cdot 4 = 120$ partial lists of $\{a, b, c, d, e, f\}$ of length 3.

This kind of reasoning works just as well in the general case. If S is an n-element set and we want to construct a partial list $v_1v_2...v_k$ of elements of S of length k, then there are:
n choices for v_1, AND
$n-1$ choices for v_2, AND
....
$n-(k-2)$ choices for v_{k-1}, AND
$n-(k-1)$ choices for v_k.

This proves the following result.

Theorem 1.4. *For $n, k \geq 0$, the number of partial lists of length k of an n-element set is $n(n-1)\cdots(n-k+2)(n-k+1)$.*

Notice that if $k > n$ then the number 0 will appear as one of the factors in the product $n(n-1)\cdots(n-k+2)(n-k+1)$. This makes sense, because if $k > n$ then there are no partial lists of length k of an n-element set. On the other hand, if $0 \leq k \leq n$ then we could also write this product as

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

We next count subsets of an n-element set S which have a particular size k. So for $n, k \geq 0$ let $\binom{n}{k}$ denote the number of k-element subsets of an n-element set S. Notice that if $k < 0$ or $k > n$ then $\binom{n}{k} = 0$ because in these cases it is impossible for S to have a k-element subset. Thus we need only consider k in the range $0 \leq k \leq n$.

To count k-element subsets of S we consider another way of constructing a partial list of length k of S. Specifically, we can choose a k-element subset A of S AND a list of A. The result will be a list of a subset of S of length k. Since every partial list of length k of S is constructed exactly once in this way, this translates into the equation

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

In summary, we have proved the following result.

Theorem 1.5. *For $0 \le k \le n$, the number of k-element subsets of an n-element set is*

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

The numbers $\binom{n}{k}$ are read as "n choose k" and are called *binomial coefficients* .

1.1.3 Think of what the numbers mean.

Usually, when faced with a formula to prove, one's first thought is to prove it by algebraic calculations, or perhaps with an induction argument, or maybe with a combination of the two. But often that is not the easiest way, nor is it the most informative. A much better strategy is one which gives some insight into the meaning of all of the parts of the formula. If we can interpret all the numbers as counting things, addition as "OR", and multiplication as "AND", then we can hope to find an explanation of the formula by constructing some objects in the correct way.

Example 1.6. Consider the equation, for any $n \ge 0$:

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n.$$

This could be proved by induction on n, but many more details would have to be given and the proof would not address the true "meaning" of the formula. Instead, let's interpret everything combinatorially:

- the number of subsets of the n-element set $\{1, 2, ..., n\}$ is 2^n;
- for each $0 \le k \le n$, the number of k-element subsets of $\{1, 2, ..., n\}$ is $\binom{n}{k}$;
- addition corresponds to "OR" (that is, disjoint union of sets).

So, this formula is saying that choosing a subset of $\{1, 2, ..., n\}$ (in one of 2^n ways) is equivalent to choosing a k-element subset of $\{1, 2, ..., n\}$ (in one of $\binom{n}{k}$ ways) for exactly one value of k in the range $0 \le k \le n$. Said that way the formula becomes self–evident, and there is nothing more to prove.

Example 1.7. Consider the equation

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

where we are using the fact that $\binom{m}{j} = 0$ if $j < 0$ or $j > m$.

This equation can be proven algebraically from the formula of Theorem 1.5, and that is a good exercise which I encourage you to try. But a more informative proof interprets these numbers combinatorially as follows:

- $\binom{n}{k}$ is the number of k-element subsets of $\{1, 2, ..., n\}$;
- $\binom{n-1}{k-1}$ is the number of $(k-1)$-element subsets of $\{1, 2, ..., n-1\}$;
- $\binom{n-1}{k}$ is the number of k-element subsets of $\{1, 2, ..., n-1\}$;
- addition corresponds to disjoint union of sets.

So, this equation is saying that choosing a k-element subset A of $\{1, 2, ..., n\}$ is equivalent to either choosing a $(k-1)$-element subset of $\{1, 2, ..., n-1\}$ or a k-element subset of $\{1, 2, ..., n-1\}$. This is perhaps not as clear as the previous example, but the two cases depend upon whether the chosen k-element subset A of $\{1, 2, ..., n\}$ is such that $n \in A$ OR $n \notin A$. If $n \in A$ then $A \smallsetminus \{n\}$ is a $(k-1)$-element subset of $\{1, 2, ..., n-1\}$, while if $n \notin A$ then A is a k-element subset of $\{1, 2, ..., n-1\}$. This construction explains the correspondence, proving the formula.

This principle – interpreting equations combinatorially and proving the formulas by describing explicit correspondences between sets of objects – is one of the most important and powerful ideas in enumeration. We will apply this way of thinking throughout the first part of these notes.

Incidentally, the equation in Example 1.7 is a very useful recurrence relation for computing binomial coefficients quickly. Together with the facts

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

and $\binom{n}{0} = \binom{n}{n} = 1$ it can be used to compute any number of binomial coefficients without difficulty. The resulting table is known as *Pascal's Triangle*:

$n\backslash k$	0	1	2	3	4	5	6	7	8
0	1								
1	1	1							
2	1	2	1						
3	1	3	3	1					
4	1	4	6	4	1				
5	1	5	10	10	5	1			
6	1	6	15	20	15	6	1		
7	1	7	21	35	35	21	7	1	
8	1	8	28	56	70	56	28	8	1

1.1.4 Multisets.

Imagine a bag which contains a large number of marbles of three colours – red, green, and blue, say. The marbles are all indistinguishable from one another except for their colours. There are N marbles of each colour, where N is very, very large (more precisely we should be considering the limit as $N \to \infty$). If I reach into the bag and pull out a handful of 11 marbles, I will have r red marbles, g green marbles, and b blue marbles, for some nonnegative integers (r, g, b) such that $r + g + b = 11$. How many possibile outcomes are there?

The word "multiset" is meant to suggest a set in which the objects can occur more than once. For example, the outcome $(4, 5, 2)$ in the above situation corresponds to the "set" $\{R, R, R, R, G, G, G, G, G, B, B\}$ in which R is a red marble, G is a green marble, and B is a blue marble. This is an 11-element multiset with elements of three types. The number of these multisets is the solution to the above problem.

Definition 1.8. Let $n \geq 0$ and $t \geq 1$ be integers. A *multiset of size n with elements of t types* is a sequence of nonnegative integers $(m_1, ..., m_t)$ such that

$$m_1 + m_2 + \cdots + m_t = n.$$

The interpretation is that m_i is the number of elements of the multiset which are of the i-th type, for each $1 \leq i \leq t$.

Theorem 1.9. *For any $n \geq 0$ and $t \geq 1$, the number of n-element multisets with elements of t types is*

$$\binom{n+t-1}{t-1}.$$

Proof. Think of what that number means! By Theorem 1.5, $\binom{n+t-1}{t-1}$ is the number of $(t-1)$-element subsets of an $(n+t-1)$-element set. So, let's write down a row of $(n+t-1)$ circles from left to right:

O O O O O O O O O O O O O

and cross out some $t-1$ of these circles to choose a $(t-1)$-element subset:

O O O O X O O O O O X O O

Now the $t-1$ crosses chop the remaining sequence of n circles into t segments of consecutive circles. (Some of these segments might be empty, which is to say of length zero.) Let m_i be the length of the i-th segment of consecutive O-s in this construction. Then $m_1 + m_2 + \cdots + m_t = n$, so that $(m_1, m_2, ..., m_t)$ is an n-element multiset with t types. Conversely, if $(m_1, m_2, ..., m_t)$ is an n-element multiset with t types then write down a sequence of m_1 O-s, then an X, then m_2 O-s, then an X, and so on, finishing with an X and then m_t O-s. The positions of the X-s will indicate a $(t-1)$-element subset of the positions $\{1, 2, ..., n+t-1\}$.

The construction of the above paragraph shows how to translate between $(t-1)$-element subsets of $\{1, 2, ..., n+t-1\}$ and n-element multisets with t types of element. This one–to–one correspondence completes the proof of the theorem. □

To answer the original question of this section, the number of 11-element multisets with elements of 3 types is $\binom{11+3-1}{3-1} = \binom{13}{2} = 78$.

1.1.5 Bijective proofs.

The arguments above, counting lists, permutations, subsets, multisets, and so on, can be phrased more formally using the idea of bijections between finite sets. In simple cases as we have seen so far this is not always necessary, but it is good style. In more complicated situations, as we will see in Chapters 2 to 4, it is a very useful way to organize one's thoughts.

Definition 1.10. Let $f : \mathcal{A} \to \mathcal{B}$ be a function from a set $\mathcal{A}$ to a set $\mathcal{B}$.

- The function f is *surjective* if for every $b \in \mathcal{B}$ there exists an $a \in \mathcal{A}$ such that $f(a) = b$.
- The function f is *injective* if for every $a, a' \in \mathcal{A}$, if $f(a) = f(a')$, then $a = a'$.
- The function f is *bijective* if it is both surjective and injective.
- The notation $\mathcal{A} \rightleftharpoons \mathcal{B}$ indicates that there is a bijection between the sets $\mathcal{A}$ and $\mathcal{B}$.

Functions with these properties are called surjections, injections, or bijections, respectively. An older terminology – now out of fashion – is that surjections are "onto" functions, injections are "one-to-one" functions, and bijections are "one-to-one and onto". By Exercise 1.4(a), the relation $\rightleftharpoons$ is an equivalence relation.

The point of Definition 1.10 is the following. Consider a bijection $f : \mathcal{A} \to \mathcal{B}$. Then every $b \in \mathcal{B}$ is the image of at least one $a \in \mathcal{A}$, since f is surjective. On the other hand, every $b \in \mathcal{B}$ is the image of at most one $a \in \mathcal{A}$, since f is injective. Therefore, every $b \in \mathcal{B}$ is the image of exactly one $a \in \mathcal{A}$. In other words, the relation $f(a) = b$ pairs off all the elements of $\mathcal{A}$ with all the elements of $\mathcal{B}$. It follows that $\mathcal{A}$ and $\mathcal{B}$ have the same number of elements. That is, if $\mathcal{A} \rightleftharpoons \mathcal{B}$ then $|\mathcal{A}| = |\mathcal{B}|$. The converse implication holds, and for infinite sets the relation $\mathcal{A} \rightleftharpoons \mathcal{B}$ is taken as the *definition* of two sets "having the same size".

Proposition 1.11. *Let $f : \mathcal{A} \to \mathcal{B}$ and $g : \mathcal{B} \to \mathcal{A}$ be functions between two sets $\mathcal{A}$ and $\mathcal{B}$. Assume the following.*

- *For all $a \in \mathcal{A}$, $g(f(a)) = a$.*
- *For all $b \in \mathcal{B}$, $f(g(b)) = b$.*

Then both f and g are bijections. Moreover, for $a \in \mathcal{A}$ and $b \in \mathcal{B}$, we have $f(a) = b$ if and only if $g(b) = a$.

Proof. Exercise 1.4(b). □

A pair of functions as in Proposition 1.11 are called *mutually inverse bijections*. The notation $g = f^{-1}$ and $f = g^{-1}$ is used to denote this relation. Notice that for a bijection f, we have $(f^{-1})^{-1} = f$.

Here are two examples of this way of thinking.

Example 1.12 (Subsets and indicator vectors.)**.** Let $\mathcal{P}(n)$ be the set of all subsets of $\{1, 2, ..., n\}$, and let $\{0, 1\}^n$ be the set of all *indicator vectors* $\alpha = (a_1, a_2, ..., a_n)$ in which each coordinate is either 0 or 1. There is a bijection between these two sets. For a subset $S \subseteq \{1, 2, ..., n\}$, define the vector $\alpha(S) = (a_1(S), a_2(S), ..., a_n(S))$ by saying that for each $1 \leq i \leq n$,

$$a_i(S) = \begin{cases} 1 & \text{if } i \in S, \\ 0 & \text{if } i \notin S. \end{cases}$$

Conversely, for an indicator vector $\alpha = (a_1, ..., a_n)$ define a subset $S(\alpha)$ by saying that

$$S(\alpha) = \{i \in \{1, 2, ..., n\} : \ a_i = 1\}.$$

For example, when $n = 8$ the subset $\{2, 3, 5, 7\}$ corresponds to the indicator vector $(0, 1, 1, 0, 1, 0, 1, 0)$. The constructions $S \mapsto \alpha(S)$ and $\alpha \mapsto S(\alpha)$ are mutually inverse bijections between the sets $\mathcal{P}(n)$ and $\{0, 1\}^n$ as in Proposition 1.11. It follows that $|\mathcal{P}(n)| = |\{0, 1\}^n| = 2^n$. This is a formalization of the proof of Theorem 1.3.

Example 1.13 (Subsets and multisets.)**.** The proof of Theorem 1.9 can be phrased in terms of bijections, as follows.

Let $\mathcal{M}(n, t)$ be the set of all multisets of size $n \in \mathbb{N}$ with elements of $t \geq 1$ types. Let $\mathcal{B}(a, k)$ be the set of all k-element subsets of $\{1, 2, ..., a\}$. We establish a bijection between $\mathcal{M}(n, t)$ and $\mathcal{B}(n+t-1, t-1)$ in what follows. Theorem 1.5 implies that $|\mathcal{B}(n + t - 1, t - 1)| = \binom{n+t-1}{t-1}$, completing the proof of Theorem 1.9. Here is a precise description of this bijection.

Let S be any $(t - 1)$-element subset of $\{1, 2, ..., n + t - 1\}$. We can sort the elements of S in increasing order: $S = \{s_1, s_2, ..., s_{t-1}\}$ in which $s_1 < s_2 < \cdots < s_{t-1}$. For notational convenience, let $s_0 = 0$ and let $s_t = n + t$. Now define a sequence $\mu = (m_1, m_2, ..., m_t)$ by letting $m_i = s_i - s_{i-1} - 1$ for all $1 \leq i \leq t$.

For example, with $n = 10$ and $t = 4$, consider the 3-element subset $S = \{2, 7, 11\}$ of $\{1, 2, ..., 13\}$. Then $s_0 < s_1 < \cdots < s_t$ is $0 < 2 < 7 < 11 < 14$, the sequence of differences is $(2, 5, 4, 3)$, and subtracting 1 from each of these yields $\mu = (1, 4, 3, 2)$. Notice that μ is a multiset of size 10 with 4 types of elements.

In general, since $s_{i-1} < s_i$ for all $1 \leq i \leq t$, it follows that $m_i = s_i - s_{i-1} - 1$ is a nonnegative integer. Also, since $s_t = n + t$, it follows that $m_1 + m_2 + \cdots + m_t = s_t - t = n$. That is, μ is a multiset of size n with elements of t types. This describes a function $S \mapsto \mu$ from the set $\mathcal{B}(n + t - 1, t - 1)$ to the set $\mathcal{M}(n, t)$. We claim that this function is a bijection between these two sets.

To show that our construction $S \mapsto \mu$ is a bijection, we will describe its inverse function. Begin with a multiset $\mu = (m_1, m_2, ..., m_t)$ of size n with t types of elements. For each $1 \leq i \leq t-1$, let $s_i = m_1 + m_2 + \cdots + m_i + i$. Notice that

$$1 \leq s_1 < s_2 < \cdots < s_{t-1} \leq n + t - 1.$$

Therefore, $S = \{s_1, s_2, ..., s_{t-1}\}$ is a member of the set $\mathcal{B}(n + t - 1, t - 1)$.

To finish this example, one must check that these constructions, $S \mapsto \mu$ and $\mu \mapsto S$, are mutually inverse bijections as in Proposition 1.11. The details are left as Exercise 1.6.

1.1.6 Inclusion/Exclusion.

In a vase is a bouquet of flowers. Each flower is (at least one of) fresh, fragrant, or colourful:
(a) 11 flowers are fresh;
(b) 7 flowers are fragrant;
(c) 8 flowers are colourful;
(d) 6 flowers are fresh and fragrant;
(e) 5 flowers are fresh and colourful;
(f) 2 flowers are fragrant and colourful;
(g) 2 flowers are fresh, fragrant, and colourful.
How many flowers are in the bouquet?

The Principle of Inclusion/Exclusion is a systematic method for answering such questions, which involve overlapping conditions which can be sat-

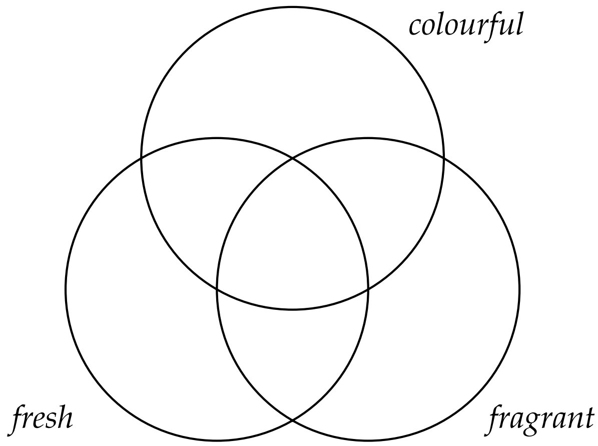


Figure 1.1: A Venn diagram for three sets.

isfied (or not) in various combinations.

Example 1.14. For a small problem as above we can reason backwards as follows:
(g): there are 2 flowers with all three properties (fresh, fragrant, and colourful);
(h): from (g) and (f) there are 0 flowers which are fragrant and colourful but not fresh;
(i): from (g) and (e) there are 3 flowers which are fresh and colourful but not fragrant;
(j): from (g) and (d) there are 4 flowers which are fresh and fragrant but not colourful;
(k): from (c)(g)(h)(i) there are 3 flowers which are colourful but neither fresh not fragrant;
(ℓ): from (b)(g)(h)(j) there is 1 flower which is fragrant but neither fresh nor colourful;
(m): from (a)(g)(i)(j) there are 2 flowers which are fresh but neither fragrant nor colourful.

The total number of flowers is counted by the disjoint union of the cases (g) through (m); that is $2+0+3+4+3+1+2 = 15$.

A Venn diagram is extremely useful for organizing this calculation. Figure 1.1 is a Venn diagram for the three sets involved in this question. Item (g) in

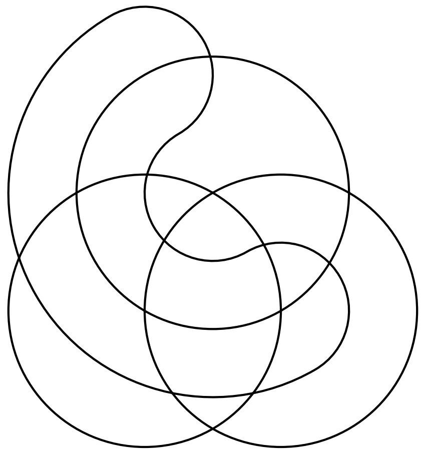


Figure 1.2: A Venn diagram for four sets.

the original data gives the number of flowers counted in the central triangle. The subsequent steps (h) to (m) calculate the rest of the numbers in the diagram, moving outwards from the center.

The above works very well for three properties (fresh, fragrant, colourful) but becomes increasingly difficult to apply as the number of properties increases. Figure 1.2 shows a Venn diagram for four sets, for instance. Instead, consider this alternative to the calculation in Example 1.14:

$$(a) + (b) + (c) - (d) - (e) - (f) + (g) = 11 + 7 + 8 - 6 - 5 - 2 + 2 = 15.$$

This looks much easier to apply, and it gives the right answer. Why? That is the Principle of Inclusion/Exclusion, which we now explain in general.

Let $A_1, A_2, ..., A_m$ be finite sets. We want a formula for the cardinality of the union of these sets $A_1 \cup A_2 \cup \cdots \cup A_m$. First a bit of notation: if S is a nonempty subset of $\{1, 2, ..., m\}$ then let A_S denote the intersection of the sets A_i for all $i \in S$. So, for example, with this notation we have $A_{\{2,3,5\}} = A_2 \cap A_3 \cap A_5$.

Theorem 1.15 (Inclusion/Exclusion). *Let A_1, A_2,..., A_m be finite sets. Then*

$$|A_1 \cup A_2 \cup \cdots \cup A_m| = \sum_{\varnothing \neq S \subseteq \{1,...,m\}} (-1)^{|S|-1} |A_S|.$$

We prove Theorem 1.15 in Section 2.5, but all that is required is the Binomial Theorem 2.2.

1.1.7 Combinatorial probabilities.

We can interpret counting problems in terms of probabilities by making one additional hypothesis. That hypothesis is that **every possible outcome is equally likely**. The exact definition of what is an "outcome" depends on the particular problem. If Ω denotes a finite set of all possible outcomes, then any subset E of Ω is what a probabilist calls an "event". The probability that a randomly chosen outcome from Ω is in the set E is $|E|/|\Omega|$ exactly because every outcome has probability $1/|\Omega|$ of being chosen, and there are $|E|$ elements in E. Here are a few examples to illustrate these ideas.

Example 1.16. What is the probability that a random subset of $\{1, 2, ..., 8\}$ has at most 3 elements?

Here an outcome is a subset of $\{1, 2, ..., 8\}$, and there are $2^8 = 256$ such subsets. The number of subsets of $\{1, 2, ..., 8\}$ with at most 3 elements is

$$\binom{8}{0} + \binom{8}{1} + \binom{8}{2} + \binom{8}{3} = 1 + 8 + 28 + 56 = 93.$$

So the probability in question is

$$\frac{93}{256} = 0.363281...$$

to six decimal places.

Example 1.17. What is the probability that a random list of $\{a, b, c, d, e, f\}$ contains the letters fad as a consecutive subsequence?

Here an outcome is a list of $\{a, b, c, d, e, f\}$, and there are $6! = 720$ such lists. Those lists of this set which contain fad as a consecutive subsequence can be constructed uniquely as the lists of the set $\{b, c, e, fad\}$, so there are $4! = 24$ of these. Thus, the probability in question is

$$\frac{24}{720} = \frac{1}{30} = 0.03333...$$

Example 1.18. What is the probability that a randomly chosen 2-element multiset with t types of element has both elements of the same type?

The outcomes are the 2-element multisets with t types, numbering

$$\binom{2+t-1}{t-1} = \binom{t+1}{t-1} = \binom{t+1}{2} = \frac{(t+1)t}{2}$$

in total. Of these, exactly t of them have both elements of the same type – choose one of the t types and take two elements of that type. Thus, the probability in question is

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

The values for the first few t are given in the following table to four decimal places:

t	1	2	3	4	5	6	7
	1.0000	0.6667	0.5000	0.4000	0.3333	0.2857	0.2500

1.2 Examples and Applications.

1.2.1 The Vandermonde convolution formula.

Example 1.19 (Vandermonde convolution formula)**.** For $m, n, k \in \mathbb{N}$,

$$\binom{m+n}{k} = \sum_{j=0}^{k} \binom{m}{j}\binom{n}{k-j}.$$

For instance, with $m = 4$ and $n = 2$ and $k = 3$ this says that

$$\binom{6}{3} = \binom{4}{0}\binom{2}{3} + \binom{4}{1}\binom{2}{2} + \binom{4}{2}\binom{2}{1} + \binom{4}{3}\binom{2}{0}.$$

(Of course, $\binom{2}{3} = 0$, but that doesn't matter.)

The Vandermonde convolution formula can be proven algebraically by induction on $m + n$, but the proof is finicky and doesn't give much insight into what the formula "means". (The formula can also be deduced easily from the Binomial Theorem **??**, as we will see in Example 2.3.)

Here is a direct **combinatorial** proof, illustrating the strategy of thinking about what the numbers mean. On the LHS, $\binom{m+n}{k}$ is the number of k-element subsets S of the set $\{1, 2, ..., m + n\}$. On the RHS, the number can be produced as follows:

- choose a value of j in the range $0 \le j \le k$, and
- choose a j-element subset A of $\{1, 2, ..., m\}$, and
- choose a $(k - j)$-element subset B of $\{m + 1, ..., m + n\}$.

(Notice that the set $\{m + 1, ..., m + n\}$ has n elements, so it has $\binom{n}{k-j}$ subsets of size $k - j$.) Now the formula is proved by describing a bijection between the k-element subsets S of $\{1, 2, ..., m+n\}$ counted on the LHS, and the pairs (A, B) of subsets counted on the RHS. To describe this correspondence, let $M = \{1, 2, ..., m\}$ and $N = \{m + 1, ..., m + n\}$. Notice that $M \cap N = \varnothing$ and $M \cup N = \{1, 2, ..., m+n\}$ and $|M| = m$ and $|N| = n$. Now, given a k-element subset S of $\{1, 2, ..., m + n\}$ we let

$$A = S \cap M \quad \text{and} \quad B = S \cap N.$$

Conversely, given a pair of subsets (A, B) satisfying the conditions in the points above, we let $S = A \cup B$. After some thought, you will see that these constructions $S \mapsto (S \cap M, S \cap N)$ and $(A, B) \mapsto A \cup B$ are mutually inverse bijections between the sets in question. Therefore, there are the same number of objects on each side, and the formula is proved.

1.2.2 Common birthdays.

Example 1.20. Let $p(n)$ denote the probability that in a randomly chosen group of n people, at least two of them are born on the same day of the year. What does the function $p(n)$ look like?

To simplify the analysis, we will ignore the existence of leap years and assume that every year has exactly 365 days. Moreover, we will also assume that people's birthdays are independently and uniformly distributed over the 365 days of the year, so that we can use the ideas of combinatorial probability theory. These are reasonable approximations – although they introduce tiny errors they do not change the qualitative "shape" of the answer.

To begin with, $p(1) = 0$ since there is only $n = 1$ person in the group. Also, if $n > 365$ then $p(n) = 1$ since there are more people in the group than days in a year, so at least two people in the group must have the same birthday.

For n in the range $2 \leq n \leq 365$ it is quite complicated to analyze the probability $p(n)$ directly. However, the complementary probability $1 - p(n)$ is relatively easy to compute. From the definition of $p(n)$ we see that $1-p(n)$ is the probability that in a randomly chosen group of n people, **no two of them** are born on the same day of the year. This model is equivalent to rolling "no pair" when throwing n independent dice each with 365 sides. If we list the people in the group as P_1, P_2, ..., P_n in any order, then their birthdays must form a partial list of the 365 days of the year, of length n. There are $365!/(365 - n)!$ such partial lists. Since the total number of outcomes is 365^n, we have derived the formula

$$1 - p(n) = \frac{365!}{(365 - n)!365^n}.$$

Therefore

$$p(n) = 1 - \frac{365!}{(365 - n)!365^n}.$$

To give some feeling for what this looks like, here is a table of $p(n)$ (rounded to six decimal places) for selected values of $2 \leq n \leq 365$.

n	$p(n)$	n	$p(n)$	n	$p(n)$
2	0.002740	25	0.568700	70	0.999160
3	0.008204	30	0.706316	80	0.999914
4	0.016356	35	0.814383	100	1.
5	0.027136	40	0.891232	150	1.
10	0.116948	45	0.940976	200	1.
15	0.252901	50	0.970374	250	1.
20	0.411438	60	0.994123	300	1.

Figure 1.3 gives a graph of this function. It is a rather surprising fact that $p(23) = 0.507297$, so that if you randomly choose a set of 23 people on earth

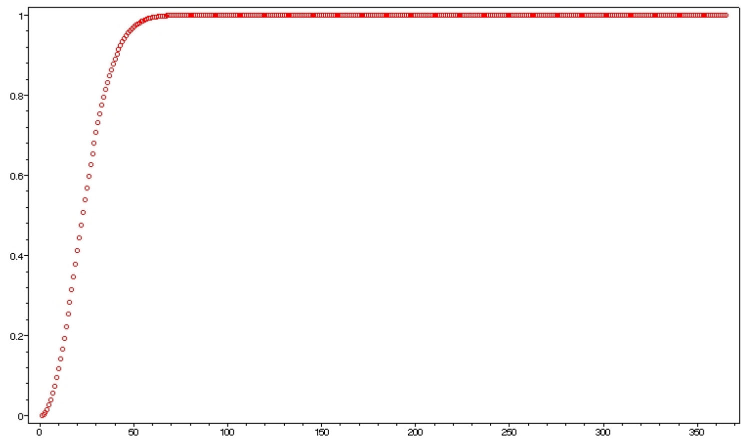


Figure 1.3: The probability of a common birthday among n people.

then there is a slightly better than 50% chance that at least two of them will have the same birthday. (Approximately – we have ignored leap years and twins.)

1.2.3 An example with multisets.

Example 1.21. A packet of *Maynard's Wine Gums* consists of a roll or packet of 10 candies, each of which has one of five "flavours" – *Green, Yellow, Orange, Red*, or *Purple*. I especially like the purple ones. What is the chance that when I buy some Wine Gums there are exactly k purple candies (for each $0 \leq k \leq 10$)?

This example is designed to illustrate the fact that the probabilities depend on which model is used to analyze the situation. There are two reasonable possibilities for this problem, which I will call the **dice model** and the **multiset model**.

In the "dice model" we keep track of the fact that the candies are stacked up in the roll from bottom to top, so there is a natural sequence $(c_1, c_2, ..., c_{10})$ of flavours one sees when the roll is opened. For example, the sequences

$$(G, P, R, Y, Y, G, O, R, Y, O)$$

and

$$(Y, G, O, P, R, R, Y, G, O, Y)$$

count as different outcomes in this model. We have a sequence of 10 candies, and a choice of one of 5 flavours for each candy, giving a total of $5^{10} = 9765625$ outcomes. (This is equivalent to rolling a sequence of ten 5-sided dice, hence the name for the model.)

In the "multiset model" we disregard the order in which the candies occur in the packet as being an inessential detail. The only important information about the packet is the number of candies of each type that it contains. For example, both of the outcomes in the previous paragraph reduce to the same multiset

$$\{G, G, Y, Y, Y, O, O, R, R, P\}$$

or $(2, 3, 2, 2, 1)$ in this model. Thus we are regarding the packet as a multiset of size 10 with 5 types of element, giving a total of $\binom{10+5-1}{5-1} = \binom{14}{4} = 1001$ outcomes.

Notice that the number of outcomes in the dice model is vastly larger than in the multiset model. It should come as no surprise, then, that the probabilities we compute depend strongly on which of these two models we consider. (The true values for the probabilities depend on the details of the manufacturing process by which the rolls or packets are made. These cannot be calculated, but must be measured instead.)

Consider the dice model first, and let $d(k)$ denote the probability of getting exactly k purple candies in a roll. There are $\binom{10}{k}$ choices for the positions of these k purple candies, and $(5-1)^{10-k}$ choices for the sequence of (non–purple) flavours of the other $10-k$ candies. This gives a total of $\binom{10}{k}4^{10-k}$ outcomes with exactly k purple candies in this model. Therefore,

$$d(k) = \binom{10}{k}\frac{4^{10-k}}{5^{10}}$$

for each $0 \leq k \leq 10$.

Next let's consider the multiset model, and let $m(k)$ denote the probability of getting exactly k purple candies in a packet. If we have k purple candies then the rest of the candies form a multiset of size $10-k$ with elements of 4 types, so there are $\binom{10-k+4-1}{4-1} = \binom{13-k}{3}$ such outcomes in this model. Therefore,

$$m(k) = \frac{\binom{13-k}{3}}{\binom{14}{4}}$$

for each $0 \le k \le 10$.

Here is a table of these probabilities (rounded to six decimal places).

k	$d(k)$	$m(k)$
0	0.107374	0.285714
1	0.268435	0.219780
2	0.301990	0.164835
3	0.201327	0.119880
4	0.088080	0.083916
5	0.026424	0.055944
6	0.005505	0.034965
7	0.000786	0.019980
8	0.000074	0.009990
9	0.000004	0.003996
10	0.000000	0.000999

The differences between the two models are clearly seen.

In closing, here are two more points about these models.

First, given a multiset $(m_1, ..., m_t)$ of size n with elements of t types, the number of outcomes in the dice model which "reduce" to this multiset is

$$\binom{n}{m_1, ..., m_t} = \frac{n!}{m_1! \cdot m_2! \cdots m_t!},$$

called a *multinomial coefficient*. This can be seen intuitively by arranging the n elements of the multiset in a line in one of $n!$ ways, and noticing that since we can't tell the m_i elements of type i apart we can freely rearrange them in $m_i!$ ways without changing the arrangement. One could prove this more carefully by induction on t, using the case $t = 2$ of binomial coefficients as part of the induction step. Or, one could give a combinatorial proof by constructing a bijection which makes the informal argument above more precise.

For the second point, the above analysis of the multiset model can be generalized to prove Exercise 1.11: for any integers $n \geq 0$ and $t \geq 2$,

$$\binom{n+t-1}{t-1} = \sum_{k=0}^{n} \binom{n-k+t-2}{t-2}.$$

1.2.4 Poker hands.

Poker is played with a *standard deck* of 52 cards, divided into four *suits*:
spades $\spadesuit$, hearts $\heartsuit$, diamonds $\diamondsuit$, and clubs $\clubsuit$.
Within each suit are 13 cards of different *values*:
A (Ace), 2, 3, 4, 5, 6, 7, 8, 9, 10, J (Jack), Q (Queen), K (King).
An Ace can be *high* (above K) or *low* (below 2) at the player's choice.

Many variations on the game exist, but the common theme is to make the best five-card hand according to the ranking of poker hands. This order is determined by how unlikely it is to be dealt such a hand. From best to worst, the types of poker hand are as follows:

- Straight Flush: five cards of the same suit with consecutive values. For example, $8\heartsuit\ 9\heartsuit\ 10\heartsuit\ J\heartsuit\ Q\heartsuit$.
- Four of a Kind (or Quads): four cards of the same value, with any fifth card. For example, $7\spadesuit\ 7\heartsuit\ 7\diamondsuit\ 7\clubsuit\ 4\diamondsuit$.
- Full House (or Tight, or Boat): three cards of the same value, and a pair of cards of another value. For example, $9\spadesuit\ 9\heartsuit\ 9\diamondsuit\ A\diamondsuit\ A\clubsuit$.
- Flush: five cards of the same suit, but not with consecutive values. For example, $3\heartsuit\ 7\heartsuit\ 10\heartsuit\ J\heartsuit\ K\heartsuit$.
- Straight: five cards with consecutive values, but not of the same suit. For example, $8\heartsuit\ 9\clubsuit\ 10\spadesuit\ J\heartsuit\ Q\diamondsuit$.
- Three of a Kind (or Trips): three cards of the same value, and two more cards not of the same value. For example, $8\spadesuit\ 8\heartsuit\ 8\diamondsuit\ K\diamondsuit\ 5\clubsuit$.
- Two Pair: this is self-explanatory. For example, $J\heartsuit\ J\clubsuit\ 6\diamondsuit\ 6\clubsuit\ A\spadesuit$.
- One Pair: this is also self-explanatory. For example, $Q\spadesuit\ Q\diamondsuit\ 8\diamondsuit\ 7\clubsuit\ 2\spadesuit$.
- Busted Hand: anything not covered above. For example, $K\spadesuit\ Q\diamondsuit\ 8\diamondsuit\ 7\clubsuit\ 2\spadesuit$.

There are $\binom{52}{5} = 2598960$ possible 5-element subsets of a standard deck of 52 cards, so this is the total number of possible poker hands. How many of these hands are of each of the above types? The answers can be found easily on the WWWeb, so there's no sense trying to keep them secret. Here they are: N is the number of outcomes of each type, and $p = N/\binom{52}{5}$ is the probability of each type of outcome (rounded to six decimal places).

Hand	N	p
Straight Flush	40	0.000015
Quads	624	0.000240
Full House	3744	0.001441
Flush	5108	0.001965
Straight	10200	0.003925
Trips	54912	0.021128
Two Pair	123552	0.047539
One Pair	1098240	0.422569
Busted	1302540	0.501177

The derivation of these numbers is excellent practice (see Exercise 1.14), so we will do only two of the cases – Straight, and Busted – as illustrations.

Example 1.22.

- To construct a Straight hand there are 10 choices for the consecutive values of the cards ($A2345$, 23456, ... up to $10JQKA$), and 4^5 choices for the suits on the cards. However, four of these choices for suits give all five cards the same suit – these lead to straight flushes and must be discounted. Hence the total number of straights is $10 \cdot (4^5 - 4) = 10200$.
- To construct a Busted hand there are $\binom{13}{5} - 10$ choices for 5 values of cards which are not consecutive (no straight) and have no pairs. Having chosen these values there are $4^5 - 4$ choices for the suits on the cards which do not give all five cards the same suit (no flush). Hence the total number of busted hands is $\left(\binom{13}{5} - 10\right) \cdot (4^5 - 4) = 1302540$.

1.2.5 Derangements.

Here is a "classical" example. (In this context, the word "derangement" makes more sense in French than in English.)

> **Example 1.23** (The Derangement Problem)**.** A group of eight people meet for dinner at a fancy restaurant and check their coats at the door. After a delicious gourmet meal the group leaves, and on the way out the eight coats are returned to the eight people completely at random by an incompetent clerk. What is the probability that no-one gets the correct coat?

Of course, we want to solve the derangement problem for any number of people, not just for eight. To state the problem mathematically, list the people P_1, P_2, ..., P_n in any order. We can record who gets whose coat by a sequence of numbers $(c_1, c_2, ..., c_n)$ in which $c_i = j$ means that P_i was given the coat belonging to P_j. The sequence $(c_1, c_2, ..., c_n)$ will thus contain each of the numbers $1, 2, ..., n$ exactly once in some order. In other words, $(c_1, ..., c_n)$ is a permutation of the set $\{1, 2, ..., n\}$, and we assume that this permutation is chosen randomly by the incompetent clerk. Person i gets the correct coat exactly when $c_i = i$. Thus, in general the derangement problem is to determine, for a random permutation $(c_1, ..., c_n)$ of $\{1, 2, ..., n\}$, the probability that $c_i \neq i$ for all $1 \leq i \leq n$.

For small values of n the derangement problem can be analyzed directly, but complications arise as n gets larger. In fact, this example is perfectly designed to illustrate the principle of Inclusion/Exclusion. To see how this applies, for each $1 \leq i \leq n$ let A_i be the set of permutations of $\{1, ..., n\}$ such that $c_i = i$. That is, A_i is the set of ways in which the coats are returned and person P_i gets the correct coat. From that interpretation, the union of sets $A_1 \cup A_2 \cup \cdots \cup A_n$ is the set of ways in which the coats are returned and **at least one person** gets the correct coat. Therefore, the complementary set of permutations gives those ways of returning the coats so that **no-one** gets the correct coat. The number of these *derangements* of n objects is thus

$$n! - |A_1 \cup A_2 \cup \cdots \cup A_n|.$$

It remains to apply Inclusion/Exclusion to determine $|A_1 \cup \cdots \cup A_n|$. To do this we need to determine $|A_S|$ for every nonempty subset $\varnothing \neq S \subseteq$

$\{1,2,...,n\}$. Consider the example $n = 8$ and $S = \{2,3,6\}$. In this case $A_{\{2,3,6\}} = A_2 \cap A_3 \cap A_6$ is the set of those permutations of $\{1,...,8\}$ such that $c_2 = 2$ and $c_3 = 3$ and $c_6 = 6$. Such a permutation looks like $\square\ 2\ 3\ \square\ \square\ 6\ \square\ \square$ in which the boxes are filled with the numbers $\{1,4,5,7,8\}$ in some order. Since there are $5!$ lists of the set$\{1,4,5,7,8\}$ it follows that $|A_{\{2,3,6\}}| = 5! = 120$ in this case. The general case is similar. If $\varnothing \neq S \subseteq \{1,2,...,n\}$ is a k-element subset then the permutations of $\{1,2,...,n\}$ in A_S are obtained by fixing $c_i = i$ for all $i \in S$, and then listing the remaining $n-k$ elements of $\{1,...,n\} \setminus S$ in the remaining spaces. Since there are $(n-k)!$ such lists we see that $|A_S| = (n-k)!$.

Since $|A_S| = (n-k)!$ for every k-element subset of $\{1,2,...,n\}$, and there are $\binom{n}{k}$ such k-element subsets, Inclusion/Exclusion implies that

$$|A_1 \cup \cdots \cup A_n| \;=\; \sum_{k=1}^{n} \binom{n}{k} (-1)^{k-1}(n-k)! \;=\; n! \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!}.$$

It follows that the number of derangements of n objects is

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

Since the total number of permutations of n objects is $n!$, the probability that a randomly chosen permutation of $\{1,2,...,n\}$ is a derangement is

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

The following table lists the first several values of the function D_n (with the decimals rounded to six places). Notice that for $n \geq 7$ the value of D_n changes very little. If you recall the Taylor series expansion of the exponential function

$$\mathrm{e}^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

then it is easy to see that as n tends to infinity the probability D_n approaches the limiting value of $\mathrm{e}^{-1} = 0.3678794412...$.

n	D_n	D_n
0	$1/1$	1.
1	$0/1$	0.000000
2	$1/2$	0.500000
3	$1/3$	0.333333
4	$3/8$	0.375000
5	$11/30$	0.366667
6	$53/144$	0.368056
7	$103/280$	0.367857
8	$2119/5760$	0.367882
9	$16687/45630$	0.367879
10	$16481/44800$	0.367879

In summary, for the original Example 1.23, the probability that no-one gets their own coat is very close to 36.79%.

1.3 Exercises.

Exercise 1.1. Fix integers $n \geq 0$ and $t \geq 1$. Consider a randomly chosen multiset of size n with elements of t types. For each part below, calculate the probability that the multiset has the stated property, and give a brief explanation.

(a) Every type of element occurs at most once.
(b) Every type of element occurs at least once.
(c) Every type of element occurs an even number of times.
(d) Every type of element occurs an odd number of times.
(e) For $k \in \mathbb{N}$, exactly k types of element occur with multiplicity at least one.
(f) For $k \in \mathbb{N}$, exactly k types of element occur with multiplicity at least two.

Exercise 1.2. Consider rolling six fair 6-sided dice, which are distinguishable, so that there are $6^6 = 46656$ equally likely outcomes. Count how many outcomes are of each of the following types. (The answers add up to 46656.)

(a) Six-of-a-kind.
(b) Five-of-a-kind and a single.
(c) Four-of-a-kind and a pair.
(d) Four-of-a-kind and two singles.
(e) Two triples.
(f) A triple, a pair, and a single.
(g) A triple and three singles.
(h) Three pairs.
(i) Two pairs and two singles.
(j) One pair and four singles.
(k) Six singles.

Exercise 1.3. Let $m \geq 1$, $d \geq 2$, and $k \geq 0$ be integers. When rolling m fair dice, each of which has d sides, what is the probability of rolling exactly k pairs and $m - 2k$ singles?

Exercise 1.4.

(a) Prove that $\rightleftharpoons$ is an equivalence relation.
(b) Prove Proposition 1.11.

Exercise 1.5. Define a function $f : \mathbb{Z} \to \mathbb{N}$ as follows: for $a \in \mathbb{Z}$,

$$f(a) = \begin{cases} 2a & \text{if } a \geq 0, \\ -1 - 2a & \text{if } a < 0. \end{cases}$$

Show that $f : \mathbb{Z} \to \mathbb{N}$ is a bijection as follows.

(a) Define a function $g : \mathbb{N} \to \mathbb{Z}$.
(b) Show that for all $a \in \mathbb{Z}$, $g(f(a)) = a$.
(c) Show that for all $b \in \mathbb{N}$, $f(g(b)) = b$.

Exercise 1.6. Complete the proof in Example 1.13.

Exercise 1.7. Give bijective proofs of the following identities.

(a) For all $n \in \mathbb{N}$, $\sum_{k=0}^{n} \binom{n}{k} k = n2^{n-1}$.
(b) For all $n \in \mathbb{N}$, $\sum_{k=0}^{n} \binom{n}{k} k(k-1) = n(n-1)2^{N-2}$.

Exercise 1.8. For an integer $n \geq 1$, give a bijective proof that

$$\sum_{\text{even } k} \binom{n}{k} = \sum_{\text{odd } k} \binom{n}{k}$$

Exercise 1.9. Let n be a positive integer. Let S_n be the set of all ordered pairs of sets (A, B) in which $A \subseteq B \subseteq \{1, 2, ..., n\}$. Let T_n be the set of all functions $f : \{1, 2, ..., n\} \to \{1, 2, 3\}$.

(a) What is $|T_n|$? (Explain.)
(b) Define a bijection $g : S_n \to T_n$. Explain why $g((A, B)) \in T_n$ for any $(A, B) \in S_n$. (You do not need to explain why g is a bijection.)
(c) Define the inverse function $g^{-1} : T_n \to S_n$ of your bijection g from part (b). (You do not need to explain why g and g^{-1} are mutually inverse bijections.)
(d) By counting S_n and T_n in two different ways, deduce that

$$3^n = \sum_{k=0}^{n} \binom{n}{k} 2^k.$$

Exercise 1.10. Fix integers $n \geq 0$ and $k \geq 1$. Let $\mathcal{A}(n, k)$ be the set of sequences $(a_1, ..., a_k) \in \mathbb{N}^k$ such that $a_1 + \cdots + a_k = n$, and j divides a_j for all $1 \leq j \leq k$. Let $\mathcal{B}(n, k)$ be the set of sequences $(b_1, ..., b_k) \in \mathbb{N}^k$ such that $b_1 + \cdots + b_k = n$, and $b_1 \geq b_2 \geq \cdots \geq b_k$. For example, here are the sets for $n = 7$ and $k = 3$:

$\mathcal{A}(7,3)$	$\mathcal{B}(7,3)$
$(7,0,0)$	$(7,0,0)$
$(5,2,0)$	$(6,1,0)$
$(4,0,3)$	$(5,2,0)$
$(3,4,0)$	$(4,3,0)$
$(2,2,3)$	$(5,1,1)$
$(1,6,0)$	$(4,2,1)$
$(1,0,6)$	$(3,3,1)$
$(0,4,3)$	$(3,2,2)$

Construct a pair of mutually inverse bijections between the sets $\mathcal{A}(n,k)$ and $\mathcal{B}(n,k)$.

Exercise 1.11. For integers $n \geq 0$ and $t \geq 2$, give a bijective proof that

$$\binom{n+t-1}{t-1} = \sum_{k=0}^{n} \binom{n-k+t-2}{t-2}.$$

Exercise 1.12. For integers $n \geq 1$ and $t \geq 1$, give a bijective proof that

$$\binom{n+t-1}{t-1} = \sum_{k=0}^{t} \binom{t}{k}\binom{n-1}{k-1}.$$

Exercise 1.13. Choose a permutation σ of $\{1, 2, ..., 7\}$ at random, so that each of the $7! = 5040$ permutations are equally likely. What are the probabilities of the following events?

1. Numbers 1 and 2 are consecutive (*i.e.* 3672154 or 3671254).
2. Number 1 is to the left of 2.
3. No two odd numbers are consecutive.
4. Any other condition you can think of.
5. Any similar questions for permutations of $\{1, 2, ..., n\}$.

Exercise 1.14. (The $r = 13$ and $s = 4$ case of this exercise completes the table in Subsection 1.2.4.) Let $r \geq 2$ and $s \geq 2$ be integers. Consider a (non-standard) deck of rs cards, divided into s suits each with cards of r different values. The cards in each suit are numbered A, 2, 3, ..., r, and A can be either below 2 or above r. Choose five cards from such a deck in one of $\binom{rs}{5}$ ways. How many ways are there to produce each kind of hand for this "poker in an alternate universe"?

(a) Count "quints" (five-of-a-kinds).
(b) Count straight flushes.
(c) Count quads.
(d) Count full houses.
(e) Count flushes.
(f) Count straights.
(g) Count trips.
(h) Count two-pairs.
(i) Count one-pairs.
(j) Count busted hands.

Exercise 1.15. The game called "Crowns and Anchors" or "Birdcage" was popular on circus midways early in the 20th century. It is a game between a Player and the House, played as follows. First, the Player wagers w dollars on an integer p from one to six. Next, the House rolls three six-sided dice. For every die that shows p dots on top, the House pays the Player w dollars, but if no dice show p dots on top then the Player's wager is forfeited, and goes to the House. (Assume that the dice are fair, so that every outcome is equally likely.)

For example, if I wager two dollars on the number five, and the dice show five, five, and three dots, respectively, then the House pays me four dollars for a total of six (a profit of four dollars). However, if in this case the dice show four, three, and two dots, respectively, then the House takes my wager for a total of zero (a loss of two dollars).

(a) For every dollar that the Player wagers, how much money should the Player expect to win back in the long run? Would you play this game?

(b) In a parallel universe there is a game of Crowns and Anchors being played with $m \geq 1$ dice, each of which has $d \geq 2$ sides. (Assume that the dice are fair, so that every outcome is equally likely.) In which universes does the Player win in the long run? In which universes does the House win in the long run? In which universes is the game completely fair?

Chapter 2

The Idea of Generating Series.

We will be dealing algebraically with infinite power series $G(x) = \sum_{n=0}^{\infty} g_n x^n$ in which the coefficients $\mathbf{g} = (g_0, g_1, g_2, ...)$ form a sequence of integers. These can, for the most part, be handled just like polynomials. Problems of convergence can arise if one tries to substitute a particular real or complex value for the indeterminate x. We usually don't do this, and finiteness of all the coefficients of $G(x)$ is all that we require.

Example 2.1 (The Geometric Series)**.** The simplest infinite case of power series is if all the coefficients equal one. Then

$$G = 1 + x + x^2 + x^3 + x^4 + \cdots .$$

Multiply this by x:

$$xG = x + x^2 + x^3 + x^4 + \cdots .$$

It follows that $G - xG = (1 - x)G = 1$. In conclusion

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots .$$

Don't worry about convergence – this is just algebra!

2.1 The Binomial Theorem and Binomial Series.

We develop two of the most useful facts that we will need in what follows. The proofs are also good illustrations of calculating with generating series.

Theorem 2.2 (The Binomial Theorem). *For any natural number $n \in \mathbb{N}$,*

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

This formula is an identity between two polynomials in the variable x. You probably have seen a proof of it by induction on n, but we are going to prove it here using the bijection between subsets and indicator vectors discussed in Example 1.12.

Proof. Recall that $\mathcal{P}(n)$ is the set of all subsets of $\{1, 2, ..., n\}$, and that $\{0, 1\}^n$ is the set of all indicator vectors $\alpha = (a_1, a_2, ..., a_n)$ in which each coordinate is either 0 or 1. Example 1.12 gives a bijection between these two sets, which you should recall. For example, when $n = 8$ the subset $\{2, 3, 5, 7\}$ corresponds to the indicator vector $(0, 1, 1, 0, 1, 0, 1, 0)$. The constructions $S \mapsto \alpha(S)$ and $\alpha \mapsto S(\alpha)$ are mutually inverse bijections between the sets $\mathcal{P}(n)$ and $\{0, 1\}^n$. From this, we concluded that $|\mathcal{P}(n)| = |\{0, 1\}^n| = 2^n$, but we can deduce more. Notice that if S is a subset with k elements then it corresponds to an indicator vector α that sums to k. It is sometimes helpful to record this information in a little table, like this:

$$\begin{aligned} \mathcal{P}(n) &\rightleftharpoons \{0,1\}^n \\ S &\leftrightarrow \alpha = (a_1, a_2, ..., a_n) \\ |S| &= a_1 + a_2 + \cdots + a_n. \end{aligned}$$

Because of this bijection, if we introduce an "indeterminate" x, and if S corresponds to α, then

$$x^{|S|} = x^{a_1+a_2+\cdots+a_n}.$$

Moreover, also because of this bijection, summing over all subsets is equivalent to summing over all indicator functions. That is,

$$\sum_{S \in \mathcal{P}(n)} x^{|S|} = \sum_{\alpha \in \{0,1\}^n} x^{a_1+a_2+\cdots+a_n}.$$

Now we can simplify both sides separately. On the LHS, we know from Theorem 1.5 that there are $\binom{n}{k}$ k-element subsets of an n-element set, for each $0 \leq k \leq n$. Therefore,

$$\sum_{S \in \mathcal{P}(n)} x^{|S|} = \sum_{k=0}^{n} \binom{n}{k} x^k.$$

On the RHS, summing over all the indicator vectors $\alpha \in \{0,1\}^n$ is equivalent to summing over all $a_1 \in \{0,1\}$ and all $a_2 \in \{0,1\}$ and so on,... until all $a_n \in \{0,1\}$. This gives

$$\sum_{\alpha \in \{0,1\}^n} x^{a_1+a_2+\cdots+a_n} = \sum_{a_1=0}^{1} \sum_{a_2=0}^{1} \cdots \sum_{a_n=0}^{1} x^{a_1+a_2+\cdots+a_n}$$
$$= \sum_{a_1=0}^{1} x^{a_1} \sum_{a_2=0}^{1} x^{a_2} \cdots \sum_{a_n=0}^{1} x^{a_n} = \left(\sum_{a=0}^{1} x^a \right)^n = (1+x)^n.$$

This proves the Binomial Theorem. With practice and familiarity, it becomes a one-line proof:

$$\sum_{k=0}^{n} \binom{n}{k} x^k = \sum_{S \in \mathcal{P}(n)} x^{|S|} = \sum_{\alpha \in \{0,1\}^n} x^{a_1+a_2+\cdots+a_n} = (1+x)^n.$$

$\square$

Example 2.3 (Vandermonde Convolution.)**.** As mentioned in Subsection 1.2.1, the Binomial Theorem easily implies the Vandermonde Convolution formula. To see this, begin with the obvious identity of polynomials

$$(1+x)^{m+n} = (1+x)^m \cdot (1+x)^n$$

and use the Binomial Theorem to expand each of the factors.

$$\sum_{k=0}^{m+n} \binom{m+n}{k} x^k = \left(\sum_{j=0}^{m} \binom{m}{j} x^j \right) \left(\sum_{i=0}^{n} \binom{n}{i} x^i \right)$$
$$= \sum_{j=0}^{m} \sum_{i=0}^{n} \binom{m}{j} \binom{n}{i} x^{j+i} = \sum_{k=0}^{m+n} \sum_{j=0}^{k} \binom{m}{j} \binom{n}{k-j} x^k.$$

(The last step is accomplished by re-indexing the double summation.) Since the polynomials on the LHS and on the RHS are equal, they must have the same coefficients. By comparing the coefficients of x^k on both sides we see that

$$\binom{m+n}{k} = \sum_{j=0}^{k} \binom{m}{j}\binom{n}{k-j},$$

giving the result.

Consider the set $\mathcal{M}(t)$ of all multisets with $t \geq 1$ types of elements, regardless of the size of the multiset. That is, an element of $\mathcal{M}(t)$ is a sequence $\mu = (m_1, m_2, ..., m_t)$ of t natural numbers, and the size of the multiset is $|\mu| = m_1 + m_2 + \cdots + m_t$. By Theorem 1.9, for each $n \in \mathbb{N}$ there are $\binom{n+t-1}{t-1}$ elements of $\mathcal{M}(t)$ of size n. By analogy with the Binomial Theorem 2.2, we could collect these numbers as the coefficients of a power series:

$$\sum_{n=0}^{\infty} \binom{n+t-1}{t-1} x^n.$$

The Binomial Series is an algebraic formula for this summation.

Theorem 2.4 (The Binomial Series)**.** *For any positive integer $t \geq 1$,*

$$\frac{1}{(1-x)^t} = \sum_{n=0}^{\infty} \binom{n+t-1}{t-1} x^n.$$

(Properly, this is the binomial series with negative integer exponent. The general binomial series is discussed in Subsection 4.4.1).

Proof. The key observation is that the set of all multisets with $t \geq 1$ types of elements is $\mathcal{M}(t) = \mathbb{N}^t$, the Cartesian product of t copies of the natural numbers $\mathbb{N}$. This leads to a calculation similar to the proof of the Binomial Theorem above, based on this structure:

$$\begin{aligned} \mathcal{M}(t) &= \mathbb{N}^t \\ \mu &= (m_1, .., m_t) \\ |\mu| &= m_1 + \cdots + m_t \end{aligned}$$

We use this to calculate as follows. The first equality is by Theorem 1.9.

$$\sum_{n=0}^{\infty} \binom{n+t-1}{t-1} x^n = \sum_{n=0}^{\infty} |\mathcal{M}(n,t)|\, x^n = \sum_{\mu \in \mathcal{M}(t)} x^{|\mu|}$$
$$= \sum_{(m_1,m_2,\ldots,m_t)\in\mathbb{N}^t} x^{m_1+m_2+\cdots m_t} = \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\cdots\sum_{m_t=0}^{\infty} x^{m_1+m_2+\cdots+m_t}$$
$$= \sum_{m_1=0}^{\infty} x^{m_1} \sum_{m_2=0}^{\infty} x^{m_2} \cdots \sum_{m_t=0}^{\infty} x^{m_t} = \left(\sum_{m=0}^{\infty} x^m\right)^t = \frac{1}{(1-x)^t}.$$

The last equality is by the geometric series $1 + x + x^2 + \cdots = 1/(1-x)$. □

2.2 The Theory in General.

In general terms we have a sequence of numbers $\mathsf{g} = (g_0, g_1, g_2, \ldots)$ which we would like to determine. To do this we introduce an *indeterminate* x and encode these numbers as the coefficients of a power series

$$G(x) \;=\; g_0 + g_1 x + g_2 x^2 + g_3 x^3 + \cdots \;=\; \sum_{n=0}^{\infty} g_n x^n,$$

called the *generating series* for the sequence g.

By "indeterminate" we mean that x behaves algebraically as if it were a number, but that it does not have any particular value. It should be thought of as a punctuation mark that is there to keep the different coefficients g_n separated from each other. Sometimes, people call x a "variable" – but that word is meant to convey the idea that x is a number, but that we don't know specifically what the value of that number is. The word "indeterminate" is meant to convey the idea that x *does not have any value at all* – it is just a punctuation mark. As will be seen, it is the coefficients of these power series that carry information about our counting problems.

In this chapter and the next we will see how to use this strategy to encode the answers to various counting problems as generating series. In Chapter 4 we will see how to get numbers out of these power series in order to answer the counting problems explicitly.

2.2.1 Generating series.

Let $\mathcal{A}$ be a set of "objects" which we want to count. For example, $\mathcal{A}$ might be the set of subsets of the set $\{1, 2, ..., n\}$. Or, $\mathcal{A}$ might be the set of all multisets with t types of elements. The set $\mathcal{A}$ can be quite arbitrary, but we assume that each element of $\mathcal{A}$ has a "size" or "weight" attached to it. The weight of $\alpha \in \mathcal{A}$ is a nonnegative integer $\omega(\alpha) \in \mathbb{N}$. We just require that there are only finitely many elements of $\mathcal{A}$ of any given weight.

Definition 2.5 (Weight Function)**.** Let $\mathcal{A}$ be a set. A function $\omega : \mathcal{A} \to \mathbb{N}$ from $\mathcal{A}$ to the set $\mathbb{N}$ of natural numbers is a *weight function* provided that for all $n \in \mathbb{N}$, the set

$$\mathcal{A}_n = \omega^{-1}(n) = \{\alpha \in \mathcal{A} : \ \omega(\alpha) = n\}$$

is finite. That is, for every $n \in \mathbb{N}$ there are only finitely many elements $\alpha \in \mathcal{A}$ of weight n.

Notice that if $\mathcal{A}$ is a set with a weight function $\omega : \mathcal{A} \to \mathbb{N}$, then

$$\mathcal{A} = \bigcup_{n=0}^{\infty} \mathcal{A}_n$$

is a (disjoint) union of countably many finite sets, and so $\mathcal{A}$ is itself either finite or countably infinite.

Definition 2.6 (Generating series)**.** Let $\mathcal{A}$ be a set with a weight function $\omega : \mathcal{A} \to \mathbb{N}$ as in Definition 2.5. The *generating series of $\mathcal{A}$ with respect to ω* is

$$A(x) = \Phi^{\omega}_{\mathcal{A}}(x) = \sum_{\alpha \in \mathcal{A}} x^{\omega(\alpha)}.$$

(We usually suppress the superscript from the notation.) Remember – the indeterminate x does not have a value. It is just used to keep track of the weight of each object $\alpha \in \mathcal{A}$ in the exponent.

Proposition 2.7. *Let $\mathcal{A}$ be a set with a weight function $\omega : \mathcal{A} \to \mathbb{N}$, and let*

$$\Phi_{\mathcal{A}}(x) = a_0 + a_1 x + a_2 x^2 + \cdots = \sum_{n=0}^{\infty} a_n x^n.$$

For every $n \in \mathbb{N}$, the number of elements of $\mathcal{A}$ of weight n is $a_n = |\mathcal{A}_n|$.

Proof.

$$\Phi_{\mathcal{A}}(x) = \sum_{\alpha \in \mathcal{A}} x^{\omega(\alpha)} = \sum_{n=0}^{\infty} \sum_{\alpha \in \mathcal{A}:\ \omega(\alpha)=n} x^{\omega(\alpha)} = \sum_{n=0}^{\infty} x^n \sum_{\alpha \in \mathcal{A}_n} 1 = \sum_{n=0}^{\infty} |\mathcal{A}_n|\, x^n.$$

Thus, for each $n \in \mathbb{N}$, the coefficient of x^n in $\Phi_{\mathcal{A}}(x)$ is the number of elements in $\mathcal{A}$ that have weight n. $\square$

Since we will be doing a lot of long calculations with power series, and because of Proposition 2.7, it is useful to have a handy notation for extracting coefficients from them.

Definition 2.8. Let $G(x) = g_0 + g_1 x + g_2 x^2 + \cdots = \sum_{n=0}^{\infty} g_n x^n$ be any power series. Then for any $k \in \mathbb{N}$,

$$[x^k]\, G(x) \;=\; g_k$$

is the coefficient of x^k in the power series $G(x)$.

Example 2.9. For example, for any natural numbers $a, b \in \mathbb{N}$,

$$[x^a] \frac{1}{(1-x)^{1+b}} = [x^a] \sum_{n=0}^{\infty} \binom{n+b}{b} x^n = \binom{a+b}{b}.$$

2.2.2 The Sum, Product, and String Lemmas.

Lemma 2.10 (The Sum Lemma.). *Let $\mathcal{A}$ and $\mathcal{B}$ be disjoint sets, so that $\mathcal{A} \cap \mathcal{B} = \varnothing$. Assume that $\omega : (\mathcal{A} \cup \mathcal{B}) \to \mathbb{N}$ is a weight function on the union of $\mathcal{A}$ and $\mathcal{B}$. We may regard ω as a weight function on each of $\mathcal{A}$ or $\mathcal{B}$ separately (by restriction). Under these conditions,*

$$\Phi_{\mathcal{A}\cup\mathcal{B}}(x) = \Phi_{\mathcal{A}}(x) + \Phi_{\mathcal{B}}(x).$$

Proof. From the definition of generating series,

$$\Phi_{\mathcal{A}\cup\mathcal{B}}(x) = \sum_{\sigma\in\mathcal{A}\cup\mathcal{B}} x^{\omega(\sigma)} = \sum_{\sigma\in\mathcal{A}} x^{\omega(\sigma)} + \sum_{\sigma\in\mathcal{B}} x^{\omega(\sigma)} = \Phi_{\mathcal{A}}(x) + \Phi_{\mathcal{B}}(x).$$

(The condition that $\mathcal{A} \cap \mathcal{B} = \varnothing$ is needed for the second equality.) □

In fact, the above proof can be generalized slightly to give more.

Lemma 2.11 (The Infinite Sum Lemma.). *Let $\mathcal{A}_0$, $\mathcal{A}_1$, $\mathcal{A}_2$,... be pairwise disjoint sets (so that $\mathcal{A}_i \cap \mathcal{A}_j = \varnothing$ if $i \neq j$), and let $\mathcal{B} = \bigcup_{j=0}^{\infty} \mathcal{A}_j$. Assume that $\omega : \mathcal{B} \to \mathbb{N}$ is a weight function. We may regard ω as a weight function on each of the sets $\mathcal{A}_j$ separately (by restriction). Under these conditions,*

$$\Phi_{\mathcal{B}}(x) = \sum_{j=0}^{\infty} \Phi_{\mathcal{A}_j}(x).$$

Proof. Exercise 2.6. □

Lemma 2.12 (The Product Lemma.). *Let $\mathcal{A}$ and $\mathcal{B}$ be sets with weight functions $\omega : \mathcal{A} \to \mathbb{N}$ and $\nu : \mathcal{B} \to \mathbb{N}$, respectively. Define $\eta : \mathcal{A} \times \mathcal{B} \to \mathbb{N}$ by putting $\eta(\alpha, \beta) = \omega(\alpha) + \nu(\beta)$ for all $(\alpha, \beta) \in \mathcal{A} \times \mathcal{B}$. Then η is a weight function on $\mathcal{A} \times \mathcal{B}$, and*

$$\Phi^{\eta}_{\mathcal{A}\times\mathcal{B}}(x) = \Phi^{\omega}_{\mathcal{A}}(x) \cdot \Phi^{\nu}_{\mathcal{B}}(x).$$

(The superscripts ω, ν, and η indicate which weight function is being used for each set.)

Proof. To see that η is a weight function, consider any $n \in \mathbb{N}$. There are $n+1$ choices for an integer $0 \leq k \leq n$. For each such k, there are only finitely many pairs $(\alpha, \beta) \in \mathcal{A} \times \mathcal{B}$ with $\omega(\alpha) = k$ and $\omega(\beta) = n - k$. That is, the set of elements of $\mathcal{A} \times \mathcal{B}$ of weight n is

$$(\mathcal{A} \times \mathcal{B})_n = \bigcup_{k=0}^{n} \mathcal{A}_k \times \mathcal{B}_{n-k},$$

a finite (disjoint) union of finite sets. It follows that there are only finitely many elements of $\mathcal{A} \times \mathcal{B}$ of weight n. Now,

$$\begin{aligned} \Phi^{\eta}_{\mathcal{A}\times\mathcal{B}}(x) &= \sum_{(\alpha,\beta)\in\mathcal{A}\times\mathcal{B}} x^{\eta(\alpha,\beta)} = \sum_{\alpha\in\mathcal{A}} \sum_{\beta\in\mathcal{B}} x^{\omega(\alpha)+\nu(\beta)} \\ &= \sum_{\alpha\in\mathcal{A}} x^{\omega(\alpha)} \cdot \sum_{\beta\in\mathcal{B}} x^{\nu(\beta)} = \Phi_{\mathcal{A}}(x) \cdot \Phi_{\mathcal{B}}(x). \end{aligned}$$

□

The Product Lemma 2.12 can be extended to the Cartesian product of any finite number of sets, by induction on the number of factors. (Exercise 2.7.)

Finally, the String Lemma combines both disjoint union and Cartesian product, as follows. Let $\mathcal{A}$ be a set with a weight function $\omega : \mathcal{A} \to \mathbb{N}$. For any $k \in \mathbb{N}$, the Cartesian product of k copies of $\mathcal{A}$ is denoted by $\mathcal{A}^k$. The entries of $\mathcal{A}^k$ are k-tuples $(\alpha_1, \alpha_2, ..., \alpha_k)$ with each $\alpha_i \in \mathcal{A}$. Notice that $\mathcal{A}^0 = \{\varepsilon\}$ is the one-element set whose only element is the empty string $\varepsilon = ()$ of length zero. We can define a weight function ω_k on $\mathcal{A}^k$ by putting

$$\omega_k(\alpha_1, ..., \alpha_k) = \omega(\alpha_1) + \cdots + \omega(\alpha_k).$$

It is a good exercise to check that this is a weight function. Note that the weight of the empty string is zero. Repeated application of the Product Lemma 2.12 shows that for all $k \in \mathbb{N}$,

$$\Phi_{\mathcal{A}^k}(x) = \left(\Phi_{\mathcal{A}}(x)\right)^k.$$

We can take the union of the sets $\mathcal{A}^k$ for all $k \in \mathbb{N}$:

$$\mathcal{A}^* = \bigcup_{k=0}^{\infty} \mathcal{A}^k.$$

Notice that the sets in this union are pairwise disjoint, since each $\mathcal{A}^k$ consists of strings with exactly k coordinates. We define a function $\omega^* : \mathcal{A}^* \to \mathbb{N}$ by saying that $\omega^* = \omega_k$ when restricted to $\mathcal{A}^k$.

Lemma 2.13. *Let $\mathcal{A}$ be a set with weight function $\omega : \mathcal{A} \to \mathbb{N}$, and define $\mathcal{A}^*$ and $\omega^* : \mathcal{A}^* \to \mathbb{N}$ as above. Then ω^* is a weight function on $\mathcal{A}^*$ if and only if there are no elements in $\mathcal{A}$ of weight zero (that is, $\mathcal{A}_0 = \varnothing$).*

Proof. If $\gamma \in \mathcal{A}$ has weight zero, $\omega(\gamma) = 0$, then for any natural number $k \in \mathbb{N}$, a sequence of k γ-s in $\mathcal{A}^k$ also has weight zero: $\omega_k(\gamma, \gamma, ..., \gamma) = 0$. So, by the way $\omega^* : \mathcal{A}^* \to \mathbb{N}$ is defined, there are infinitely many elements of weight zero in $\mathcal{A}^*$, so that ω^* is not a weight function.

Conversely, assume that every element of $\mathcal{A}$ has weight at least 1. Then, for each $k \in \mathbb{N}$, every element of $\mathcal{A}^k$ has weight at least k. Now consider any $n \in \mathbb{N}$ and all the strings $(\alpha_1, ..., \alpha_k) \in \mathcal{A}^*$ of weight n. By the previous sentence, if there are any such strings of length k then $0 \leq k \leq n$. For each $0 \leq k \leq n$, $\mathcal{A}^k$ has only finitely many elements of weight n. It follows that $\mathcal{A}^*$ has only finitely many elements of weight n. Therefore, ω^* is a weight function on $\mathcal{A}^*$. ∎

Lemma 2.14 (The String Lemma.)**.** *Let $\mathcal{A}$ be a set with a weight function $\omega : \mathcal{A} \to \mathbb{N}$ such that there are no elements of $\mathcal{A}$ of weight zero. Then*

$$\Phi_{\mathcal{A}^*}(x) = \frac{1}{1 - \Phi_{\mathcal{A}}(x)}.$$

Proof. By the Infinite Sum and Product Lemmas 2.11 and 2.12,

$$\Phi_{\mathcal{A}^*}(x) = \sum_{k=0}^{\infty} \Phi_{\mathcal{A}^k}(x) = \sum_{k=0}^{\infty} \left(\Phi_{\mathcal{A}}(x)\right)^k = \frac{1}{1 - \Phi_{\mathcal{A}}(x)}.$$

∎

2.3 Compositions.

Definition 2.15. A *composition* is a finite sequence of positive integers:

$$\gamma = (c_1, c_2, ..., c_k),$$

in which $k \in \mathbb{N}$ is a natural number, and each $c_i \geq 1$ is a positive integer. The entries c_i are called the *parts* of the composition. The *length* of the composition is $\ell(\gamma) = k$, the number of parts. The *size* of the composition is

$$|\gamma| = c_1 + c_2 + \cdots + c_k,$$

the sum of the parts.

Notice that there is exactly one composition of length zero: this is $\varepsilon = ()$, the empty string with no entries. Compositions are related to multisets, but there are two important differences: the parts of a composition must be positive integers, not just nonnegative, and the length of a composition might not be specified, while the number of types of element in a multiset must be fixed.

Example 2.16. Here are all the compositions of size five:

(5)	(2,3)	(2,2,1)	(1,2,1,1)
(4,1)	(3,1,1)	(2,1,2)	(1,1,2,1)
(1,4)	(1,3,1)	(1,2,2)	(1,1,1,2)
(3,2)	(1,1,3)	(2,1,1,1)	(1,1,1,1,1)

In this section we will apply the results of Subsection 2.2.2 to obtain formulas for the generating series of various sets of compositions defined by imposing some extra conditions. In Chapter 4 we will see how to use this information to actually count such things.

Theorem 2.17. *Let $P = \{1, 2, 3, ...\}$ be the set of positive integers.*

(a) The set $\mathcal{C}$ of all compositions is $\mathcal{C} = P^*$.

(b) The generating series for $\mathcal{C}$ with respect to size is

$$\Phi_{\mathcal{C}}(x) = 1 + \frac{x}{1-2x}.$$

(c) For each $n \in \mathbb{N}$, the number of compositions of size n is

$$|\mathcal{C}_n| = \begin{cases} 1 & \text{if } n = 0, \\ 2^{n-1} & \text{if } n \geq 1. \end{cases}$$

Proof. A single part is a positive integer $c \in P = \{1, 2, 3, ...\}$. A composition of length k is a sequence $\gamma = (c_1, c_2, ..., c_k)$ of k positive integers, so is an element of P^k. Since the length k can be any natural number, the set $\mathcal{C}$ of all compositions is

$$\mathcal{C} = \bigcup_{k=0}^{\infty} P^k = P^*.$$

This proves part (a).

The generating series for one-part compositions with respect to size is

$$\Phi_P(x) \;=\; \sum_{c=1}^{\infty} x^c = x + x^2 + x^3 + \cdots \;=\; \frac{x}{1-x},$$

by the geometric series. From the String Lemma 2.14 it follows that

$$\Phi_{\mathcal{C}}(x) \;=\; \Phi_{P^*}(x) \;=\; \frac{1}{1 - x/(1-x)} \;=\; \frac{1-x}{1-2x} \;=\; 1 + \frac{x}{1-2x}.$$

This proves part (b).

Expanding $C(x) = \Phi_{\mathcal{C}}(x)$ using the geometric series, we obtain

$$C(x) \;=\; 1 + \sum_{j=0}^{\infty} 2^j\, x^{j+1} \;=\; 1 + \sum_{n=1}^{\infty} 2^{n-1}\, x^n.$$

Since $|\mathcal{C}_n| = [x^n]C(x)$ is the coefficient of x^n in $C(x)$, this proves part (c). □

Proposition 2.23 is a bijection which gives a combinatorial proof of Theorem 2.17(c).

Many variations on the proof of Theorem 2.17 are possible. We will do a few as examples, and present many more as exercises. The general approach consists of three steps:

- Identify the allowed values for each part. This might depend on the position of the part within the composition.
- Identify the allowed lengths for the compositions.
- Apply the Sum, Product, and String Lemmas to obtain a formula for the generating series.

Example 2.18. Let $\mathcal{F}$ be the set of all compositions in which each part is either one or two.

- The allowed sizes for a part are 1 or 2, so $P = \{1, 2\}$ is the set of allowed parts. The generating series for a single part is $x + x^2$.
- The length can be any natural number $k \in \mathbb{N}$. By the Product Lemma, the generating series for compositions in $\mathcal{F}$ of length k is $(x + x^2)^k$.
- Since $\mathcal{F} = \{1, 2\}^*$, the String Lemma implies that

$$F(x) \;=\; \Phi_{\mathcal{F}}(x) \;=\; \sum_{n=0}^{\infty} f_n\, x^n = \sum_{k=0}^{\infty} (x + x^2)^k = \frac{1}{1 - x - x^2}.$$

  Here, $f_n = [x^n]F(x) = |\mathcal{F}_n|$ is the number of compositions in $\mathcal{F}$ of size n. In Section 4.1 we will see how to use this information to get a formula for the numbers f_n.

Example 2.19. Let $\mathcal{H}$ be the set of all compositions in which each part is at least two.

- The allowed sizes for a part are $P = \{2, 3, 4, ...\}$. The generating series for a single part is

$$\Phi_P(x) \;=\; \sum_{c=2}^{\infty} x^c \;=\; x^2 + x^3 + x^4 + \cdots \;=\; \frac{x^2}{1 - x}.$$

- The length can be any natural number $k \in \mathbb{N}$. By the Product Lemma, the generating series for compositions in $\mathcal{H}$ of length k

is

$$\left(\frac{x^2}{1-x}\right)^k.$$

- Since $\mathcal{H} = P^*$, the String Lemma implies that

$$\begin{aligned} H(x) &= \Phi_{\mathcal{H}}(x) = \sum_{n=0}^{\infty} h_n\, x^n = \sum_{k=0}^{\infty} \left(\frac{x^2}{1-x}\right)^k \\ &= \frac{1}{1-x^2/(1-x)} = \frac{1-x}{1-x-x^2} = 1 + \frac{x^2}{1-x-x^2}. \end{aligned}$$

Here, $h_n = [x^n]H(x)$ is the number of compositions in $\mathcal{H}$ of size n.

Example 2.20. Let $\mathcal{J}$ be the set of all compositions in which each part is odd.

- The allowed sizes for a part are $P = \{1, 3, 5, ...\}$. The generating series for a single part is

$$\Phi_P(x) = \sum_{i=0}^{\infty} x^{2i+1} = x^1 + x^3 + x^5 + \cdots = \frac{x}{1-x^2}.$$

- The length can be any natural number $k \in \mathbb{N}$. By the Product Lemma, the generating series for compositions in $\mathcal{J}$ of length k is

$$\left(\frac{x}{1-x^2}\right)^k.$$

- Since $\mathcal{J} = P^*$, the String Lemma implies that

$$\begin{aligned} J(x) &= \Phi_{\mathcal{J}}(x) = \sum_{n=0}^{\infty} j_n\, x^n = \frac{1}{1-x/(1-x^2)} \\ &= \frac{1-x^2}{1-x-x^2} = 1 + \frac{x}{1-x-x^2}. \end{aligned}$$

Here, $j_n = [x^n]J(x)$ is the number of compositions in $\mathcal{J}$ of size n.

Example 2.21. The sets $\mathcal{F}$, $\mathcal{H}$, and $\mathcal{J}$ in Examples 2.18 to 2.20 have very similar generating series. In fact, after a little thought one sees that for all $n \geq 2$,

$$[x^n]\, H(x) \;=\; [x^{n-1}]\, J(x) \;=\; [x^{n-2}]\, F(x) \;=\; [x^{n-2}]\, \frac{1}{1-x-x^2}.$$

This means that for all $n \geq 2$, we have $h_n = j_{n-1} = f_{n-2}$, so for the sizes of sets we have $|\mathcal{H}_n| = |\mathcal{J}_{n-1}| = |\mathcal{F}_{n-2}|$. We have proven these equalities even though we don't yet know what those numbers actually are! This seems slightly magical, but it works.

Since these sets have the same sizes there must be bijections between them to explain this fact. Constructing such bijections is left to Exercise 2.17. As a starting point, here are the sets for $n = 7$:

$\mathcal{H}_7$	$\mathcal{J}_6$	$\mathcal{F}_5$
(7)	$(5,1)$	$(2,2,1)$
$(5,2)$	$(1,5)$	$(2,1,2)$
$(2,5)$	$(3,3)$	$(1,2,2)$
$(4,3)$	$(3,1,1,1)$	$(2,1,1,1)$
$(3,4)$	$(1,3,1,1)$	$(1,2,1,1)$
$(3,2,2)$	$(1,1,3,1)$	$(1,1,2,1)$
$(2,3,2)$	$(1,1,1,3)$	$(1,1,1,2)$
$(2,2,3)$	$(1,1,1,1,1,1)$	$(1,1,1,1,1)$

(It need not be the case that the bijections match up these sets of compositions line by line in this table.) In Section 4.1 we will determine the coefficients of the power series $1/(1-x-x^2)$, answering the counting problem for these sets $\mathcal{F}$, $\mathcal{H}$, and $\mathcal{J}$.

Example 2.22. Let $\mathcal{Q}$ be the set of all compositions in which each part is at least two, and the number of parts is even.

- The allowed sizes for a part are $P = \{2, 3, 4, ...\}$. The generating series for a single part is

$$\Phi_P(x) \;=\; \sum_{c=2}^{\infty} x^c = x^2 + x^3 + x^4 + \cdots \;=\; \frac{x^2}{1-x}.$$

- The length is an even natural number $k = 2j$ for some $j \in \mathbb{N}$. By the Product Lemma, the generating series for a composition in $\mathcal{Q}$

of length $2j$ is

$$\left(\frac{x^2}{1-x}\right)^{2j}.$$

- Since $\mathcal{Q} = (P^2)^*$, the String Lemma implies that

$$\begin{aligned} Q(x) \; &= \; \Phi_{\mathcal{Q}}(x) = \sum_{n=0}^{\infty} q_n\, x^n \; = \; \sum_{j=0}^{\infty} \left(\frac{x^2}{1-x}\right)^{2j} \\ &= \frac{1}{1 - x^4/(1-x)^2} \; = \; \frac{(1-x)^2}{(1-x)^2 - x^4} \\ &= \frac{1-2x+x^2}{1-2x+x^2-x^4} \; = \; 1 + \frac{x^4}{1-2x+x^2-x^4}. \end{aligned}$$

Here $q_n = [x^n]Q(x) = |\mathcal{Q}_n|$ is the number of compositions in $\mathcal{Q}$ of size n. In Example 4.11 we will see how to calculate the first several values of $|\mathcal{Q}_n|$.

2.4 Subsets with Restrictions.

The theory above for compositions can be used to obtain generating series for subsets of natural numbers subject to some restrictions on the "gaps" between consecutive elements of the subset. This is because of the following correspondence between such subsets and nonempty compositions.

Proposition 2.23. *Let $\mathcal{U}$ be the set of pairs (n, A) in which $n \in \mathbb{N}$ is a natural number and $A \subseteq \{1, 2, ..., n\}$ is a subset. Let $\mathcal{C} \smallsetminus \{\varepsilon\}$ be the set of nonempty compositions. There is a bijection $\mathcal{U} \rightleftharpoons \mathcal{C} \smallsetminus \{\varepsilon\}$ between these two sets.*

Proof. We define mutually inverse bijections $\mathcal{U} \rightleftharpoons \mathcal{C} \smallsetminus \{\varepsilon\}$ as follows.

First, given (n, A) in $\mathcal{U}$, sort the elements of $A = \{a_1, a_2, ..., a_k\}$ into increasing order:

$$1 \leq a_1 < a_2 < \cdots < a_k \leq n.$$

For convenience, put $a_0 = 0$ and $a_{k+1} = n + 1$. Now define $c_i = a_i - a_{i-1}$ for all $1 \leq i \leq k + 1$, and let $\gamma = (c_1, c_2, ..., c_{k+1})$. Notice that each c_i is a positive integer, so that γ is a composition of length $k + 1$. Since $k + 1$ is at least one, this composition γ is not empty, so that γ is in the set $\mathcal{C} \smallsetminus \{\varepsilon\}$. The size of this composition is

$$|\gamma| = \sum_{i=1}^{k+1} c_i = \sum_{i=1}^{k+1}(a_i - a_{i-1}) = a_{k+1} - a_0 = n+1,$$

and its length is $\ell(\gamma) = |A| + 1$.

Conversely, given a nonempty composition $\gamma = (c_1, c_2, ..., c_\ell)$ in $\mathcal{C} \smallsetminus \{\varepsilon\}$, notice that $\ell \geq 1$. We define $a_j = c_1 + c_2 + \cdots + c_j$ for each $1 \leq j \leq \ell - 1$, and let $A = \{a_1, a_2, ..., a_{\ell-1}\}$ and $n = |\gamma| - 1$. This defines a pair (n, A) in the set $\mathcal{U}$.

One can check that these constructions give a pair of mutually inverse bijections $\mathcal{U} \rightleftharpoons \mathcal{C} \smallsetminus \{\varepsilon\}$ by using Proposition 1.11. This is left as Exercise 2.18 □

The bijection of Proposition 2.23 can be neatly summarized in a little table, as follows:

$$\begin{aligned} \mathcal{U} &\rightleftharpoons \mathcal{C} \smallsetminus \{\varepsilon\} \\ (n, A) &\leftrightarrow \gamma \\ n &= |\gamma| - 1 \\ |A| &= \ell(\gamma) - 1. \end{aligned}$$

Many variations on this bijection are possible, by putting some restrictions on the allowed "gaps" between elements of the subset A and then analyzing the corresponding set of (nonempty) compositions. When deriving these generating series, we have to be careful that the n in the pair (n, A) corresponds to $|\gamma| - 1$ for the corresponding composition γ.

Example 2.24. For each $n \in \mathbb{N}$, let r_n be the number of subsets of $\{1, ..., n\}$ that do not contain two consecutive numbers (like a and $a+1$). We obtain a formula for the generating series $R(x) = \sum_{n=0}^{\infty} r_n\, x^n$ using the ideas of Proposition 2.23.

For $n \in \mathbb{N}$, let $\mathcal{R}_n$ be the set of pairs (n, A) with A as in the statement of the problem, and let $\mathcal{R} = \bigcup_{n=0}^{\infty} \mathcal{R}_n$. Then $|\mathcal{R}_n| = r_n$ for all $n \in \mathbb{N}$, and we want to determine the generating series for the set $\mathcal{R}$ with respect to the weight function $\omega(n, A) = n$.

The first question to ask is: which nonempty compositions correspond to pairs in the set $\mathcal{R}$? Notice that (n, A) is in $\mathcal{R}$ if and only if the corresponding composition $\gamma = (c_1, c_2, ..., c_\ell)$ of size $n+1$ has $c_i \geq 2$ for all $2 \leq i \leq \ell - 1$. It is possible that maybe $c_1 = 1$ or $c_\ell = 1$, though. Let $\mathcal{M}$ be the set of compositions corresponding to pairs in $\mathcal{R}$. Notice that

$$M(x) = \Phi_{\mathcal{M}}(x) = \sum_{\gamma \in \mathcal{M}} x^{|\gamma|} = \sum_{(n,A) \in \mathcal{R}} x^{n+1} = x\,R(x).$$

The compositions in $\mathcal{M}$ can be described as follows.

- The first and last parts are positive integers.
- Parts other than the first and last parts are integers greater than or equal to 2.
- The length is at least one (since it is one more than the size of the corresponding subset).

A part that is a positive integer has generating series $x/(1-x)$, and a part that is at least 2 has generating series $x^2/(1-x)$, as we have seen. Now we do a case analysis by the number of parts, using the Product Lemma.

- For $\ell = 1$ part the generating series is $x/(1-x)$.
- For $\ell \geq 2$ parts the generating series is

$$\frac{x}{1-x}\left(\frac{x^2}{1-x}\right)^{\ell-2}\frac{x}{1-x} = \frac{x^{2\ell-2}}{(1-x)^\ell}.$$

Combining the contributions for all lengths $\ell \geq 1$ using the Sum Lemma, we have

$$\begin{aligned} xR(x) = M(x) &= \frac{x}{1-x} + \sum_{\ell=2}^{\infty} \frac{x^{2\ell-2}}{(1-x)^\ell} \\ &= \frac{x}{1-x} + \frac{x^2}{(1-x)^2} \sum_{j=0}^{\infty} \left(\frac{x^2}{1-x}\right)^j \end{aligned}$$

$$= \frac{x}{1-x} + \frac{x^2}{(1-x)^2} \cdot \frac{1}{1 - x^2/(1-x)}$$
$$= \frac{x - x^2 - x^3}{(1-x)(1-x-x^2)} + \frac{x^2}{(1-x)(1-x-x^2)}$$
$$= \frac{x + x^2}{1 - x - x^2}.$$

It follows that
$$R(x) = \frac{1+x}{1-x-x^2}.$$

2.5 Proof of Inclusion/Exclusion.

In this section we prove Theorem 1.15, the Principle of Inclusion/Exclusion.

Lemma 2.25. *For any nonempty set T,*
$$\sum_{\varnothing \neq S \subseteq T} (-1)^{|S|-1} = 1.$$

Proof. Consider the identity
$$\sum_{S \subseteq \{1,2,\ldots,n\}} x^{|S|} \;=\; (1+x)^n$$
which was part of the proof of the Binomial Theorem. If T is any n-element set then
$$\sum_{S \subseteq T} x^{|S|} \;=\; (1+x)^n$$
as well (as can be seen by numbering the elements of T arbitrarily). Both sides are polynomials in x, so we can substitute $x = -1$. The result is
$$\sum_{S \subseteq T} (-1)^{|S|} \;=\; (1-1)^n \;=\; 0^n \;=\; 0,$$
because $n \geq 1$. (Note that $0^0 = 1$, since it is an empty product.) On the LHS we separate the term corresponding to $S = \varnothing$, and see that
$$1 + \sum_{\varnothing \neq S \subseteq T} (-1)^{|S|} = 0.$$

Rearranging this gives the desired formula. □

Recall the notation from Subsection 1.1.6: for any finite number of sets $A_1, A_2, \ldots, A_m$ and $\varnothing \neq S \subseteq \{1, 2, ..., m\}$, let

$$A_S = \bigcap_{i \in S} A_i.$$

So, for example, $A_{\{2,3,5\}} = A_2 \cap A_3 \cap A_5$.

Theorem 2.26 (Inclusion/Exclusion)**.** *Let $A_1, A_2, \ldots, A_m$ be finite sets. Then*

$$|A_1 \cup \cdots \cup A_m| = \sum_{\varnothing \neq S \subseteq \{1,2,\ldots,m\}} (-1)^{|S|-1} |A_S|.$$

Proof. Let $V = A_1 \cup \cdots \cup A_m$, and let $N_m = \{1, 2, .., m\}$. For each $v \in V$ let $T(v) = \{i \in N_m : \ v \in A_i\}$. Notice that $T(v) \neq \varnothing$, for all $v \in V$. Also notice that for $\varnothing \neq S \subseteq N_m$ we have $v \in A_S$ if and only if $\varnothing \neq S \subseteq T(v)$. Therefore, using Lemma 2.25 above, we have

$$\begin{aligned}\sum_{\varnothing \neq S \subseteq N_m} (-1)^{|S|-1} |A_S| &= \sum_{\varnothing \neq S \subseteq N_m} (-1)^{|S|-1} \sum_{v \in A_S} 1 \\ &= \sum_{v \in V} \sum_{\varnothing \neq S \subseteq T(v)} (-1)^{|S|-1} = \sum_{v \in V} 1 = |V|,\end{aligned}$$

as was to be shown. □

Example 2.27 (The Euler totient function)**.** For a positive integer n, the *Euler totient* of n is the number $\varphi(n)$ of integers b in the range $1 \leq b \leq n$ such that b and n are relatively prime. That is,

$$\varphi(n) = |\{b \in \{1, 2, ..., n\} : \ \gcd(b, n) = 1\}|.$$

We can use Inclusion/Exclusion to obtain a formula for $\varphi(n)$, as follows. Let the prime factorization of n be $n = p_1^{c_1} p_2^{c_2} \cdots p_m^{c_m}$, in which the p_i are pairwise distinct primes and the c_i are positive integers. For each $1 \leq i \leq m$, let

$$A_i := \{b \in N_n : \ p_i \text{ divides } b\}.$$

Then

$$\varphi(n) = |(N_n \smallsetminus (A_1 \cup \cdots \cup A_m))| = n - |A_1 \cup \cdots \cup A_m|.$$

Since the factors p_i are pairwise coprime, for any $\varnothing \neq S \subseteq N_m$ and $b \in N_n$ we have $b \in A_S$ if and only if $\prod_{i \in S} p_i$ divides b. Therefore,

$$|A_S| = \frac{n}{\prod_{i \in S} p_i}.$$

By Inclusion/Exclusion, it follows that

$$|A_1 \cup \cdots \cup A_m| = n \sum_{\varnothing \neq S \subseteq N_m} (-1)^{|S|-1} \prod_{i \in S} \frac{1}{p_i}.$$

Therefore

$$\begin{aligned} \varphi(n) &= n - n \sum_{\varnothing \neq S \subseteq N_m} (-1)^{|S|-1} \prod_{i \in S} \frac{1}{p_i} \\ &= n \sum_{S \subseteq N_m} (-1)^{|S|} \prod_{i \in S} \frac{1}{p_i} = n \prod_{i=1}^{m} \left(1 - \frac{1}{p_i}\right). \end{aligned}$$

2.6 Exercises.

Exercise 2.1. Calculate the following coefficients.

(a) $[x^8](1-x)^{-7}$.
(b) $[x^{10}]x^6(1-2x)^{-5}$.
(c) $[x^8](x^3+5x^4)(1+3x)^6$.
(d) $[x^9]((1-4x)^5+(1-3x)^{-2})$.
(e) $[x^n](1-2tx)^{-k}$.
(f) $[x^{n+1}]x^k(1-4x)^{-2k}$.
(g) $[x^n]x^k(1-x^2)^{-m}$.
(h) $[x^n]((1-x^2)^{-k}+(1-7x3)^{-k})$.

Exercise 2.2. In each case, find an instance of a Binomial Series that begins as shown.

(a) $1 - 2x + 3x^2 - 4x^3 + 5x^4 - 6x^5 + \cdots$.
(b) $1 + 3x + 6x^2 + 10x^3 + 15x^4 + 21x^5 + \cdots$.
(c) $1 - x^3 + x^6 - x^9 + x^{12} - x^{15} + \cdots$.
(d) $1 + 2x^2 + 4x^4 + 8x^6 + 16x^8 + 32x^{10} + \cdots$.
(e) $1 - 4x^2 + 12x^4 - 32x^6 + 80x^8 - 192x^{10} + \cdots$.
(f) $1 + 6x + 24x^2 + 80x^3 + 240x^4 + 672x^5 + \cdots$.

Exercise 2.3. Give algebraic proofs of these identities from Exercise 1.7.

(a) For all $n \in \mathbb{N}$, $\sum_{k=0}^{n} \binom{n}{k} k \ = \ n\, 2^{n-1}$.
(b) For all $n \in \mathbb{N}$, $\sum_{k=0}^{n} \binom{n}{k} k(k-1) \ = \ n(n-1)2^{n-2}$.

Exercise 2.4. Calculate $[x^n](1+x)^{-2}(1-2x)^{-2}$. Give the simplest expression you can find.

Exercise 2.5.

(a) Let $a \geq 1$ be an integer. For each $n \in \mathbb{N}$, extract the coefficient of x^n from both sides of this power series identity:

$$\frac{(1+x)^a}{(1-x^2)^a} = \frac{1}{(1-x)^a}$$

to show that

$$\binom{n+a-1}{a-1} = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{a}{n-2k}\binom{k+a-1}{a-1}.$$

(b) Can you think of a combinatorial proof?

Exercise 2.6. Prove the Infinite Sum Lemma 2.11.

Exercise 2.7. Extend the Product Lemma 2.12 to the product of finitely many sets with weight functions.

Exercise 2.8. Show that for $m, n, k \in \mathbb{N}$,

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

Exercise 2.9.

(a) Make a list of all the four-letter "words" that can be formed from the "alphabet" $\{a, b\}$. Define the weight of a word to be the number of occurrences of ab in it. Determine how many words there are of weight 0, 1 and 2. Determine the generating series.

(b) Do the same for five-letter words over the same alphabet, but, preferably, without listing all the words separately.

(c) Do the same for six-letter words.

Exercise 2.10.

(a) Consider throwing two six–sided dice, one red and one green. The weight of a throw is the total number of pips showing on the top faces of both dice (that is, the usual score). Make a table showing the number of throws of each weight, and write down the generating series.

(b) Do the same as for part (a), but throwing three dice: one red, one green, and one white.

Exercise 2.11. Construct a table, as in Exercise 2.10(a), if the weight of a throw is defined to be the absolute value of the difference between the numberof pips showing on the two dice. Also, write down the generating series.

Exercise 2.12. Let $\mathcal{S}$ be the set of ordered pairs (a, b) of integers with $0 \leq |b| \leq a$. Each part gives a function ω defined on the set $\mathcal{S}$. Determine whether or not ω is a weight function on the set $\mathcal{S}$. If it is not, then explain why not. If it is a weight function, then determine the generating series $\Phi_{\mathcal{S}}(x)$ of $\mathcal{S}$ with respect to ω, and write it as a polynomial or a quotient of polynomials.

(a) For (a, b) in $\mathcal{S}$, let $\omega((a, b)) = a$.
(b) For (a, b) in $\mathcal{S}$, let $\omega((a, b)) = a + b$.
(c) For (a, b) in $\mathcal{S}$, let $\omega((a, b)) = 2a + b$.

Exercise 2.13. Let $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}^4$ be the set of outcomes when rolling four six-sided dice. For $(a, b, c, d) \in \mathcal{S}$, define its weight to be $\omega(a, b, c, d) = a + b + c + d$. Consider the generating series $\Phi_{\mathcal{S}}(x)$ of $\mathcal{S}$ with respect to ω.

(a) Explain why $\Phi_{\mathcal{S}}(x) = \left(\dfrac{x - x^7}{1 - x}\right)^4$.
(b) How many outcomes in $\mathcal{S}$ have weight 19?
(c) Let m, d, k be positive integers. When rolling m dice, each of which has exactly d sides (numbered with $1, 2, ..., d$ pips, respectively), how many different ways are there to roll a total of k pips on the top faces of the dice? (Part (b) is the case $m = 4$, $d = 6$, $k = 19$.)

Exercise 2.14. Let $\mathcal{A}$ be a set with weight function $\omega : \mathcal{A} \to \mathbb{N}$. Show that for any $n \in \mathbb{N}$, the number of $\alpha \in \mathcal{A}$ with $\omega(\alpha) \leq n$ is

$$[x^n] \, \frac{1}{1 - x} \, \Phi_{\mathcal{A}}(x).$$

Exercise 2.15. For each of the following sets of compositions, obtain a rational function formula for the generating series of that set (with respect to size).

(a) Let $\mathcal{A}$ be the set of compositions of length congruent to 1 (modulo 3).
(b) Let $\mathcal{B}$ be the set of compositions of length congruent to 2 (modulo 3).
(c) Let $\mathcal{C}$ be the set of compositions of even length, with each part being at most 3.
(d) Let $\mathcal{D}$ be the set of compositions of odd length, with each part being at least 2.
(e) Let $\mathcal{E}$ be the set of compositions $\gamma = (c_1, c_2, ..., c_k)$ of any length, in which each part c_i is congruent to i (modulo 2). So c_1 is odd, c_2 is even, c_3 is odd, and so on. (Note that the empty composition $\varepsilon = ()$ is in the set $\mathcal{E}$.)

Exercise 2.16.

(a) Let $\mathcal{A}_n$ be the set of all compositions of size n in which every part is at most three. Obtain a formula for the generating series $\sum_{n=0}^{\infty} |\mathcal{A}_n| \, x^n$.
(b) Let $\mathcal{B}_n$ be the set of all compositions of size n in which every part is a positive integer that is not divisible by three. Obtain a formula for the generating series $\sum_{n=0}^{\infty} |\mathcal{B}_n| \, x^n$.
(c) Deduce that for all $n \geq 3$, $|\mathcal{B}_n| = |\mathcal{A}_n| - |\mathcal{A}_{n-3}|$.
(d)* Can you find a combinatorial proof of part (c)?

Exercise 2.17. Find bijections to explain the equalities in Example 2.21.

Exercise 2.18. Prove that the constructions in the proof of Proposition 2.23 define mutually inverse bijections $\mathcal{U} \rightleftharpoons \mathcal{C} \setminus \{\varepsilon\}$.

Exercise 2.19. For each part, determine the generating series for the number of subsets S of $\{1, 2, ..., n\}$ subject to the stated restriction.

(a) Consecutive elements of S differ by at most 2.
(b) Consecutive elements of S differ by at least 3.
(c) Consecutive elements of S differ by at most 3.
(d) Consecutive elements of S differ by a number congruent to 1 (modulo 3).
(e) Consecutive elements of S differ by a number congruent to 2 (modulo 3).
(f) If $S = \{a_1 < a_2 < \cdots < a_k\}$ then $a_i \equiv i \pmod 2$ for all $1 \le i \le k$.
(g) Fix integers $1 \le g < h$. If $S = \{a_1 < a_2 < \cdots < a_k\}$ then $g \le a_i - a_{i-1} \le h$ for $2 \le i \le k$.

Exercise 2.20. Fix $n, k \in \mathbb{N}$. Let $R(n, k)$ be the number of k-element subsets $S = \{a_1 < a_2 < \cdots < a_k\}$ of $\{1, 2, ..., n\}$ such that $a_i - a_{i-1} \ge i$ for all $2 \le i \le k$. Show that

$$R(n,k) = \binom{n - k(k-1)/2}{k}.$$

Exercise 2.21. Let $p(n)$ be a polynomial function of n.

(a) Prove, by induction on $d = \deg(p)$, that $p(n)$ can be written as a linear combination of $\binom{n+j}{j}$ for $j = 0, 1, 2, ..., d$.
(b) Briefly explain why there is a polynomial $A_p(x)$ such that

$$\sum_{n=0}^{\infty} p(n)x^n = \frac{A_p(x)}{(1-x)^{d+1}}.$$

(c) What can you say about the degree of $A_p(x)$? What can you say about the value of $A_p(1)$?

Chapter 3

Binary Strings.

This chapter presents a wide variety of examples to which the generating series technique of Chapter 2 applies. In Chapter 4 we will see how to use these generating series to answer the counting problems which arise. Similar calculations which provide even more information are presented in Chapter 11.

Definition 3.1. A *binary string* is a finite sequence $\sigma = b_1 b_2 \cdots b_n$ in which each *bit* b_i is either 0 or 1. The number of bits is the *length* of the string, denoted $\ell(\sigma) = n$. Thus, a binary string of length n is an element of the Cartesian power $\{0,1\}^n$. A binary string of arbitrary length is an element of the set $\{0,1\}^* = \bigcup_{n=0}^{\infty} \{0,1\}^n$. There is exactly one binary string $\varepsilon = ()$ of length zero, the empty string with no bits.

Clearly, there are 2^n binary strings of length n, so that the generating series for binary strings with respect to length is

$$\Phi_{\{0,1\}^*}(x) = \sum_{\sigma \in \{0,1\}^*} x^{\ell(\sigma)} = \sum_{n=0}^{\infty} 2^n x^n = \frac{1}{1-2x}.$$

We will see how to describe various subsets of binary strings in a way which allows us to determine their generating series (with respect to length).

3.1 Regular Expressions and Rational Languages.

Definition 3.2 (Regular Expressions.)**.** A *regular expression* is defined recursively, as follows.

- All of ε, 0, and 1 are regular expressions.
- If R and S are regular expressions, then so is R$\smile$S.
- If R and S are regular expressions, then so is RS.
  For any finite $k \in \mathbb{N}$ we also use R^k for the k-fold concatenation of R: that is $\mathsf{R}^2 = \mathsf{RR}$ and $\mathsf{R}^3 = \mathsf{RRR}$, and so on.
- If R is a regular expression, then so is R^*.

For example, $(\varepsilon \smile 0^*00)\,(1^*\,0^*00)^*\,1^*$ is a regular expression. These regular expressions are just formal syntactic constructions with no intrinsic meaning. However, we will interpret them in two different ways.

- A regular expression R will *produce* a subset $\mathcal{R} \subseteq \{0,1\}^*$. Such a subset is called a *rational language*. (See Definition 3.5.)
- A regular expression R will *lead to* a rational function $R(x)$. (See Definition 3.11.)

In general, the rational function $R(x)$ is quite meaningless. However, under favourable conditions on the expression R, it turns out that $R(x) = \Phi_{\mathcal{R}}(x)$ is the generating series of the set $\mathcal{R}$ with respect to length. Then the machinery of Chapter 4 can be applied.

Definition 3.3 (Concatenation Product)**.** Let $\alpha, \beta \in \{0,1\}^*$ be binary strings – so $\alpha = a_1a_2\cdots a_m$ and $\beta = b_1b_2\cdots b_n$. The *concatenation* of α and β is the string

$$\alpha\beta = a_1a_2\cdots a_mb_1b_2\cdots b_n.$$

Let $\mathcal{A}, \mathcal{B} \subseteq \{0,1\}^*$ be sets of binary strings. The *concatenation product* $\mathcal{AB}$ is the set

$$\mathcal{AB} = \{\alpha\beta : \ \alpha \in \mathcal{A} \text{ and } \beta \in \mathcal{B}\}.$$

Example 3.4. Consider the sets $\mathcal{A} = \{011, 01\}$ and $\mathcal{B} = \{101, 1101\}$. There are four ways to concatenate a string in $\mathcal{A}$ followed by a string in $\mathcal{B}$:

$$011.101,\ 011.1101,\ 01.101,\ 01.1101.$$

Here, the dot . indicates the point at which the concatenation takes place. However, this information is not recorded when passing from $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$ to their concatenation $\alpha\beta$. Thus the concatenation product $\mathcal{AB}$ consists of the strings

$$011101,\ 0111101,\ 01101,\ 011101.$$

The string 011101 is produced twice. The concatenation product $\mathcal{AB}$ has only three elements:

$$\mathcal{AB} = \{011101, 0111101, 01101\}.$$

Definition 3.5 (Rational Languages.)**.** A *rational language* is a set $\mathcal{R} \subseteq \{0,1\}^*$ of binary strings that is *produced* by a regular expression; this is defined recursively as follows.

- To begin with, ε produces $\{\varepsilon\}$ and 0 produces $\{0\}$ and 1 produces $\{1\}$.
- If R produces $\mathcal{R}$ and S produces $\mathcal{S}$, then R$\smile$S produces $\mathcal{R} \cup \mathcal{S}$.
- If R produces $\mathcal{R}$ and S produces $\mathcal{S}$, then RS produces $\mathcal{RS}$.
- If R produces $\mathcal{R}$ then R* produces $\mathcal{R}^* = \bigcup_{k=0}^{\infty} \mathcal{R}^k$.
  Here $\mathcal{R}^k$ is the concatenation product of k copies of $\mathcal{R}$.

It is important to note that a rational language can be produced by many different regular expressions, as we will see.

In Definition 3.5, when R produces $\mathcal{R}$ and S produces $\mathcal{S}$, it might happen that $\mathcal{R} \cup \mathcal{S}$ is not a disjoint union of sets. Also, the concatenation product $\mathcal{RS}$ is not the same as the Cartesian product $\mathcal{R} \times \mathcal{S}$, as Example 3.4 shows. These facts lead to complications which are addressed in Section 3.2.

Example 3.6. Here are some easy examples.

- The regular expression 1^* produces the rational language
$$\{1\}^* = \{\varepsilon, 1, 11, 111, 1111, ...\}$$
of all finite strings of 1s (including the empty string ε).
- The regular expression $(1 \smile 11)^*$ also produces the set $\{1\}^*$.
- The regular expression $1(11)^*$ produces the rational language of all strings of 1s of odd (positive) length:
$$\{1, 111, 11111, ...\}.$$
- The regular expression $(0 \smile 1)^*$ produces the rational language $\{0,1\}^*$ of all binary strings.
- The regular expression $1^*(01^*)^*$ also produces the rational language $\{0,1\}^*$ of all binary strings.

Example 3.7. Not every set of binary strings is a rational language.

- The regular expression $(01)^*$ produces the rational language
$$\{\varepsilon, 01, 0101, 010101, 01010101, ...\}.$$
For every even natural number $2j$ there is exactly one string of length $2j$ in this set.
- The set
$$\{\varepsilon, 01, 0011, 000111, 00001111, ...\}$$
is not a rational language, even though for every even natural number $2j$ there is exactly one string of length $2j$ in this set.

The problem with Example 3.7 is that to describe the second set we would need an expression like $\bigcup_{j=0}^{\infty} 0^j 1^j$. However, an infinite union like this is not allowed according to Definition 3.2. The underlying difficulty is that a regular expression has a "finite memory" and cannot remember arbitrarily large numbers, like the $j \in \mathbb{N}$ needed in the above expression. There is a close connection between rational languages and finite state machines, and this is a central topic in the theory of computation.

3.2 Unambiguous Expressions.

Definition 3.8 (Unambiguous Expression)**.** Let R be a regular expression that produces a rational language $\mathcal{R}$. Then R is *unambiguous* if every string in $\mathcal{R}$ is produced exactly once by R. If an expression is not unambiguous then it is *ambiguous*.

As is usual with regular expressions, whether or not it is unambiguous can be decided recursively.

Lemma 3.9 (Unambiguous Expression)**.** *Let R and S be unambiguous expressions producing the sets $\mathcal{R}$ and $\mathcal{S}$, respectively.*

- *The expressions ε and 0 and 1 are unambiguous.*
- *The expression $\mathsf{R} \smile \mathsf{S}$ is unambiguous if and only if $\mathcal{R} \cap \mathcal{S} = \varnothing$, so that $\mathcal{R} \cup \mathcal{S}$ is a disjoint union of sets.*
- *The expression RS is unambiguous if and only if there is a bijection $\mathcal{RS} \rightleftharpoons \mathcal{R} \times \mathcal{S}$ between the concatenation product $\mathcal{RS}$ and the Cartesian product $\mathcal{R} \times \mathcal{S}$. In other words, for every string $\alpha \in \mathcal{RS}$ there is exactly one way to write $\alpha = \rho\sigma$ with $\rho \in \mathcal{R}$ and $\sigma \in \mathcal{S}$.*
- *The expression R^* is unambiguous if and only if each of the concatenation products R^k is unambiguous and the union $\bigcup_{k=0}^{\infty} \mathcal{R}^k$ is a disjoint union of sets.*

Proof. Exercise 3.1. □

Example 3.10. Here are some easy examples.

- The expression 1^* is unambiguous.
- The expression $(1 \smile 11)^*$ is ambiguous: $1.11 = 11.1 = 1.1.1$.
- The expression $(0 \smile 1)^*$ is unambiguous. This expression produces each string in $\{0, 1\}^*$ one bit at a time, so each string is produced in exactly one way.
- The expression $1^*(01^*)^*$ is also unambiguous. First, the expression 01^* is unambiguous, since it produces a 0 followed by a (possibly empty) string of 1s – each such string is produced exactly once.

Now $(01^*)^k$ is unambiguous for any $k \in \mathbb{N}$, since any string it produces will begin with a 0, it will have k bits equal to 0, and the strings of 1s following these 0s can only be constructed in one way. Next, $(01^*)^*$ is unambiguous since for each $k \in \mathbb{N}$, the strings produced by $(01^*)^k$ have exactly k bits equal to 0 (so the union of sets corresponding to the outer $*$ is a disjoint union). Finally, $1^*(01^*)^*$ is unambiguous since for any string it produces, the length of the initial string of 1s is also determined uniquely.

3.2.1 Translation into generating series.

We translate regular expressions into rational functions as follows.

Definition 3.11. A regular expression *leads to* a rational function; this is defined recursively, as follows. Assume that R and S are regular expressions that lead to $R(x)$ and $S(x)$, respectively.

- To begin with, ε leads to 1 and 0 leads to x and 1 leads to x.
- The expression $\mathsf{R} \smile \mathsf{S}$ leads to $R(x) + S(x)$.
- The expression RS leads to $R(x) \cdot S(x)$.
- The expression R^* leads to $1/(1 - R(x))$.

It is easy to see (again recursively!) that if R leads to $R(x)$, then $R(x)$ is a rational function.

Example 3.12. Here are some easy examples.

- The unambiguous expression 1^* leads to $1/(1-x)$.
- The ambiguous expression $(1 \smile 11)^*$ leads to $1/(1-(x+x^2))$. But this expression produces the same rational language as 1^*.
- The unambiguous expression $(0 \smile 1)^*$ leads to
$$\frac{1}{1-(x+x)} = \frac{1}{1-2x}.$$
- The unambiguous expression $1^*(01^*)^*$ leads to
$$\frac{1}{1-x} \cdot \frac{1}{1-x \cdot 1/(1-x)} = \frac{1}{1-2x}.$$

Theorem 3.13. *Let* R *be a regular expression producing the rational language* $\mathcal{R}$ *and leading to the rational function* $R(x)$*. If* R *is an unambiguous expression for* $\mathcal{R}$ *then* $R(x) = \Phi_{\mathcal{R}}(x)$*, the generating series for* $\mathcal{R}$ *with respect to length.*

Sketch of proof. The proof of this is, as usual, recursive. Or, one could say it goes by induction on the complexity of the expression R, and uses the fact that R is unambiguous. Certainly, each of ε, 0, and 1 are unambiguous and lead to the correct generating series for the sets $\{\varepsilon\}$, $\{0\}$, and $\{1\}$, respectively. The induction step follows from Lemma 3.9 and the Sum, Product, and String Lemmas of Subsection 2.2.2, because each of the operations is unambiguous. □

Example 3.14. If the regular expression R producing $\mathcal{R}$ is ambiguous, then the rational function $R(x)$ is in general meaningless.

- For example, consider the regular expression $(\varepsilon \smile 1)^*$, which produces the rational language $\{1\}^*$. It is an ambiguous expression. In fact, it produces every string of 1s in infinitely many ways. The generating series for the set $\{1\}^*$ is $1/(1-x)$. The expression $(\varepsilon \smile 1)^*$ leads to $1/(1-(1+x)) = -x^{-1}$. If this were a generating series it would say that there are exactly -1 objects of size -1, and nothing else. This makes no sense.
- Similarly, the ambiguous expression $(1 \smile 11)^*$ also produces the set $\{1\}^*$. However, the expression $(1 \smile 11)^*$ leads to $1/(1-x-x^2)$, which is also incorrect.

3.2.2 Block decompositions.

Definition 3.15 (Blocks of a string)**.** Let $\sigma = b_1 b_2 b_3 \cdots b_n$ be a binary string of length n. A *block* of σ is a nonempty maximal subsequence of consecutive equal bits. To be clearer, a block is a nonempty subsequence $b_i b_{i+1} \cdots b_j$ of consecutive bits all of which are the same (all are 0, or all are 1), and which cannot be made longer. So, either $i = 1$ or $b_{i-1} \neq b_i$, and either $j = n$ or $b_{j+1} \neq b_j$.

Example 3.16. The blocks of the string $11000101110010001111001011$ are separated by dots here:

$$11.000.1.0.111.00.1.000.1111.00.1.0.11$$

Proposition 3.17 (Block Decompositions.)**.** *The regular expressions*

$$0^*(1^*10^*0)^*1^* \quad \textit{and} \quad 1^*(0^*01^*1)^*0^*$$

are unambiguous expressions for the set $\{0,1\}^*$ *of all binary strings. They produce each binary string block by block.*

Sketch of proof. By symmetry it is enough to consider the first expression. The middle part 1^*10^*0 produces a block of 1s followed by a block of 0s. This concatenation is unambiguous. The repetition of this $(1^*10^*0)^*$ is also unambiguous, since each pass through the repetition starts with a 1 and ends with a 0. (Try it out on Example 3.16.) But the string we want to build might start with a block of 0s: the initial 0^* allows this but does not require it, since $0^* = \varepsilon \smile 0^*0$. The final 1^* similarly allows the string to end with a block of 1s, but does not require it. All the operations are unambiguous, so the whole expression is unambiguous. □

Example 3.18. The block decomposition $0^*(1^*10^*0)^*1^*$ is unambiguous, and produces $\{0,1\}^*$. It had better lead to the right generating series! After a bit of calculation, we see that it leads to

$$\frac{1}{1-x} \cdot \frac{1}{1-x^2/(1-x)^2} \cdot \frac{1}{1-x} = \frac{1}{1-2x},$$

which is good.

Example 3.19. Let $\mathcal{G}$ be the set of binary strings in which every block of 1s has odd length. What is the generating series for $\mathcal{G}$ with respect to length? We will modify the block decomposition $0^*(1^*10^*0)^*1^*$ for all binary strings. The expression 1^*1 in the middle produces a block of 1s. The expression $1^* = \varepsilon \smile 1^*1$ produces either the empty string or a block of 1s. If we want a block of 1s of odd length, then that is produced by $(11)^*1$. So the expression

$$\mathsf{G} = 0^* \left((11)^*1\,0^*0\right)^* (\varepsilon \smile (11)^*1)$$

is a block decomposition for the set $\mathcal{G}$ in question. It is therefore an unambiguous expression that produces $\mathcal{G}$. This expression leads to

$$\begin{aligned} G(x) &= \frac{1}{1-x} \cdot \frac{1}{1-(x/(1-x^2))(x/(1-x))} \cdot \left(1+\frac{x}{1-x^2}\right) \\ &= \frac{1+x-x^2}{(1-x)(1-x^2)-x^2} = \frac{1+x-x^2}{1-x-2x^2+x^3}. \end{aligned}$$

By Theorem 3.13, $G(x) = \Phi_{\mathcal{G}}(x)$ is the generating series of the set $\mathcal{G}$. It follows from Theorem 4.8 that the number g_n of strings of length n in $\mathcal{G}$ satisfies the linear recurrence relation with initial conditions given by

$$g_n - g_{n-1} - 2g_{n-2} + g_{n-3} = \begin{cases} 1 & \text{if } n = 0, \\ 1 & \text{if } n = 1, \\ -1 & \text{if } n = 2, \\ 0 & \text{if } n \geq 3, \end{cases}$$

(with the convention that $g_n = 0$ for all $n < 0$). This gives the initial conditions $g_0 = 1$, $g_1 = 2$, $g_2 = 3$ (which can be checked directly), and the recurrence $g_n = g_{n-1} + 2g_{n-2} - g_{n-3}$ for all $n \geq 3$. It is easy to calculate the first several of these numbers.

n	0	1	2	3	4	5	6	7	8
g_n	1	2	3	6	10	19	33	61	108

To get an exact formula for g_n we would need to factor the denominator $1 - x - 2x^2 + x^3$ to find its inverse roots. They turn out to be slightly horrible complex numbers, so we won't do it.

Example 3.20. Let $\mathcal{H}$ be the set of binary strings in which each block of 0s has length one. It is not hard to see that $(\varepsilon \smile 0)\,(1^*1\,0)^*\,1^*$ is a block decomposition for $\mathcal{H}$, and is therefore unambiguous. This expression leads to the formula

$$H(x) = \sum_{n=0}^{\infty} h_n x^n = (1+x) \cdot \frac{1}{1-x^2/(1-x)} \cdot \frac{1}{1-x} = \frac{1+x}{1-x-x^2}.$$

This resembles Examples 2.21 and 2.24 yet again! These are the Fibonacci numbers – see Example 4.1. With $1/(1-x-x^2) = \sum_{n=0}^{\infty} f_n x^n$, we see that for $n \geq 1$,

$$h_n = [x^n]H(x) = f_n + f_{n-1} = f_{n+1}.$$

Let's reality check this for $n = 5$. The strings in $\mathcal{H}_5$ are 11111, five strings with one 0, six strings with two 0s, and 01010. And indeed, $f_6 = 13$. Neat!

3.2.3 Prefix decompositions.

Example 3.21. Consider the regular expression $(0^*1)^*0^*$. We claim that this is unambiguous. To see this, let $\sigma = b_1 b_2 \cdots b_n$ be a binary string produced by this expression. As an example, take

$$0011101011000101100.$$

How can this string be produced? The repetition $(0^*1)^*$ produces a binary string by chopping it into pieces after each occurrence of the bit 1. But any string produced by this expression is either empty or ends with a 1. The final "suffix" 0^* in the expression allows the possibility that the string might end with some 0s. The string above is produced as

$$001.1.1.01.01.1.0001.01.1.000$$

and this is the only way it is produced. This rule – "chop the string into pieces after each occurrence of the bit 1" – gives a unique way to produce each binary string from the expression $(0^*1)^*0^*$. It follows that $(0^*1)^*0^*$ is an unambiguous expression for the set $\{0,1\}^*$ of all binary strings. And, indeed, $(0^*1)^*0^*$ leads to

$$\frac{1}{1-x/(1-x)} \cdot \frac{1}{1-x} = \frac{1}{1-2x}.$$

In general, a *prefix decomposition* for a set of binary strings is a regular expression of the form $\mathsf{A}^*\mathsf{B}$. The idea is to chop each string that is produced into initial segments produced by the expression A. There might be a terminal segment produced by the expression B. (It is possible that $\mathsf{B} = \varepsilon$.) One does have to argue that such expressions are unambiguous, but – as in Example 3.21 – this can usually be done by checking that:

- there is at most one way for a binary string to begin with an initial segment produced by A, and
- if the string does not begin with an initial segment produced by A then it is produced by B.

Both of the expressions A and B must be unambiguous, too, of course.

Similarly, a *postfix decomposition* has the form $\mathsf{A}(\mathsf{B}^*)$.

3.3 Recursive Decompositions.

Recursive decompositions are more general than regular expressions, and can produce sets of binary strings that are more general than rational languages. The added feature is that in addition to the elementary building blocks "ε" and "0" and "1" in a regular expression, we are also allowed to use letters that stand for sub-expressions on both sides of an equation.

Example 3.22. A simple example is the expression $\mathsf{S} = \varepsilon \smile (0 \smile 1)\mathsf{S}$. Since S appears on both the LHS and the RHS, in some sense this "defines S in terms of itself". This is the key idea behind recursive expressions. One reads this expression as saying that a string produced by S is either empty, or it consists of a single bit (either 0 or 1) followed by another (shorter) string produced by S. By induction on the length of the string, one sees that every binary string is produced exactly once in this way.

Translating the above expression into rational functions by analogy with Definition 3.11 leads to an equation for $S(x)$:

$$S(x) = 1 + (x + x)S(x).$$

This is easily solved to yield $S(x) = 1/(1 - 2x)$.

Example 3.23. As in Example 3.7, consider the set of strings

$$\mathcal{B} = \{\varepsilon, 01, 0011, 000111, ...\}.$$

This is not a rational language, but it can be described by the unambiguous recursive expression $\mathsf{B} = \varepsilon \smile 0\,\mathsf{B}\,1$. This leads to the equation $B(x) = 1 + x^2 B(x)$, and thence to $B(x) = 1/(1 - x^2)$.

Subsection 4.4.2 presents a more substantial example of a recursive decomposition which produces a set of binary strings that is not a rational language. The generating series for that example is not even a rational function.

3.3.1 Excluded substrings.

Let $\kappa \in \{0,1\}^*$ be a nonempty binary string. We say that $\sigma \in \{0,1\}^*$ *contains* κ if there are (possibly empty) binary strings α, β such that $\sigma = \alpha\,\kappa\,\beta$. If σ does not contain κ then σ *avoids* or *excludes* κ. Let $\mathcal{A}_\kappa \subset \{0,1\}^*$ be set of strings that avoid κ. We will develop a general method for calculating the generating series $A_\kappa(x)$.

Example 3.24. As an easy first example, consider the case $\kappa = 01011$. Let $\mathcal{A}$ be the set of strings avoiding 01011, and let $\mathcal{B}$ be the set of strings that have exactly one occurrence of 01011, at the very end (that is, as a suffix). Consider the strings in $\mathcal{A} \cup \mathcal{B}$. Such a string is either empty, or it ends with either a 0 or a 1. If this string is not empty, then removing the last bit leaves a (possibly empty) string in $\mathcal{A}$ (because of the way the sets $\mathcal{A}$ and $\mathcal{B}$ are defined). This translates into the relation

$$\mathsf{A} \smile \mathsf{B} \;=\; \varepsilon \smile \mathsf{A}\,(0 \smile 1)$$

for expressions A and B producing $\mathcal{A}$ and $\mathcal{B}$, respectively. This leads to the equation

$$A(x) \; + \; B(x) \; = \; 1 \; + \; 2x\,A(x)$$

for the generating series.

We need another equation in order to determine $A(x)$ and $B(x)$. If $\beta = \alpha\,01011$ is an arbitrary string in $\mathcal{B}$, then α is in $\mathcal{A}$, again from the way the sets $\mathcal{A}$ and $\mathcal{B}$ are defined. We claim that the converse also holds: if $\alpha \in \mathcal{A}$ then $\beta = \alpha\,01011$ is in $\mathcal{B}$. To see this we need to show that the only occurrence of 01011 as a substring of β is the one at the very end. We know that $\alpha \in \mathcal{A}$ does not contain 01011 as a substring. So if there is another occurrence of 01011 inside β then it must "overlap" the final 01011 in at least one position (but not in all positions). For this particular excluded substring this is not possible, as the following table shows:

$$\begin{array}{cccc|ccccc}
. & . & . & . & 0 & 1 & 0 & 1 & 1 \\ \hline
0 & 1 & 0 & 1 & \underline{1} & . & . & . & . \\
. & 0 & 1 & 0 & \underline{1} & 1 & . & . & . \\
. & . & 0 & 1 & 0 & 1 & \underline{1} & . & . \\
. & . & . & 0 & \underline{1} & \underline{0} & \underline{1} & 1 & .
\end{array}$$

In each row after the first, there is at least one position at which the shifted 01011 disagrees with the substring in the first row. So 01011 cannot overlap itself in a nontrivial way. This gives the relation $\mathsf{B} = \mathsf{A}\,01011$, yielding the equation $B(x) = x^5 A(x)$. Substituting this into the first equation gives $1 + 2x\,A(x) = (1 + x^5)\,A(x)$, which is easily solved. We conclude that $A(x) = 1/(1 - 2x + x^5)$.

Example 3.25. As a slightly trickier example, consider the case $\kappa = 01101$. Again, let $\mathcal{A}$ be the set of strings avoiding 01101, and let $\mathcal{B}$ be the set of strings that have exactly one occurrence of 01101, at the very end (that is, as a suffix). As in the previous example, a string in $\mathcal{A} \cup \mathcal{B}$ is either empty, or it ends with either a 0 or a 1. The reasoning of the previous example also holds in this case, translating into the same relation

$$\mathsf{A} \smile \mathsf{B} \; = \; \varepsilon \smile \mathsf{A}\,(0 \smile 1)$$

and the same equation $A(x) \; + \; B(x) \; = \; 1 \; + \; 2x\,A(x)$ for the generating series.

The second equation is the slightly trickier part, because the string 01101 can overlap itself in a nontrivial way. As in the previous example, we still have the set inclusion $\mathcal{B} \subseteq \mathcal{A}\,01101$. But the reverse inclusion does not hold in this case: for example $011.01101 = 01101.101$ is in $\mathcal{A}\,01101$ but not in $\mathcal{B}$, because it contains a substring 01101 that is not at the very end.

$$\begin{array}{cccc|ccccc}
. & . & . & . & 0 & 1 & 1 & 0 & 1 \\
\hline
0 & 1 & 1 & 0 & \underline{1} & . & . & . & . \\
. & 0 & 1 & 1 & 0 & 1 & . & . & . \\
. & . & 0 & 1 & \underline{1} & \underline{0} & 1 & . & . \\
. & . & . & 0 & \underline{1} & 1 & \underline{0} & \underline{1} & .
\end{array}$$

Looking at all the possible ways that 01101 can overlap itself, we see that in this case $\mathcal{A}\,01101 = \mathcal{B} \cup \mathcal{B}\,101$. This gives $\mathsf{A}\,01101 = \mathsf{B}\,(\varepsilon \smile 101)$ and $x^5 A(x) = (1+x^3)B(x)$. Substituting $B(x) = x^5 A(x)/(1+x^3)$ into the first equation and solving for $A(x)$ yields

$$A(x) = \frac{1+3x+2x^2}{1-2x-2x^2+x^3+x^5}.$$

The proof of the general result follows exactly the same pattern.

Theorem 3.26. *Let $\kappa \in \{0,1\}^*$ be a nonempty string of length n, and let $\mathcal{A} = \mathcal{A}_\kappa$ be the set of binary strings that avoid κ. Let $\mathcal{C}$ be the set of all nonempty suffixes γ of κ such that $\kappa\,\gamma = \eta\,\kappa$ for some nonempty prefix η of κ. Let $C(x) = \sum_{\gamma\in\mathcal{C}} x^{\ell(\gamma)}$. Then*

$$A(x) = \frac{1+C(x)}{(1-2x)(1+C(x))+x^n}.$$

Proof. Let $\mathcal{B}$ be the set of strings that contain exactly one occurrence of κ, at the very end. Any string in $\mathcal{A} \cup \mathcal{B}$ is either empty or in the set $\mathcal{A}\{0,1\}$, so that

$$\mathcal{A} \cup \mathcal{B} = \{\varepsilon\} \cup \mathcal{A}\,\{0,1\}.$$

This yields $A(x) + B(x) = 1 + 2x\,A(x)$.

The key observation is that if $\beta \in \mathcal{B}$ and ρ is a proper prefix of β then $\rho \in \mathcal{A}$. Consequently, $\mathcal{B} \subseteq \mathcal{A}\,\kappa$. Conversely, consider any $\sigma = \alpha\,\kappa$ in $\mathcal{A}\,\kappa$.

Maybe $\sigma \in \mathcal{B}$. If $\sigma \notin \mathcal{B}$, then there is an "early" occurrence of κ in $\alpha\,\kappa$ that is not the occurrence at the very end. All such early occurrences of κ must overlap the final occurrence of κ nontrivially because $\alpha \in \mathcal{A}$. Consider the **first** such early occurrence of κ in $\sigma = \alpha\,\kappa$. Then there is a nonempty suffix $\gamma \in \mathcal{C}$ of κ, and a nonempty prefix η of κ, and a (possibly empty) prefix ρ of α such that

$$\sigma = \alpha\,\kappa = \rho\,\eta\,\kappa = \rho\,\kappa\,\gamma.$$

Since we are looking at the first early occurrence of κ in σ, the substring $\rho\,\kappa$ is in $\mathcal{B}$. Moreover, this argument shows that these substrings ρ, η, and γ are determined uniquely from σ. This shows that every $\sigma \in \mathcal{A}\,\kappa$ is either in $\mathcal{B}$, or is in $\mathcal{B}\,\gamma$ for exactly one $\gamma \in \mathcal{C}$, and this decomposition is unique. That is,

$$\mathcal{A}\,\kappa = \mathcal{B} \cup \bigcup_{\gamma \in \mathcal{C}} \mathcal{B}\,\gamma$$

is a disjoint union of sets. Translating this into generating series yields

$$x^n\,A(x) \;=\; (1 + C(x))\,B(x).$$

Solving this for $B(x)$ and substituting this into the first equation gives

$$1 + 2x\,A(x) = A(x) \cdot \left(1 + \frac{x^n}{1 + C(x)}\right).$$

Solving this for $A(x)$ yields the result. □

In Section 12.1 we use a completely different method to solve a vast generalization of this excluded substring problem

3.4 Exercises.

Exercise 3.1. Prove Lemma 3.9.

Exercise 3.2. Let $\mathsf{A} = (10 \smile 101)$ and $\mathsf{B} = (001 \smile 100 \smile 1001)$. For each of AB and BA, is the expression unambiguous? What is the generating series (by length) of the set it produces?

Exercise 3.3. Let $\mathsf{A} = (00 \smile 101 \smile 11)$ and $\mathsf{B} = (00 \smile 001 \smile 10 \smile 110)$. Prove that A^* is unambiguous, and that B^* is ambiguous. Find the generating series (by length) for the set $\mathcal{A}^*$ produced by A^*.

Exercise 3.4. For each of the following sets of binary strings, write an unambiguous expression which produces that set.

(a) Binary strings that have no block of 0s of size 3, and no block of 1s of size 2.
(b) Binary strings that have no substring of 0s of length 3, and no substring of 1s of length 2.
(c) Binary strings in which the substring 011 does not occur.
(d) Binary strings in which the blocks of 0s have even length and the blocks of 1s have odd length.

Exercise 3.5. Let $\mathsf{G} = 0^*((11)^*1(00)^*00 \smile (11)^*11(00)^*0)^*$, and let $\mathcal{G}$ be the set of binary strings produced by G.

(a) Give a verbal description of the strings in the set $\mathcal{G}$.
(b) Find the generating series (by length) of $\mathcal{G}$.
(c) For $n \in \mathbb{N}$, let g_n be the number of strings in $\mathcal{G}$ of length n. Give a recurrence relation and enough initial conditions to uniquely determine g_n for all $n \in \mathbb{N}$.

Exercise 3.6.

(a) Show that the generating series (by length) for binary strings in which every block of 0s has length at least 2 and every block of 1s has length at least 3 is

$$\frac{(1-x+x^3)(1-x+x^2)}{1-2x+x^2-x^5}.$$

(b) Give a recurrence relation and enough initial conditions to determine the coefficients of this power series.

Exercise 3.7.

(a) For $n \in \mathbb{N}$, let h_n be the number of binary strings of length n such that each even-length block of 0s is followed by a block of exactly one 1 and each odd-length block of 0s is followed by a block of exactly two 1s. Show that

$$h_n = [x^n]\frac{1+x}{1-x^2-2x^3}.$$

(b) Give a recurrence relation and enough initial conditions to determine h_n for all $n \in \mathbb{N}$.

Exercise 3.8. Let $\mathcal{K}$ be the set of binary strings in which any block of 1s which immediately follows a block of 0s must have length at least as great as the length of that block of 0s. (Note: this is *not* a rational language.)

(a) Derive a formula for $K(x) = \sum_{\alpha \in \mathcal{K}} x^{\ell(\alpha)}$.
(b) Give a recurrence relation and enough initial conditions to determine the coefficients $[x^n]K(x)$ for all $n \in \mathbb{N}$.

Exercise 3.9.

(a) Fix an integer $m \geq 1$. Find the generating series (by length) of the set of binary strings in which no block has length greater than m.
(b) Fix integers $m, k \geq 1$. Find the generating series (by length) of the set of binary strings in which no block of 0s has length greater than m, and no block of 1s has length greater than k.

Exercise 3.10. Let $\mathcal{L}$ be the set of binary strings in which each block of 1s has odd length, and which do not contain the substring 0001. Let $\mathcal{L}_n$ be the set of strings in $\mathcal{L}$ of length n, and let $L(x) = \sum_{n=0}^{\infty} |\mathcal{L}_n| x^n$.

(a) Give an expression that produces the set $\mathcal{L}$ unambiguously, and explain briefly why it is unambiguous and produces $\mathcal{L}$.

(b) Use your expression from part (a) to show that

$$L(x) = \frac{1 + x - x^2}{1 - x - 2x^2 + x^3 + x^4}.$$

Exercise 3.11. Let $\mathcal{M}$ be the set of binary strings in which each block of 0s has length at most two, and which do not contain 00111 as a substring. Let $\mathcal{M}_n$ be the set of strings in $\mathcal{M}$ of length n, and let $M(x) = \sum_{n=0}^{\infty} |\mathcal{M}_n| x^n$.

(a) Give an expression that produces the set $\mathcal{M}$ unambiguously, and explain briefly why it is unambiguous and produces $\mathcal{M}$.

(b) Use your expression from part (a) to show that

$$M(x) = \frac{1 + x + x^2}{1 - x - x^2 - x^3 + x^5}.$$

Exercise 3.12. Let $\mathcal{N}$ be the set of binary strings in which each block of 0s has odd length, and each block of 1s has length one or two. Let $\mathcal{N}_n$ be the set of strings in $\mathcal{N}$ of length n, and let $N(x) = \sum_{n=0}^{\infty} |\mathcal{N}_n| x^n$.

(a) Show that

$$N(x) = \frac{1 + 2x + x^2 - x^4}{1 - 2x^2 - x^3} = -2 + x + \frac{3 + x - 3x^2}{1 - 2x^2 - x^3}.$$

(b) Derive an exact formula for $|\mathcal{N}_n|$ as a function of n.

Exercise 3.13. For $n \in \mathbb{N}$, let p_n be the number of binary strings of length n in which every block of 0s is followed by a block of 1s with the same parity of length.

(a) Determine the generating series $P(x) = \sum_{n=0}^{\infty} p_n x^n$.

(b) Show that if $n \geq 2$, then $p_n = 2 \cdot 3^{\lfloor n/2 \rfloor - 1}$.

Exercise 3.14.

(a) Let $\mathcal{Q}$ be the set of binary strings that do not contain 11000 as a substring. For $n \in \mathbb{N}$, let $\mathcal{Q}_n$ be the set of strings in $\mathcal{Q}$ of length n. Obtain a formula for the generating series $Q(x) = \sum_{n=0}^{\infty} |\mathcal{Q}_n| x^n$, with a brief explanation.

(b) Let $\mathcal{R}$ be the set of compositions, of any length, in which each part is at most 4. For $n \in \mathbb{N}$, let $\mathcal{R}_n$ be the set of compositions in $\mathcal{R}$ of size n. Obtain a formula for the generating series $R(x) = \sum_{n=0}^{\infty} |\mathcal{R}_n| x^n$, with a brief explanation.

(c) Deduce that for all integers $n \geq 1$, $|\mathcal{R}_n| = |\mathcal{Q}_n| - |\mathcal{Q}_{n-1}|$.

(d)* Part (c) implies that for every integer $n \geq 1$, there is a bijection $\mathcal{Q}_n \rightleftharpoons \mathcal{R}_n \cup \mathcal{Q}_{n-1}$. Can you determine such a bijection precisely?

Exercise 3.15. Let $\mathcal{V}$ be the set of binary strings that do not contain 0110 as a substring. Show that the generating series (by length) for $\mathcal{V}$ is

$$V(x) = \Phi_{\mathcal{V}}(x) = \frac{1 + x^3}{1 - 2x + x^3 - x^4}.$$

Exercise 3.16.

(a) Let $\mathcal{W}$ be the set of binary strings that do not contain 0101 as a substring. Obtain a formula for the generating series (by length) of $\mathcal{W}$.

(b) Fix a positive integer $r \geq 1$ and consider the binary string $(01)^r$. (Part (a) is the case $r = 2$.) Obtain a formula for the generating series of the set of binary strings that do not contain $(01)^r$.

Exercise 3.17. Let $\mathsf{S} = \mathsf{A}^*\mathsf{B}$ be an unambiguous prefix decomposition producing some set of strings $\mathcal{S} \subseteq \{0,1\}^*$. Show that the recursion $\mathsf{R} = \mathsf{B} \smile \mathsf{A}\,\mathsf{R}$ defines an expression R that produces the same set of strings $\mathcal{S} \subseteq \{0,1\}^*$. Also check that both S and R lead to the rational function $B(x)/(1 - A(x))$.

Chapter 4

Recurrence Relations.

In Chapters 2 and 3 we saw how to encode a sequence of numbers as the coefficients of a power series $G(x) = \sum_{n=0}^{\infty} g_n x^n$. We used the Sum, Product, and String Lemmas to obtain algebraic formulas for these generating series. In this section we will see two techniques for using these algebraic formulas to compute the coefficients g_n, which are the numbers we really want. First we will do the example of Fibonacci numbers in detail, and then we will develop the theory in general.

4.1 Fibonacci Numbers.

Example 4.1. The sequence of Fibonacci numbers $\mathbf{f} = (f_0, f_1, f_2, f_3, ...)$ is defined by putting $f_0 = 1$, $f_1 = 1$, and $f_n = f_{n-1} + f_{n-2}$ for all $n \geq 2$. One can use this information to compute f_n iteratively for as long as you want:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
f_n	1	1	2	3	5	8	13	21	34	55	89	144	233	377	610

Can we obtain a formula for f_n as a function of n?

The answer is "yes", of course – we obtain such a formula in this section. Moreover, this is just the first example of a very general technique, which is the main subject of this chapter.

Example 4.2. We start by finding a formula for the generating series $F(x) = \sum_{n=0}^{\infty} f_n x^n$. From the information defining the Fibonacci numbers, we see that

$$F(x) = f_0 + f_1 x + \sum_{n=2}^{\infty} f_n x^n = 1 + x + \sum_{n=2}^{\infty} (f_{n-1} + f_{n-2}) x^n.$$

The next step is to write the RHS in terms of $F(x)$.

$$\begin{aligned} F(x) &= 1 + x + \sum_{n=2}^{\infty} f_{n-1} x^n + \sum_{n=2}^{\infty} f_{n-2} x^n \\ &= 1 + x + x \sum_{i=1}^{\infty} f_i x^i + x^2 \sum_{j=0}^{\infty} f_j x^j \\ &= 1 + x + x(F(x) - f_0) + x^2 F(x) \\ &= 1 + xF(x) + x^2 F(x). \end{aligned}$$

This equation can be solved for $F(x)$, yielding

$$F(x) = \frac{1}{1 - x - x^2}.$$

We have seen this generating series before, relating to the sets of compositions $\mathcal{F}$, $\mathcal{H}$, and $\mathcal{J}$ in Example 2.21. Obtaining a formula for Fibonacci numbers will thus solve the counting problem for each of these sets of compositions.

The key is the denominator of the series, in this case $1-x-x^2$.

Example 4.3. We factor the denominator in the form

$$1 - x - x^2 = (1 - \alpha x)(1 - \beta x)$$

for some complex numbers α and β, called the *inverse roots* of the polynomial. To do this we can use the Quadratic Formula, but since we are looking for the **inverse roots** of a polynomial we have to be careful. Substitute $x = 1/t$ and multiply both sides by t^2 to get

$$t^2 - t - 1 = (t - \alpha)(t - \beta).$$

Now the inverse roots of the denominator $1 - x - x^2$ are the (ordinary) roots of this *auxiliary polynomial* $t^2 - t - 1$. By the Quadratic Formula, they are

$$\left.\begin{array}{c}\alpha\\ \beta\end{array}\right\} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{5}}{2}.$$

Example 4.4. Having found the inverse roots of the denominator, the next step is to apply the Partial Fractions Theorem 4.12, which will be explained (and proved) in Section 4.3. In this case it implies that there are complex numbers A and B such that

$$\frac{1}{1 - x - x^2} = \frac{1}{(1 - \alpha x)(1 - \beta x)} = \frac{A}{1 - \alpha x} + \frac{B}{1 - \beta x}.$$

There are a few different ways to find the numbers A and B, as we will see. Here we can multiply by $1 - x - x^2 = (1 - \alpha x)(1 - \beta x)$ and collect like powers of x:

$$1 = A(1 - \beta x) + B(1 - \alpha x) = (A + B) - (A\beta + B\alpha)x.$$

Comparing coefficients, we see that $A + B = 1$ and $A\beta + B\alpha = 0$.
From $A\alpha + B\alpha = \alpha$ and $A\beta + B\alpha = 0$, we see that

$$A = \frac{\alpha}{\alpha - \beta} = \frac{5 + \sqrt{5}}{10}.$$

Similarly, from $A\beta + B\beta = \beta$ and $A\beta + B\alpha = 0$, we see that

$$B = \frac{\beta}{\beta - \alpha} = \frac{5 - \sqrt{5}}{10}.$$

Now all that remains is to put the pieces of this calculation together.

Example 4.5. We apply the geometric series expansion to the result of the partial fractions decomposition. (More generally, we would use binomial series.)

$$\frac{A}{1-\alpha x}+\frac{B}{1-\beta x}=A\sum_{n=0}^{\infty}\alpha^n x^n+B\sum_{n=0}^{\infty}\beta^n x^n=\sum_{n=0}^{\infty}\left(A\alpha^n+B\beta^n\right)x^n.$$

It follows that for all $n\in\mathbb{N}$, the Fibonacci numbers are given by the formula

$$\begin{aligned} f_n=[x^n]F(x) &= [x^n]\frac{1}{1-x-x^2} = A\alpha^n+B\beta^n \\ &= \frac{5+\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n+\frac{5-\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n. \end{aligned}$$

That seems kind of weird, since we know that the Fibonacci numbers are integers. But notice that $\beta=(1-\sqrt{5})/2\approx-0.618$ so that as $n\to\infty$, $\beta^n\to 0$. In fact, for all $n\in\mathbb{N}$, f_n is the integer closest to $\frac{5+\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n$.

4.2 Homogeneous Linear Recurrence Relations.

Definition 4.6 (Homogeneous linear recurrence relation)**.** Let $\mathrm{g}=(g_0,g_1,g_2,\ldots)$ be an infinite sequence of complex numbers. Let a_1, a_2, ..., a_d be in $\mathbb{C}$, and let $N\geq d$ be an integer. We say that g satisfies a *homogeneous linear recurrence relation* provided that

$$g_n + a_1g_{n-1} + a_2g_{n-2} + \cdots + a_dg_{n-d} = 0$$

for all $n\geq N$. The values g_0, g_1, ..., g_{N-1} are the *initial conditions* of the recurrence. The relation is *linear* because the LHS is a linear combination of the entries of the sequence g; it is *homogeneous* because the RHS of the equation is zero.

In Definition 4.6, if the RHS of the equation is instead a non-zero function $p:\mathbb{N}\to\mathbb{C}$, then this is an *inhomogeneous* linear recurrence relation.

The recurrence relation can be rewritten as

$$g_n=-(a_1g_{n-1}+a_2g_{n-2}+\cdots+a_dg_{n-d})$$

for all $n\geq N$, and this can be used to compute the numbers g_n by induction on n, using the initial conditions as the basis of induction.

Consider the Fibonacci numbers from Example 4.1: $f_0 = f_1 = 1$ are the initial conditions, and $f_n - f_{n-1} - f_{n-2} = 0$ for all $n \geq 2$ is a homogeneous linear recurrence relation. We derived the formula

$$F(x) = \sum_{n=0}^{\infty} f_n x^n = \frac{1}{1 - x - x^2}$$

for the generating series. This is an instance of a general fact about a sequence with a homogeneous linear recurrence relation. Here is another example before we see the general theory.

Example 4.7. Define a sequence $\mathbf{g} = (g_0, g_1, ...)$ by the initial conditions $g_0 = 2$, $g_1 = 5$, and $g_2 = 6$, and the relation $g_n - 3g_{n-2} - 2g_{n-3} = 0$ for all $n \geq 3$. Obtain a formula for the generating series $G(x) = \sum_{n=0}^{\infty} g_n x^n$.

The general method is to multiply the recurrence by x^n and sum over all $n \geq N$. In this case

$$\sum_{n=3}^{\infty} (g_n - 3g_{n-2} - 2g_{n-3})\, x^n = 0.$$

Now we split the LHS into separate summations, reindex them, and write everything in terms of the power series $G(x)$

$$\begin{aligned}
\sum_{n=3}^{\infty} g_n x^n - 3 \sum_{n=3}^{\infty} g_{n-2} x^n - 2 \sum_{n=3}^{\infty} g_{n-3} x^n &= 0 \\
(G(x) - g_0 - g_1 x - g_2 x^2) - 3x^2 \sum_{j=1}^{\infty} g_j x^j - 2x^3 \sum_{k=0}^{\infty} g_k x^k &= 0 \\
(G(x) - 2 - 5x - 6x^2) \;-\; 3x^2\, (G(x) - 2) \;-\; 2x^3\, G(x) &= 0 \\
G(x) \;-\; 3x^2\, G(x) \;-\; 2x^3\, G(x) &= 2 + 5x.
\end{aligned}$$

It follows that

$$G(x) = \frac{2 + 5x}{1 - 3x^2 - 2x^3}.$$

Notice how the polynomial $1 - 3x^2 - 2x^3$ in the denominator of this formula is related to the linear recurrence relation $g_n - 3g_{n-2} - 2g_{n-3} = 0$ for $n \geq 3$. We can explain the numerator, too, if we make the convention that

$g_n = 0$ for all integers $n < 0$. Then, using the initial conditions, we have

$$\begin{aligned} g_0 - 3g_{-2} - 2g_{-3} &= 2 - 0 - 0 = 2 \text{ for } n = 0, \\ g_1 - 3g_{-1} - 2g_{-2} &= 5 - 0 - 0 = 5 \text{ for } n = 1, \\ g_2 - 3g_0 - 2g_{-1} &= 6 - 3 \cdot 2 - 0 = 0 \text{ for } n = 2, \\ g_n - 3g_{n-2} - 2g_{n-3} &= 0 \text{ for } n \geq 3. \end{aligned}$$

Theorem 4.8. *Let $\mathbf{g} = (g_0, g_1, g_2, ...)$ be a sequence of complex numbers, and let $G(x) = \sum_{n=0}^{\infty} g_n x^n$ be the corresponding generating series. The following are equivalent.*

(a) The sequence $\mathbf{g}$ satisfies a homogeneous linear recurrence relation

$$g_n + a_1 g_{n-1} + \cdots + a_d g_{n-d} = 0 \text{ for all } n \geq N,$$

with initial conditions g_0, g_1, ..., g_{N-1}.

(b) The series $G(x) = P(x)/Q(x)$ is a quotient of two polynomials. The denominator is

$$Q(x) = 1 + a_1 x + a_2 x^2 + \cdots + a_d x^d$$

and the numerator is $P(x) = b_0 + b_1 x + b_2 x^2 + \cdots + b_{N-1} x^{N-1}$, in which

$$b_k = g_k + a_1 g_{k-1} + \cdots + a_d g_{k-d}$$

for all $0 \leq k \leq N - 1$, with the convention that $g_n = 0$ for all $n < 0$.

Proof. To prove this theorem, we just copy the calculation in Example 4.7, but do it in the most general case. For convenience, let $a_0 = 1$. Assume that part (a) holds, and let

$$Q(x) = 1 + a_1 x + a_2 x^2 + \cdots + a_d x^d.$$

Consider the product $Q(x)G(x)$:

$$\begin{aligned} Q(x)G(x) &= \left(\sum_{j=0}^{d} a_j x^j \right) \left(\sum_{n=0}^{\infty} g_n x^n \right) \\ &= \sum_{j=0}^{d} \sum_{n=0}^{\infty} a_j g_n x^{n+j} = \sum_{k=0}^{\infty} \left(\sum_{j=0}^{d} a_j g_{k-j} \right) x^k. \end{aligned}$$

In the last step we have re-indexed the double sum using $k = n + j$, and used the convention that $g_n = 0$ for all $n < 0$.

The coefficient of x^k in this formula is $g_k + a_1 g_{k-1} + \cdots + a_d g_{k-d}$. This is the LHS of the recurrence relation for **g** applied when $n = k$. Thus, this coefficient is zero for $k \geq N$. On the other hand, for $0 \leq k \leq N-1$, we see that it is $\sum_{j=0}^{d} a_j g_{k-j} = b_k$ by the way the numbers b_k are defined. That is,

$$Q(x)G(x) = \sum_{k=0}^{N-1} b_k x^k = P(x),$$

and it follows that $G(x) = P(x)/Q(x)$. This shows that (a) implies (b).

Conversely, assume that (b) holds and that $G(x) = P(x)/Q(x)$ is as given. We essentially run the argument for the first part of the proof in reverse. The equations $b_k = g_k + a_1 g_{k-1} + \cdots + a_d g_{k-d}$ for $0 \leq k \leq N-1$ (with the convention that $g_n = 0$ for $n < 0$) determine the initial conditions g_0, g_1,... g_{N-1} inductively. For $n \geq N$, the coefficient of x^n in $P(x) = Q(x)G(x)$ is zero. This implies that $g_k + a_1 g_{k-1} + \cdots + a_d g_{k-d} = 0$ for all $k \geq N$, showing that (b) implies (a). $\square$

Theorem 4.8 is useful in both directions, as the following two examples show.

Example 4.9. Let $D(x) = \sum_{n=0}^{\infty} d_n x^n = \dfrac{1 - 3x + 4x^2}{1 - 2x + 3x^3}$. Obtain a homogeneous linear recurrence relation and initial conditions satisfied by the sequence $\mathbf{d} = (d_0, d_1, d_2, ...)$.

From Theorem 4.8 we can read from the RHS that for all $n \in \mathbb{N}$:

$$d_n - 2d_{n-1} + 3d_{n-3} = \begin{cases} 1 & \text{if } n = 0, \\ -3 & \text{if } n = 1, \\ 4 & \text{if } n = 2, \\ 0 & \text{if } n \geq 3, \end{cases}$$

with the convention that $d_n = 0$ if $n < 0$. We can determine the initial conditions inductively as follows: $d_0 = 1$; $d_1 - 2d_0 = -3$, so $d_1 = -1$; $d_2 - 2d_1 = 4$, so $d_2 = 2$. The recurrence $d_n - 2d_{n-1} + 3d_{n-3} = 0$ holds for all $n \geq 3$. Using the initial conditions $d_0 = 1$, $d_1 = -1$, $d_2 = 2$, and the recurrence $d_n = 2d_{n-1} - 3d_{n-3}$ for all $n \geq 3$, we can compute d_n for as long as we want:

n	0	1	2	3	4	5	6	7	8
d_n	1	-1	1	1	5	4	5	-5	-22

Example 4.10. A sequence $\mathsf{s} = (s_0, s_1, s_2, ...)$ is defined by the initial conditions $s_0 = 1$, $s_1 = 2$, $s_2 = 1$, and the recurrence $s_n - s_{n-1} - 2s_{n-3} = 0$ for all $n \geq 3$. Obtain a formula for the generating series $S(x) = \sum_{n=0}^{\infty} s_n x^n$.

Since we have the information at hand, we might as well compute a few more values of s_n:

n	0	1	2	3	4	5	6	7
s_n	1	2	1	3	7	9	15	29

To get the generating series, Theorem 4.8 implies immediately that the denominator is $1 - x - 2x^3$. To obtain the numerator we apply the recurrence for small values of n, with the convention that $s_n = 0$ if $n < 0$.

$$s_n - s_{n-1} - 2s_{n-3} = \begin{cases} 1 & \text{if } n = 0, \\ 2 - 1 = 1 & \text{if } n = 1, \\ 1 - 2 = -1 & \text{if } n = 2. \end{cases}$$

Thus, the numerator is $1 + x - x^2$. The generating series is

$$S(x) = \frac{1 + x - x^2}{1 - x - 2x^3}.$$

Example 4.11. Let's revisit Example 2.22, concerning the set $\mathcal{Q}$ of all compositions in which each part is at least two, and the number of parts is even. We derived the generating series

$$Q(x) = \sum_{n=0}^{\infty} q_n x^n = \frac{1 - 2x + x^2}{1 - 2x + x^2 - x^4}.$$

From Theorem 4.8 we see immediately that

$$q_n - 2q_{n-1} + q_{n-2} - q_{n-4} = \begin{cases} 1 & \text{if } n = 0, \\ -2 & \text{if } n = 1, \\ 1 & \text{if } n = 2, \\ 0 & \text{if } n \geq 3, \end{cases}$$

with the convention that $q_n = 0$ if $n < 0$. We can inductively calculate the first several values of q_n:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
q_n	1	0	0	0	1	2	3	4	6	10	17	28	45	72	116

We have determined that $|\mathcal{Q}_{14}| = 116$, but we have not listed all these compositions individually. That is pretty cool, when you think about it.

4.3 Partial Fractions.

A *rational function* is a quotient of two polynomials $P(x)/Q(x)$. This is analogous to a rational number being a quotient of two integers.

Theorem 4.12 (Partial Fractions)**.** *Let $G(x) = P(x)/Q(x)$ be a rational function in which $\deg P < \deg Q$ and the constant term of $Q(x)$ is 1. Factor the denominator to obtain its inverse roots:*

$$Q(x) = (1 - \lambda_1 x)^{d_1}(1 - \lambda_2 x)^{d_2} \cdots (1 - \lambda_s x)^{d_s}$$

in which λ_1, ..., λ_s are distinct nonzero complex numbers and $d_1 + \cdots + d_s = d = \deg Q$. Then there are d complex numbers:

$$C_1^{(1)}, C_1^{(2)}, ..., C_1^{(d_1)}; \quad C_2^{(1)}, C_2^{(2)}, ..., C_2^{(d_2)}; \quad ... ; \quad C_s^{(1)}, C_s^{(2)}, ..., C_s^{(d_s)}$$

(which are uniquely determined) such that

$$G(x) = \frac{P(x)}{Q(x)} = \sum_{i=1}^{s} \sum_{j=1}^{d_s} \frac{C_i^{(j)}}{(1 - \lambda_i x)^j}.$$

Proof. Consider the set $\mathcal{V}_Q$ of all rational functions $P(x)/Q(x)$ in which Q is a fixed polynomial as in the statement of the theorem, and P is any polynomial of degree strictly less than $d = \deg Q$. It is easily seen that $\mathcal{V}_Q$ is a vector space over the complex numbers $\mathbb{C}$, since if $P(x)$ and $R(x)$ both have degree less than d and $\alpha \in \mathbb{C}$ then $P(x) + \alpha R(x)$ has degree less than d. It is also clear that the vectors

$$\frac{1}{Q(x)}, \quad \frac{x}{Q(x)}, \quad \frac{x^2}{Q(x)}, \quad \ldots, \quad \frac{x^{d-1}}{Q(x)}$$

span $\mathcal{V}_Q$ as a vector space over $\mathbb{C}$. Thus, the dimension of $\mathcal{V}_Q$ is at most d.

Now we claim that for every $1 \leq i \leq s$ and $1 \leq j \leq d_i$, the quotient $1/(1-\lambda_i x)^j$ is in $\mathcal{V}_Q$. This is because

$$\frac{1}{(1-\lambda_i x)^j} = \frac{(1-\lambda_i x)^{d_i - j} \prod_{h \neq i} (1-\lambda_h x)^{d_h}}{Q(x)}$$

and the numerator has degree $d - j \leq d - 1 < d$.

The essential point in the proof is that the set of vectors

$$\mathcal{B} \;=\; \left\{ \frac{1}{(1-\lambda_i x)^j} \;:\; 1 \leq i \leq s \;\text{ and }\; 1 \leq j \leq d_i \right\}$$

in $\mathcal{V}_Q$ is linearly independent. From this we can conclude that the dimension of $\mathcal{V}_Q$ is at least $d_1 + \cdots + d_s = d$. It then follows that $\dim \mathcal{V}_Q = d$ and that $\mathcal{B}$ is a basis for $\mathcal{V}_Q$. Therefore, every vector in $\mathcal{V}_Q$ can be written uniquely as a linear combination of vectors in $\mathcal{B}$. That is exactly what the Partial Fractions Theorem is claiming.

It remains only to show that $\mathcal{B}$ is a linearly independent set. Consider any linear combination of vectors in $\mathcal{B}$ which gives the zero vector:

$$\sum_{i=1}^{s} \sum_{j=1}^{d_s} \frac{C_i^{(j)}}{(1-\lambda_i x)^j} \;=\; 0. \tag{4.1}$$

We must show that $C_i^{(j)} = 0$ for all $1 \leq i \leq s$ and $1 \leq j \leq d_i$. Suppose not. Then there is some index $1 \leq p \leq s$ for which at least one of the coefficients $C_p^{(1)}$, $C_p^{(2)}$, ..., $C_p^{(d_p)}$ is not zero. Letting $C_p^{(t)} \neq 0$ be the one with the largest superscript, we also have $C_p^{(t+1)} = \cdots = C_p^{(d_p)} = 0$.

Now multiply equation (4.1) by $(1-\lambda_p x)^t$. Separating out the terms with $i = p$ and using the fact that $C_p^{(t+1)} = \cdots = C_p^{(d_p)} = 0$, we see that

$$\sum_{j=1}^{t} C_p^{(j)}(1-\lambda_p x)^{t-j} \; + \; \sum_{i\neq p}\sum_{j=1}^{d_i} C_i^{(j)} \frac{(1-\lambda_p x)^t}{(1-\lambda_i x)^j} \; = \; 0.$$

The LHS is a rational function of the variable x which does not have a pole at the point $x = 1/\lambda_p$, so we can substitute this value for x. But every term on the LHS has a factor of $(1-\lambda_p x)$ except for the term with $i = p$ and $j = t$. Thus, upon making the substitution $x = 1/\lambda_p$ this equation becomes

$$C_p^{(t)} \; = \; 0.$$

But this contradicts our choice of p and t. This contradiction shows that all the coefficients $C_i^{(j)}$ in equation (4.1) must be zero, and it follows that the set $\mathcal{B}$ is linearly independent.

Since $\mathcal{B}$ is a set of d linearly independent vectors in a vector space $\mathcal{V}_Q$ of dimension at most d, it follows that $\mathcal{B}$ is a basis for $\mathcal{V}_Q$, and the proof is complete. $\square$

Example 4.13. Let's re-examine the generating series

$$G(x) = \frac{2+5x}{1-3x^2-2x^3}$$

from Example 4.7. This satisfies the hypotheses of the Partial Fractions Theorem 4.12. The denominator $1-3x^2-2x^3$ vanishes when $x=-1$, so that $1+x$ is a factor. Some calculation shows that

$$1-3x^2-2x^3 = (1+x)(1-x-2x^2) = (1+x)^2(1-2x).$$

Thus, there are complex numbers A, B, C such that

$$\frac{2+5x}{1-3x^2-2x^3} = \frac{A}{1+x} + \frac{B}{(1+x)^2} + \frac{C}{1-2x}.$$

Now multiply by the denominator on the LHS.

$$2+5x = A(1+x)(1-2x) + B(1-2x) + C(1+x)^2.$$

This is an equality of polynomials, so it holds for any value of x.

- At $x = -1$ we find that $2 - 5 = B(1 + 2)$, so that $B = -1$.
- At $x = 1/2$ we find that $2 + 5/2 = C(3/2)^2$, so that $9/2 = C(9/4)$, so that $C = 2$.
- At $x = 0$ we find that $2 = A + B + C$, so that $A = 2 - B - C = 2 + 1 - 2 = 1$.

Therefore
$$\frac{2+5x}{1-3x^2-2x^3} = \frac{1}{1+x} - \frac{1}{(1+x)^2} + \frac{2}{1-2x}.$$

Now we can expand each of these terms using Binomial Series, and collect the results.

$$\begin{aligned} G(x) &= \sum_{n=0}^{\infty}(-x)^n - \sum_{n=0}^{\infty}\binom{n+1}{1}(-x)^n + 2\sum_{n=0}^{\infty}(2x)^n \\ &= \sum_{n=0}^{\infty}\left((-1)^n - (n+1)(-1)^n + 2\cdot 2^n\right)x^n \\ &= \sum_{n=0}^{\infty}\left(2^{n+1} - n(-1)^n\right)x^n. \end{aligned}$$

It follows that $g_n = [x^n]G(x) = 2^{n+1} + n(-1)^{n+1}$ for all $n \in \mathbb{N}$. This can be "reality checked" by comparison with the initial conditions $g_0 = 2$, $g_1 = 5$, and $g_2 = 6$, and the recurrence relation $g_n - 3g_{n-2} - 2g_{n-3} = 0$ for all $n \geq 3$ defining this sequence in Example 4.7. The first few values are

n	0	1	2	3	4	5	6
g_n	2	5	6	19	28	69	122

4.3.1 The Main Theorem.

Theorem 4.14. *Let* $\mathbf{g} = (g_0, g_1, g_2)$ *be a sequence of complex numbers, and let* $G(x) = \sum_{n=0}^{\infty} g_n x^n$ *be the corresponding generating series. Assume that the equivalent conditions of Theorem 4.8 hold, and that*

$$G(x) = R(x) + \frac{P(x)}{Q(x)}$$

for some polynomials $P(x)$, $Q(x)$, *and* $R(x)$ *with* $\deg P(x) < \deg Q(x)$ *and* $Q(0) = 1$. *Factor* $Q(x)$ *to obtain its inverse roots and their multiplicities:*

$$Q(x) = (1 - \lambda_1 x)^{d_1}(1 - \lambda_2 x)^{d_2} \cdots (1 - \lambda_s x)^{d_s}.$$

Then there are polynomials $p_i(n)$ *for* $1 \leq i \leq s$, *with* $\deg p_i(n) < d_i$, *such that for all* $n > \deg R(x)$,

$$g_n = p_1(n)\lambda_1^n + p_2(n)\lambda_2^n + \cdots + p_s(n)\lambda_s^n.$$

Proof. The conclusion of the theorem only concerns terms with $n > \deg R(x)$, so we can basically ignore the polynomial $R(x)$. In truth, all it is doing is getting in the way, and preventing the formula from holding for smaller values of n. So we are going to concentrate on the quotient $P(x)/Q(x)$, to which the Partial Fractions Theorem 4.12 applies.

Consider the factor $(1 - \lambda_i x)^{d_i}$ of the denominator $Q(x)$. In the partial fractions expansion of $P(x)/Q(x)$, this contributes

$$\frac{C_i^{(1)}}{1 - \lambda_i x} + \frac{C_i^{(2)}}{(1 - \lambda_i x)^2} + \cdots + \frac{C_i^{(d_i)}}{(1 - \lambda_i x)^{d_i}}.$$

Each term is a binomial series, and can be expanded accordingly:

$$\begin{aligned} \sum_{j=1}^{d_i} \frac{C_i^{(j)}}{(1 - \lambda_i x)^j} &= \sum_{j=1}^{d_i} C_i^{(j)} \sum_{n=0}^{\infty} \binom{n+j-1}{j-1} \lambda_i^n x^n \\ &= \sum_{n=0}^{\infty} \left(\sum_{j=1}^{d_i} C_i^{(j)} \binom{n+j-1}{j-1} \right) \lambda_i^n x^n. \end{aligned}$$

Notice that $\binom{n+j-1}{j-1}$ is a polynomial function of n of degree $j - 1$. It follows that

$$p_i(n) \;=\; \sum_{j=1}^{d_i} C_i^{(j)} \binom{n+j-1}{j-1}$$

is a polynomial function of n of degree at most $d_i - 1$. The contribution of the inverse root λ_i to the coefficient $g_n = [x^n]G(x)$ is thus $p_i(n)\lambda_i^n$. By the form of the partial fractions expansion we see that

$$g_n = p_1(n)\lambda_1^n \;+\; p_2(n)\lambda_2^n \;+\; \cdots \;+\; p_s(n)\lambda_s^n,$$

completing the proof. □

The converse of Theorem 4.14 also holds – see Exercise 4.11.

One can use Theorem 4.14 to go straight from a recurrence relation to a formula for its entries, without doing Partial Fractions explicitly. Here is an example of this kind of calculation.

Example 4.15. *A sequence* $\mathbf{h}$ *of integers is given by the initial conditions* $h_0 = 1$, $h_1 = 1$, $h_2 = 0$, $h_3 = 2$, $h_4 = -4$, $h_5 = 3$, *and the recurrence* $h_n - 3h_{n-1} + 4h_{n-3} = 0$ *for all* $n \geq 6$. *Obtain a formula for* h_n *as a function of* n.

From Theorem 4.8 we see that the denominator of the generating series $H(x) = h_n x^n$ is $1 - 3x + 4x^3$. This vanishes at $x = -1$, so $1 + x$ is a factor. After some work, we obtain

$$1 - 3x + 4x^3 = (1 + x)(1 - 4x + 4x^2) = (1 - 2x)^2(1 + x).$$

Theorem 4.14 implies that there are constants A, B, C such that for sufficiently large n, $h_n = (A + Bn)2^n + C(-1)^n$. To determine these constants we need to take data from the sequence $\mathbf{h}$ from a point later than the degree of the polynomial $R(x)$ appearing in Theorem 4.14. From Theorem 4.8, in this case the degree of the numerator of the generating series $H(x)$ is no more than five, since the general case of the recurrence holds for $n \geq 6$. Writing

$$H(x) = R(x) + \frac{P(x)}{1 - 3x + 4x^3} = \frac{(1 - 3x + 4x^3)R(x) + P(x)}{1 - 3x + 4x^3},$$

it follows that the degree of $R(x)$ is at most two. So we can fit the form $h_n = (A + Bn)2^n + C(-1)^n$ to the data $h_3 = 2$, $h_4 = -4$, and $h_5 = 3$.

This gives three equations in three unknowns – a standard linear algebra problem.

$$\begin{aligned} h_3 = 2 &= (A+3B)8 - C = 8A + 24B - C \\ h_4 = -4 &= (A+4B)16 + C = 16A + 64B + C \\ h_5 = 3 &= (A+5B)32 - C = 32A + 160B - C \end{aligned}$$

In this case, it is a rather unpleasant linear algebra problem. Sparing you the details, the solution is $A = -5/16$, $B = 1/16$, $C = -3$, and so

$$h_n = (n-5)2^{n-4} - 3(-1)^n$$

for all $n \geq 3$. The values for h_n with $0 \leq n \leq 2$ are given in the initial conditions.

4.3.2 Inhomogeneous Linear Recurrence Relations.

Example 4.16. Define a sequence of integers $\mathbf{g} = (g_0, g_1, g_2, ...)$ by the initial conditions $g_0 = 1$ and $g_1 = 2$, and the recurrence relation

$$g_n = g_{n-1} + 2\,g_{n-2} - n + 1$$

for all $n \geq 2$. This clearly determines the sequence g inductively:

n	0	1	2	3	4	5	6	7	8
g_n	1	2	3	5	8	14	25	47	90

What is g_n as a function of $n \in \mathbb{N}$?

We solve Example 4.16 by generalizing the method above just a little bit. First, write the recurrence in the form

$$g_n - g_{n-1} - 2\,g_{n-2} = -n + 1$$

for all $n \geq 2$. This is an **inhomogeneous** linear recurrence relation according to Definition 4.6. But we can proceed just as before. We begin by obtaining a formula for the generating series $G(x) = \sum_{n=0}^{\infty} g_n x^n$. Multiply both sides by x^n and sum over all $n \geq 2$. On the RHS we obtain

$$\sum_{n=2}^{\infty} (-n+1)\, x^n = \sum_{j=0}^{\infty} (-(j+2)+1)\, x^{j+2} \;=\; -x^2 \sum_{j=0}^{\infty} (j+1)\, x^j$$
$$= \frac{-x^2}{(1-x)^2}.$$

On the LHS we obtain

$$\begin{aligned}
&\sum_{n=2}^{\infty} (g_n \;-\; g_{n-1} \;-\; 2\, g_{n-2})\, x^n \\
&= \sum_{n=2}^{\infty} g_n\, x^n \;-\; \sum_{n=2}^{\infty} g_{n-1}\, x^n -\; 2 \sum_{n=2}^{\infty} g_{n-2}\, x^n \\
&= (G(x) - g_0 - g_1 x) \;-\; x \sum_{j=1}^{\infty} g_j\, x^j \;-\; 2\, x^2 \sum_{k=0}^{\infty} g_k\, x^k \\
&= (G(x) - 1 - 2x) - x\,(G(x) - 1) - 2x^2\, G(x) \\
&= (1 - x - 2x^2)G(x) - 1 - x.
\end{aligned}$$

Equating the LHS and the RHS yields

$$(1 - x - 2x^2)G(x) \;=\; 1 + x - \frac{x^2}{(1-x)^2} = \frac{(1+x)(1-x)^2 - x^2}{(1-x)^2}$$

Noting that $1 - x - 2x^2 = (1+x)(1-2x)$, we obtain the formula

$$G(x) \;=\; \frac{1 - x - 2x^2 + x^3}{(1+x)(1-x)^2(1-2x)}.$$

This is a rational function, and so the Main Theorem 4.14 applies. There are complex numbers A, B, C, D such that $g_n = A(-1)^n + (B + Cn) + D2^n$ for all $n \in \mathbb{N}$. For $n \in \{0, 1, 2, 3\}$ this yields the system of linear equations

$$\begin{array}{rrrrl}
A & +B & & +D & = 1 \\
-A & +B & +C & +2D & = 2 \\
A & +B & +2C & +4D & = 3 \\
-A & +B & +3C & +8D & = 5
\end{array}$$

Solving this system yields $A = -1/12$, $B = 3/4$, $C = 1/2$, $D = 1/3$, so that

$$g_n \;=\; \frac{1}{12}\left(2^{n+2} + (6n+9) - (-1)^n\right)$$

for all $n \in \mathbb{N}$.

Example 4.17. The denominator of $G(x)$ in the above example is $1-3x+x^2+3x^3-2x^4$. From Theorem 4.8 we see that the sequence $\mathbf{g}$ satisfies the **homogeneous** linear recurrence relation and initial conditions given by

$$g_n - 3g_{n-1} + g_{n-2} + 3g_{n-3} - 2g_{n-4} = \begin{cases} 1 & \text{if } n = 0, \\ -1 & \text{if } n = 1, \\ -2 & \text{if } n = 2, \\ 1 & \text{if } n = 3, \\ 0 & \text{if } n \geq 4. \end{cases}$$

This agrees with the results above.

Examples 4.16 and 4.17 illustrate a general fact: if the generating series of the RHS in an **inhomogeneous** linear recurrence relation is a rational function, then the generating series for the entries of the sequence is also a rational function. Thus, the sequence in fact satisfies a **homogeneous** linear recurrence relation, so we are actually back in the case we have already considered. Proving this in general is the main point of this subsection.

The following terminology is not standard but will be convenient. A function $q : \mathbb{N} \to \mathbb{C}$ is *polyexp* if there are polynomial functions $q_i(n)$ and complex numbers $\beta_i \in \mathbb{C}$ for $1 \leq i \leq t$ such that

$$q(n) \;=\; q_1(n)\beta_1^n \;+\; q_2(n)\beta_2^n \;+\; \cdots \;+\; q_t(n)\beta_t^n \tag{4.2}$$

for all $n \in \mathbb{N}$. For example, $\cos(n\theta) = (\mathrm{e}^{\mathrm{i}\theta n} + \mathrm{e}^{-\mathrm{i}\theta n})/2$ is polyexp, but $\sqrt{n}$ is not. More generally, the function $q : \mathbb{N} \to \mathbb{C}$ is *eventually polyexp* if there is an integer $M \in \mathbb{N}$ such that equation (4.2) holds for all $n \geq M$. The Main Theorem 4.14 thus states that if $G(x) = \sum_{n=0}^{\infty} g_n x^n$ is a rational function then g_n is an eventually polyexp function of n. (Exercise 4.11 is the converse implication.)

Theorem 4.18. *Let $\mathbf{g} = (g_0, g_1, g_2, \ldots)$ be a sequence of complex numbers. The following are equivalent.*

(a) The sequence $\mathbf{g}$ satisfies a homogeneous linear recurrence relation (with initial conditions).

(b) The sequence $\mathbf{g}$ satisfies a possibly inhomogeneous linear recurrence relation (with initial conditions) in which the RHS is an eventually

polyexp function.

(c) *The generating series* $G(x) = \sum_{n=0} g_n x^n$ *is a rational function (a quotient of polynomials in* x*).*

(d) *The sequence* $\mathbf{g} = (g_0, g_1, g_2, ...)$ *is an eventually polyexp function.*

Proof. Theorem 4.8 shows that conditions (a) and (c) are equivalent. Theorem 4.14 shows that (c) implies (d). That (d) implies (c) is left as Exercise 4.11. It is clear that (a) implies (b). All that remains is to show that (b) implies (c).

Thus, assume that g satisfies the linear recurrence relation

$$g_n + a_1 g_{n-1} + \cdots + a_d g_d = q(n)$$

for all $n \geq N$, with initial conditions g_0, g_1, ..., g_{N-1}, in which $q : \mathbb{N} \to \mathbb{C}$ is an eventually polyexp function as in equation (4.2) for all $n \geq M$.

UNDER CONSTRUCTION

□

4.4 Quadratic Recurrence Relations.

Let $\mathbf{g} = (g_0, g_1, g_2, ...)$ be a sequence of numbers with generating series $G(x) = \sum_{n=0}^{\infty} g_n x^n$. In Theorem 4.8 we saw that g satisfies a homogeneous linear recurrence relation (with initial conditions) if and only if $G(x) = P(x)/Q(x)$ is a rational function. Rewriting this as $Q(x)G(x) - P(x) = 0$, we see that $G(x)$ is a solution to a linear equation: $QG - P = 0$.

Definition 4.19. The sequence g satisfies a *quadratic recurrence* if its generating series $G(x)$ satisfies a quadratic equation:

$$A(x)G(x)^2 + B(x)G(x) + C(x) = 0.$$

Here, the coefficients $A(x)$, $B(x)$, and $C(x)$ are power series in x.

There are two solutions to the equation in Definition 4.19, and they can be found using the Quadratic Formula:

$$\left.\begin{matrix} G_+(x) \\ G_-(x) \end{matrix}\right\} = \frac{-B(x) \pm \sqrt{B(x)^2 - 4\,A(x)\,C(x)}}{2\,A(x)}.$$

Rigorous justification for this kind of algebra with power series is discussed in detail in CO 330. If $G(x)$ is a generating series for some combinatorial objects then it has only nonnegative coefficients and nonnegative exponents. This can be used to decide which case of the $\pm$ sign to take. In general, only one of $G_+(x)$ or $G_-(x)$ is the correct generating series.

4.4.1 The general binomial series.

In Section 2.1 we saw the Binomial Theorem and The Binomial Series with negative integer exponents. That is, for a natural number $n \in \mathbb{N}$,

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

and for a positive integer $t \geq 1$,

$$\frac{1}{(1-x)^t} = \sum_{n=0}^{\infty} \binom{n+t-1}{t-1} x^n.$$

These are two special cases of the general *binomial series*.

Definition 4.20. For any complex number $\alpha \in \mathbb{C}$ and nonnegative integer $k \in \mathbb{N}$, the *k-th binomial coefficient of α* is

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

This binomial coefficient is a polynomial function of α of degree k.

Theorem 4.21 (The Binomial Series)**.** *For any complex number $\alpha \in \mathbb{C}$,*

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

Sketch of proof. We can think of $(1+x)^\alpha$ as a function of a complex variable x. The only possible singularity is that 0^α might not be well-defined – this can happen only for $x = -1$. Therefore, $(1+x)^\alpha$ is analytic in the disc $|x| < 1$, and so it has a Taylor series expansion. By Taylor's Theorem, the coefficient of x^k in this Taylor series expansion is

$$\frac{1}{k!}\,\frac{\mathrm{d}^k}{\mathrm{d}x^k}(1+x)^\alpha\bigg|_{x=0} = \frac{1}{k!}(\alpha)(\alpha-1)\cdots(\alpha-k+1)(1+x)^{\alpha-k}\bigg|_{x=0} = \binom{\alpha}{k}.$$

This proves the stated formula. □

We will use the following special case.

Proposition 4.22. $\sqrt{1-4x} \;=\; 1-2\sum_{k=1}^{\infty}\frac{1}{k}\binom{2k-2}{k-1}x^k.$

Proof. By Theorem 4.21, $\sqrt{1-4x} \;=\; \sum_{k=0}^{\infty}(-1)^k 4^k\binom{1/2}{k}x^k.$

For $k=0$, the coefficient of x^0 is $(-1)^0 4^0\binom{1/2}{0}=1$.
For $k\geq 1$, we can calculate as follows.

$$\begin{aligned}
(-1)^k 4^k\binom{1/2}{k} &= (-1)^k 4^k\,\frac{1}{k!}\,(1/2)(-1/2)(-3/2)\cdots(1/2-k+1)\\
&= -4^k\,\frac{1}{k!}\,(1/2)(1/2)(3/2)\cdots(k-3/2)\\
&= -2^k\,\frac{1}{k!}\,(1)(1)(3)(5)\cdots(2k-3)\\
&= -\,\frac{2}{k}\cdot\frac{(1)(3)\cdots(2k-3)}{(k-1)!}\cdot\frac{(2)(4)\cdots(2k-2)}{(k-1)!}\\
&= -\,\frac{2}{k}\binom{2k-2}{k-1}.
\end{aligned}$$

(Where did we use the fact that $k\geq 1$ in this calculation?) □

4.4.2 Catalan numbers.

The numbers $C_n=\frac{1}{n+1}\binom{2n}{n}$ are called *Catalan numbers*. The first few Catalan numbers are shown here:

n	0	1	2	3	4	5	6	7	8	9	10
C_n	1	1	2	5	14	42	132	429	1430	4862	16796

Catalan numbers occur surprisingly often in answers to counting problems.

Example 4.23 (Well-Formed Parenthesizations.)**.**
A *well-formed parenthesization* (WFP) is a sequence of n opening parentheses and n closing parentheses which "match together" using the usual rules for grouping parentheses. The *size* of a WFP is the number of opening parentheses in it. Here are all the WFPs of size 3:

()()() **()(())** **(())()** **(()())** **((()))**

And here are all the WFPs of size 4:

()()()() **(())()()** **(()())()** **()()(())**
(())(()) **(()(()))** **()(())()**
((()))() **((())())** **()(()())** **(()()())**
((()())) **()((()))** **(((())))**

We determine the number w_n of WFPs of size n, for all $n \in \mathbb{N}$. Let

$$W(x) = \sum_{n=0}^{\infty} w_n x^n$$

be the generating series for WFPs with respect to size.

We can obtain a quadratic recurrence for $W(x)$, as follows. The empty sequence ε contributes $x^0 = 1$ to the generating series $W(x)$. Any other WFP γ begins with an opening parenthesis. There is exactly one closing parenthesis which matches to the beginning parenthesis. That is, $\gamma = \mathbf{(}\,\alpha\,\mathbf{)}\,\beta$ for some other sequences α and β. Note that α or β might be empty. Because of the way parentheses are matched to each other, both α and β are in fact WFPs themselves. The total number of opening parentheses in γ is $n(\gamma) = 1 + n(\alpha) + n(\beta)$. Conversely, given any WFPs α and β we can always form a new nonempty WFP: $\mathbf{(}\,\alpha\,\mathbf{)}\,\beta$.

Writing a 0 instead of a **(**, and a 1 instead of a **)**, a WFP can be thought of as a binary string in $\{0,1\}^*$. Let $\mathcal{W}$ be the set of all binary strings corresponding to WFPs. The previous paragraph justifies the claim that the recursive decomposition

$$\mathcal{W} = \varepsilon \smile 0\,\mathcal{W}\,1\,\mathcal{W}$$

is unambiguous. This allows us to calculate as follows:

$$
\begin{aligned}
W(x) = \sum_{\gamma \in \mathcal{W}} x^{n(\gamma)} &= x^{n(\varepsilon)} + \sum_{\gamma \in \mathcal{W} \smallsetminus \{\varepsilon\}} x^{n(\gamma)} \\
&= 1 + \sum_{\alpha \in \mathcal{W}} \sum_{\beta \in W} x^{1+n(\alpha)+n(\beta)} \\
&= 1 + x \left(\sum_{\alpha \in \mathcal{W}} x^{n(\alpha)} \right) \left(\sum_{\beta \in \mathcal{W}} x^{n(\beta)} \right) \\
&= 1 + xW(x)^2.
\end{aligned}
$$

Now we can solve this equation $xW(x)^2 - W(x) + 1 = 0$ using the Quadratic Formula:

$$
\left.\begin{array}{l} W_+(x) \\ W_-(x) \end{array}\right\} = \frac{1 \pm \sqrt{1-4x}}{2x}.
$$

Proposition 4.22 gives the power series for $\sqrt{1-4x}$, so that

$$
\frac{1 \pm \sqrt{1-4x}}{2x} = \frac{1}{2x} \pm \frac{1}{2x} \left(1 - 2 \sum_{n=0}^{\infty} \frac{1}{n+1} \binom{2n}{n} x^{n+1} \right).
$$

To get nonnegative coefficients, and to cancel the term $1/2x$, we need to take the minus sign from the $\pm$. The result is

$$
W(x) = \sum_{n=0}^{\infty} \frac{1}{n+1} \binom{2n}{n} x^n.
$$

Thus, the number of WFPs of size n is the n-th Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$, for each $n \in \mathbb{N}$.

Since the generating series for the set $\mathcal{W}$ is $W(x) = (1 - \sqrt{1-4x})/2x$, which is not a rational function, it follows from Theorem 3.13 that the set of WFPs is not a rational language.

4.5 Exercises.

Exercise 4.1. For each of the sets of compositions from Exercise 2.15, do the following.

- Derive a recurrence relation and initial conditions for the coefficients of the corresponding generating series $G(x) = \sum_{n=0}^{\infty} g_n x^n$.
- Calculate the coefficients g_0, g_1, ... up to g_9.

Exercise 4.2. Let $\mathcal{K}$ be the set of compositions $\gamma = (c_1, c_2, ..., c_k)$ with at least one part, and such that the first part is odd. Let $K(x)$ be the generating series for $\mathcal{K}$ with respect to size.

(a) Show that
$$K(x) = \frac{x}{(1+x)(1-2x)}.$$

(b) Use part (a) to show that, among all 2^{n-1} compositions of size $n \geq 1$, the fraction of these compositions in the set $\mathcal{K}$ is
$$\frac{2}{3} + \frac{1}{3}\left(\frac{-1}{2}\right)^{n-1}.$$

Exercise 4.3. Consider the power series
$$\sum_{n=0}^{\infty} c_n x^n = \frac{1-2x^2}{1-5x+8x^2-4x^3} = 1 + 5x + 15x^2 + 39x^3 + \cdots$$

(a) Give a linear recurrence relation that (together with the initial conditions above) determines the sequence of coefficients $(c_n : n \geq 0)$ uniquely.

(b) Derive a formula for c_n as a function of $n \geq 0$.
[Hint: $1 - 5x + 8x^2 - 4x^3 = (1-x)(1-4x+4x^2)$.]

Exercise 4.4. Consider the power series
$$A(x) = \sum_{n=0}^{\infty} a_n x^n = \frac{x+7x^2}{1-3x^2-2x^3}.$$

(a) Write down a linear recurrence relation and enough initial conditions to determine the sequence $(a_n : n \in \mathbb{N})$ uniquely.
(b) Given that $1 - 3x^2 - 2x^3 = (1-2x)(1+x)^2$, obtain a formula for a_n as a function of $n \in \mathbb{N}$.

Exercise 4.5. Consider the power series

$$\sum_{n=0}^{\infty} c_n x^n = \frac{3 - 11x + 11x^2}{1 - 4x + 5x^2 - 2x^3} = 3 + x + 0x^2 + x^3 + 6x^4 + 19x^5 + \cdots$$

(a) Give a linear recurrence relation that (together with the initial conditions above) determines the sequence of coefficients $(c_n : n \geq 0)$ uniquely.
(b) Derive a formula for c_n as a function of $n \geq 0$.

Exercise 4.6. A sequence of integers is determined by the initial conditions $g_0 = 1$, $g_1 = 2$, $g_2 = 3$, and the recurrence relation $g_n = 2g_{n-1} - g_{n-2} + 2g_{n-3}$ for all $n \geq 3$.

(a) Obtain a rational function formula for the generating series

$$G(x) = \sum_{n=0}^{\infty} g_n x^n = 1 + 2x + 3x^2 + 6x^3 + 13x^4 + 26x^5 + 51x^6 + \cdots .$$

(b) Obtain a formula for the coefficient g_n as a function of $n \in \mathbb{N}$.

Exercise 4.7. Define a sequence of numbers $(c_n : n \in \mathbb{N})$ by the initial conditions $c_0 = 1$, $c_1 = 2$, and $c_2 = 3$, and the recurrence relation $c_n = -c_{n-1} + 2c_{n-2} + 2c_{n-3}$ for all $n \geq 3$.

(a) Obtain an algebraic formula for the rational function

$$C(x) = \sum_{n=0}^{\infty} c_n x^n = 1 + 2x + 3x^3 + 3x^3 + 7x^4 + 5x^5 + \cdots .$$

(b) Obtain a formula for c_n as a function of $n \in \mathbb{N}$.

Exercise 4.8.

(a) Obtain a formula for the coefficients of the rational function

$$B(x) = \sum_{n=0}^{\infty} b_n x^n = \frac{1+3x-x^2}{1-3x^2-2x^3}.$$

(b) Derive a recurrence relation and use it to check your answer.

Exercise 4.9. Define a sequence $(h_n : n \in \mathbb{N})$ by the initial conditions $h_0 = 1$, $h_1 = 2$, $h_2 = 0$, $h_3 = 5$, and the recurrence relation $h_n = -2h_{n-1} + h_{n-2} + 4h_{n-3} + 2h_{n-4}$ for all $n \geq 4$.

(a) Obtain an algebraic formula for the rational function

$$H(x) = \sum_{n=0}^{\infty} h_n x^n = 1 + 2x + 0x^2 + 5x^3 + 0x^4 + 9x^5 + \cdots .$$

(b) Obtain a formula for h_n as a function of $n \in \mathbb{N}$.

Exercise 4.10.

(a) Find rational numbers A, B, C such that for all $n \in \mathbb{N}$,

$$n^2 = A\binom{n+2}{2} + B(n+1) + C.$$

(b) Write $\sum_{n=0}^{\infty} n^2 x^n$ as a quotient of polynomials.
(c) Write $\sum_{n=0}^{\infty} n^3 x^n$ as a quotient of polynomials.
(d) For each $d \in \mathbb{N}$, let $F_d(x) = \sum_{n=0}^{\infty} n^d x^n$.
Show that $F_0(x) = 1/(1-x)$ and for all $d \geq 1$,

$$F_d(x) = x \frac{\mathrm{d}}{\mathrm{d}x} F_{d-1}(x).$$

(e) Let $F_d(x) = P_d(x)/(1-x)^{1+d}$. Derive a recurrence relation for the polynomials $P_d(x)$.

Exercise 4.11. Show that the converse of Theorem 4.14 holds. That is, assume that

$$g_n = p_1(n)\lambda_1^n + p_2(n)\lambda_2^n + \cdots + p_s(n)\lambda_s^n$$

for all $n \geq N$, in which $p_i(n)$ is a polynomial of degree strictly less than d_i and the λ_i are distinct nonzero complex numbers, for $1 \leq i \leq s$. Let

$$Q(x) = (1 - \lambda_1 x)^{d_1}(1 - \lambda_2 x)^{d_2} \cdots (1 - \lambda_s x)^{d_s}.$$

Then

$$G(x) = \sum_{n=0}^{\infty} g_n x^n = R(x) + \frac{P(x)}{Q(x)}$$

in which $P(x)$ and $R(x)$ are polynomials, and $\deg P(x) < \deg Q(x)$ and $\deg R(x) < N$.

Exercise 4.12.

(a) Show that $\dfrac{1}{\sqrt{1-4x}} = \displaystyle\sum_{k=0}^{\infty} \binom{2k}{k} x^k$.

(b) Deduce that for all $n \geq \mathbb{N}$, $\displaystyle\sum_{j=0}^{n} \binom{2j}{j}\binom{2n-2j}{n-j} = 4^n$.

(c)* Can you think of a combinatorial proof of part (b)?