Chapter 6

Sequences and Recurrence Relations

In this chapter we will study sequences of numbers that are built recursively. Technically, a **sequence** (of real numbers) is a function a from $\mathbb{N}$ to $\mathbb{R}$. If $n \in \mathbb{N}$, it is common to write $a_n := a(n)$. We refer to a_n as the nth **term** of the sequence. We will abuse notation and associate a sequence with its list of outputs, namely:

$$\boxed{(a_n)_{n=1}^{\infty} := (a_1, a_2, \ldots)},$$

which we may abbreviate as $\boxed{(a_n)}$. Sometimes we may start our sequences at $n = 0$ as opposed to $n = 1$. That is, we may allow the domain of a sequence to be $\mathbb{N} \cup \{0\}$.

Example 6.1. Define $a : \mathbb{N} \to \mathbb{R}$ via $a_n = \frac{1}{2^n}$. Then we have

$$a = \left(\frac{1}{2}, \frac{1}{4}, \ldots\right) = \left(\frac{1}{2^n}\right)_{n=1}^{\infty}.$$

It is important to point out that not every sequence has a description in terms of an algebraic formula. For example, we could form a sequence out of the digits to the right of the decimal in the decimal expansion of π, namely the nth term of the sequence is the nth digit to the right of the decimal. But then there is no nice algebraic formula for describing the nth term of this sequence.

Loosely speaking, a sequence of numbers is defined recursively if the nth term of the sequence is defined in terms of "earlier" terms of the sequence. We have already encountered one famous example of a sequence being defined recursively, namely the Fibonacci sequence (f_n), which we defined by $f_1 = 1$, $f_2 = 1$, and $f_n = f_{n-1} + f_{n-2}$ for $n \geq 3$. The equation $f_n = f_{n-1} + f_{n-2}$ is the **recurrence relation** while $f_1 = 1$ and $f_2 = 1$ are the **initial conditions**. It is important to emphasize that we cannot define the Fibonacci number using only the recurrence relation since otherwise, we would not be able to "get started" with the recurrence.

We have also encountered a few recurrence relations of a different flavor that arise out of two-dimensional arrays of numbers. For example:

(a) Number of k-permutations of $[n]$: For $1 \leq k \leq n$,

$$P(n, k) = P(n-1, k) + kP(n-1, k-1).$$

(b) Number of k-subsets of $[n]$: For $1 \leq k \leq n$,

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

(c) Number of set partitions of $[n]$ with k blocks: For $1 \leq k \leq n$,

$$\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k \left\{ {n-1 \atop k} \right\}.$$

Notice that each of the descriptions above are not sufficient without also providing a way to "get started". For the two-dimensional case, the initial conditions are often called **boundary conditions**. For the rest of this chapter, we will focus on one-dimensional sequences.

Here is an important general principle.

Theorem 6.2. If two sequences satisfy the same recurrence relation and initial conditions, then the two sequences must be equal.

Problem 6.3. Recall that a **composition** of n with k parts is an ordered list of k positive integers whose sum is n, denoted $\alpha = (\alpha_1, \ldots, \alpha_k)$. We say that α_i is the ith part.

(a) How many compositions of n have only odd parts?

(b) How many compositions of n have parts of size 1 and 2 only?

Problem 6.4. Prove that $f_{n+1} = \sum_{k \geq 0} \binom{n-k}{k}$ by utilizing one of the parts from Problem 6.3. What does this identity tell us about Pascal's Triangle?

Problem 6.5. For each of the following recursively defined sequences, generate the first few terms. If possible, find an explicit formula for the terms of the sequence.

(a) $a_1 = 2$, $a_n = a_{n-1} + 7$ for $n \geq 2$.

(b) $a_0 = 1$, $a_n = 2a_{n-1}$ for $n \geq 1$.

(c) $a_0 = 0$, $a_n = na_{n-1}$ for $n \geq 1$.

(d) $a_0 = 0$, $a_n = a_{n-1} + n$ for $n \geq 1$.

(e) $a_0 = 0$, $a_n = a_{n-1} + \sum_{i=0}^{n}(i + n)$ for $n \geq 1$.

By **solving** a recurrence relation together with its initial conditions we mean finding an explicit expression (sometimes called a **closed form**) for an arbitrary term a_n as a function of n (but no earlier terms of the sequence). The explicit expression for a_n is called the **solution** of the recurrence relation. For example, each time we found an explicit formula for the nth term of a sequence in the previous problem, we were solving the recurrence relation and the corresponding expression we found is the solution. By the **general solution** of a recurrence relation, we mean the set of its solutions given any initial conditions.

Problem 6.6. Find the general solution for $a_n = 2a_{n-1}$ if the first term of the sequence is a_0. What if the sequence starts at a_1?

It's important to point out that finding a solution to a recurrence relation can be quite complicated, maybe even impossible! However, verifying whether a proposed solution is correct or not is straightforward.

Problem 6.7. Consider the recurrence relation $a_n = a_{n-1}+6a_{n-2}$. Is $a_n = (-2)^n$ a solution? How about $a_n = 3^n$? How about $5(-2)^n + 7 \cdot 3^n$?

We now turn our attention to two special classes of recurrence relations. An **arithmetic progression** is a recurrence relation in which the first term a_0 (or a_1) and a **common difference** d are given and the corresponding recurrence relation is

$$\boxed{a_n = a_{n-1} + d.}$$

A **geometric progression** is a recurrence relation in which the first term a_0 (or a_1) and **common ratio** r are given and the corresponding recurrence relation is

$$\boxed{a_n = r \cdot a_{n-1}.}$$

Problem 6.8. Compute the first few terms of each of the following and find a solution.

(a) Arithmetic progression with $a_0 = 3$ and $d = 2$.

(b) Geometric progression with $a_0 = 3$ and $r = 2$.

Problem 6.9. Conjecture a solution to an arithmetic progression with first term a_0 and common difference d. Can you prove that your conjecture is correct?

Problem 6.10. Conjecture a solution to a geometric progression with first term a_0 and common ratio r. Can you prove that your conjecture is correct?

Problem 6.11. Recall that the **triangular numbers** are defined via $t_n := 1 + 2 + \cdots + n$. The first few terms of this sequence are 1, 3, 6, 10, 15.

(a) Express the triangular numbers using a recurrence relation and initial condition.

(b) Is this sequence an arithmetic progression? Geometric progression?

(c) Notice that the sequence of triangular numbers is a sequence of partial sums of the arithmetic sequence $1, 2, 3, \ldots$. What happens if we add the partial sum expression for t_n to a second copy of t_n written in reverse? Can you recover the nice closed form for t_n we are already familiar with?

We can generalize the technique above for any sequence that is given by partial sums of an arithmetic sequence.

Problem 6.12. Define $a_n = 6 + 10 + 14 + \cdots + (4n-2) = \sum_{i=2}^{n}(4i-2)$. Find a closed form for a_n.

What about sequences that are partial sums of geometric progressions? In this case, it turns out that we can multiply by the common ratio, shift, and subtract.

Problem 6.13. Define $a_n = 3^0 + 3^1 + \cdots + 3^n = \displaystyle\sum_{i=0}^{n} 3^i$. Find a closed form for a_n.

A **linear constant-coefficient** recurrence relation of **order** r has the form

$$\boxed{a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_r a_{n-r} + f(n),}$$

where $c_1, c_2, \ldots, c_r$ are real numbers with $c_r \neq 0$. Such a recurrence relation is said to be **homogeneous** if $f(n) = 0$, so that it can be written as

$$\boxed{a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_r a_{n-r},}$$

and is **non-homogeneous** otherwise.

Notice that every arithmetic progression and every geometric progression is a first order linear constant-coefficient recurrence relation. In particular, each geometric progression is homogeneous while each arithmetic progression is non-homogeneous.

Problem 6.14. Determine which of the following are linear constant-coefficient recurrence relations. For those that are, which are homogeneous and which are non-homogeneous?

(a) $a_n = n a_{n-1}$

(b) $a_n = a_{n-1} + d$

(c) $a_n = c a_{n-1}$

(d) $a_n = a_{n-1} + a_{n-2}$

(e) $a_n = \sum_{i=0}^{n-1} a_i a_{n-1-i}$

(f) $a_n = a_{n-1} - 4a_{n-2} + 7a_{n-3}$

(g) $a_n = a_{n-1}^2 + 7a_{n-2} + 2a_{n-6}$

Problem 6.15. Solve the first-order linear constant-coefficient non-homogeneous recurrence relation $a_n = 3a_{n-1} + 2$ with initial condition $a_0 = 1$.

Unfortunately, the technique of the previous example is difficult to generalize to higher orders.

The next theorem characterizes the phenomenon that we witnessed in Problem 6.7. This theorem can be proved by direct substitution and some algebraic manipulation.

Theorem 6.16 (Principle of Superposition)**.** If $s_1(n), \ldots, s_k(n)$ are solutions to the linear constant-coefficient homogeneous recurrence relation

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_r a_{n-r}$$

and $\alpha_1, \ldots, \alpha_k$ are real numbers, then the linear combination $\alpha_1 s_1(n) + \cdots + \alpha_k s_k(n)$ is also a solution.

We now focus on solving second-order linear constant-coefficient homogeneous recurrence relations. Given the second-order linear constant-coefficient homogeneous recurrence relation

$$a_n = c_1 a_{n-1} + c_2 a_{n-2},$$

its corresponding **characteristic equation** is defined via

$$\boxed{x^2 - c_1 x - c_2 = 0.}$$

The solutions of the characteristic equation are called **characteristic roots**.

Example 6.17. The characteristic equation for the Fibonacci relation $f_n = f_{n-1} + f_{n-2}$ is $x^2 - x - 1 = 0$, which has characteristic roots $x = \frac{1 \pm \sqrt{5}}{2}$. Note that the characteristic root $\frac{1+\sqrt{5}}{2} \approx 1.618$ is the well-known **golden ratio**.

We will utilize the following remarkable theorem without proving it.

Theorem 6.18. If r_1 and r_2 are two *distinct* characteristic roots (i.e., $r_1 \neq r_2$) of the characteristic equation for $a_n = c_1 a_{n-1} + c_2 a_{n-2}$, then the solution to the recurrence relation is

$$\boxed{a_n = a r_1^n + b r_2^n,}$$

where a and b are constants determined by the initial conditions.

Problem 6.19. Solve $a_n = a_{n-1} + a_{n-2}$ with initial conditions $a_0 = 0$ and $a_1 = 1$.

Problem 6.20. Use the previous problem to find a solution to the Fibonacci sequence given by $f_1 = 1$, $f_2 = 1$, and $f_n = f_{n-1} + f_{n-2}$. The closed form we just obtained for f_n is called **Binet's formula**.

Although we will not consider examples more complicated than these, this characteristic root technique can be applied to much more complicated recurrence relations.

We now turn our attention to one of my favorite sequences, which is defined by a recurrence relation of a different flavor. The **Catalan numbers** are defined via $c_0 = 1$ and

$$\boxed{c_n = \sum_{i=0} c_i c_{n-1-i}}$$

for $n \geq 1$. The equation above is called the **Catalan recurrence**. Using the initial condition and the Catalan recurrence, we can generate the first several terms of the Catalan sequence:

$$1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786$$

There are hundreds of interesting combinatorial objects counted by the Catalan numbers! Let's explore a few.

Problem 6.21. A **Dyck path** of length $2n$ is a lattice path from $(0,0)$ to (n,n) that takes n steps East from (i,j) to $(i+1,j)$ and n steps North from (i,j) to $(i,j+1)$ such that all points on the path satisfy $i \leq j$. This sound more complicated that it really is. You can think of a Dyck path as one of our paths to get coffee that starts at $(0,0)$ and ends at (n,n) but never drops below the line $y = x$. Let $\text{Dyck}(n)$ denote set of all Dyck paths of length $2n$ and let $d_n := |\,\text{Dyck}(n)|$. We define $d_0 := 1$ for convenience. *Important:* Unfortunately, we also used d_n to denote the number of derangements of n. This problem is about Dyck path, not derangements.

(a) Compute d_1, d_2, d_3, and d_4 via brute force.

(b) Show that d_n satisfies the following recurrence for $n \geq 1$:

$$d_n = \sum_{i=0}^{n-1} d_i d_{n-1-i}.$$

Hint: Consider the collection of Dyck paths that hit the line $y = x$ at $(i+1, i+1)$ for the first time after leaving $(0,0)$. Think about how many ways you can draw the Dyck path to get to $(i+1, i+1)$ versus how many ways you can draw the Dyck path from $(i+1, i+1)$ to (n,n). The first case is trickier to think about. Notice that the portion of the Dyck path from $(0,0)$ to $(i+1, i+1)$ never hits the line $y = x$ along the way. Moreover, this portion necessarily starts with a North step and ends with an East step. What are the possible values for i?

Since d_n satisfies the same recurrence and initial conditions, it follows that $d_n = c_n$. That is, the number of Dyck paths is a Catalan number.

Problem 6.22. A sequence of parentheses is **balanced** if it can be parsed syntactically. In other words, there should be the same number of left parentheses "(" and right parentheses ")", and when reading from left to right there should never be more right parentheses than left. Here are the five balanced parenthesizations containing three pairs:

$$()()(), ()(()), (()()), (())(), ((())).$$

Prove that the number of balanced sequences of n pairs of parentheses is c_n. *Hint:* Use a bijection!

Problem 6.23. A **triangulation** of a convex $(n+2)$-gon is a dissection into n triangles using only lines from vertices to vertices. Think of the polygon as being fixed in space. Prove that the number of triangulations of a convex $(n+2)$-gon is c_n. Incidentally, this is the problem that Euler was interested in when he studied the Catalan numbers!

Let's see if we can find a closed form for the Catalan numbers!

Problem 6.24. Tackle each of the following.

(a) Argue that the number of lattice paths (not just Dyck paths) from $(0,0)$ to (n,n) is equal to $\binom{2n}{n}$.

(b) Argue that the number of lattice paths from $(0,0)$ to $(n+1, n-1)$ is equal to $\binom{2n}{n-1}$.

(c) Prove that there is a bijection from the set of lattice paths from $(0,0)$ to (n,n) that pass below $y=x$ at least once to the set of lattice paths from $(0,0)$ to $(n+1, n-1)$. *Hint:* Consider the first point on lattice path from $(0,0)$ to (n,n) that passes below $y=x$. Reflect the remaining portion of the path over the appropriate line to get a path from $(0,0)$ to $(n+1, n-1)$.

(d) Prove that $d_n = \dbinom{2n}{n} - \dbinom{2n}{n-1}$.

It is easy to verify that $\dbinom{2n}{n} - \dbinom{2n}{n-1} = \dfrac{1}{n+1}\dbinom{2n}{n}$, and since $d_n = c_n$, we obtain

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