30400, Discrete Math Lecture Notes

Fabrizio Iozzi

Contents

Chapter 1

Counting

1.1. Natural Numbers and Mathematical Induction

In the following, whenever a number n has the property Q we will write $Q(n)$.

Definition 1.1. The set $\mathbb{N}$ of the natural numbers is defined as follows:

(1) 0 is a natural number.
(2) If m is a natural number then there is another natural number $\hat{m}$ which is called the successor of m.
(3) $0 \neq \hat{m}$ for every natural number m.
(4) If $\hat{m} = \hat{n}$ then $m = n$.
(5) Principle of Mathematical Induction (PMI): If Q is a property and if
   (a) $Q(0)$;
   (b) $Q(m) \implies Q(\hat{m})$;
   then $Q(n)$ for all $n \in \mathbb{N}$. ◊

Theorem 1.2 (Well-Ordering Principle (WOP))**.** *Every non-empty set $S \subseteq \mathbb{N}$ has a least element* $\min S$*.*

Theorem 1.3 (Maximal Principle (MP))**.** *Every non-empty bounded above set $T \subseteq \mathbb{N}$ has a greatest element* $\max T$*.*

Theorem 1.4. *WOP, PMI and MP are logically equivalent.*

Proof. It is clear that, if $\min S$ and $\max T$ exist in WOP and MP, then they are uniquely defined. The proof is divided into three steps: WOP $\implies$ MP $\implies$ PMI $\implies$ WOP.

- WOP $\implies$ MP. Suppose T is non empty and bounded above. Let S be
$$S = \{n \in \mathbb{N} : \forall t \in T, t < n\}$$
and let t' be an upper bound of T. By construction, for all $t \in T$ we have $t \leq t' < t' + 1$, so that $t' + 1 \in S$ and S is non-empty. By WOP there is an

integer $s' = \min S$. Since s' is the least element of S, $s' - 1$ does not belong to S, hence $s' - 1 \in T$. Moreover, $t < s'$ for all $t \in T$ means that $t \leq s' - 1$, i.e. $s' - 1 = \max T$.

- MP $\implies$ PMI. Let $P(n)$ be a proposition defined for all $n \in \mathbb{N}$ and suppose PMI is not satisfied, that is $P(0)$ holds true and $P(n+1)$ is true whenever $P(n)$ is true but there exists a (positive) integer n_0 such that $P(n_0)$ does not hold. Let
$$T = \{t \in \mathbb{N} : \forall n \in \mathbb{N} \cap [0, t], P(n) \text{ holds}\}.$$
Then by construction $T \neq \varnothing$ (it contains 0) and it is upper bounded by n_0: by MP the greatest integer $t' = \max T$ is well-defined. Hence the proposition $P(t')$ holds true and by assumption $P(t'+1)$ holds true too. It contradicts the definition of t' as far as it would imply $t' + 1 \in T$.
- PMI $\implies$ WOP. Let $S \subseteq \mathbb{N}$ be a non-empty set. Suppose that $\min S$ does not exist and let $P(n)$ be the proposition $S \cap [0, n] = \emptyset$. Clearly $P(0)$ holds true, otherwise 0 belongs to S and it would be surely its least element. Suppose that $P(n)$ holds true for some non-negative integer n. Then $n+1$ has to belong to S, otherwise it would be its least element. Hence $P(n+1)$ is true as well and, by PMI, $n \notin S$ for all non-negative integers n. But this is impossible since S is non-empty. $\square$

Theorem 1.5. *A set with n elements has 2^n subsets*

Proof. The statement is true for $n = 1$ because a set A with one element has exactly two subsets: A itself and $\varnothing$.

Suppose now the statement is true for $n-1$ and let's prove it for n. Denote the number of subsets of a set A of n elements by S_n. Take one of the elements, say $x \in A$, and consider separately the subsets of A which contain x and those which do not. There are S_{n-1} subsets not containing x (because they are the subsets of the $n-1$–element set consisting of the original set without x). The number of subsets which do contain x is also S_{n-1}, since a subset with x is formed by adding x to one of the subsets without x. So,
$$S_n = S_{n-1} + S_{n-1} = 2S_{n-1}$$
By the induction hypothesis, $S_{n-1} = 2^{n-1}$ whence
$$S_n = S_{n-1} + S_{n-1} = 2S_{n-1} = 2 \cdot 2^{n-1} = 2^n$$
and the theorem is proved. $\square$

Definition 1.6. A sequence is a function $f : \mathbb{N} \to \mathbb{R}$. The values taken by a sequence are called terms of the sequence. $\Diamond$

For sequences, we drop the usual $f(x)$ notation in favor of a_n, where a is the name of the sequence and the subscript n is the input of the function.

A sequence can be defined as a correspondence between any integer and the value taken by the function at that integer, as in
$$a_n = \frac{n+3}{n^2+n+1}$$

This is called closed form. However, sequences have another peculiar way of being defined, that is the recursive form. In this form, a starting term is given, usually a_0, and a rule for getting the next term via the previous one (or, sometimes, the previous ones). For example,

$$\begin{cases} a_0 = 1 \\ a_{n+1} = \frac{a_n}{2} + \frac{1}{a_n} \end{cases}$$

The closed form easily provides the value of any term of the sequence but does not reveal if there is a relation between that term and those that come before it. On the contrary, the recursive form states precisely how to get one term from the preceding ones, but to get any term one must compute all the previous terms.

Example 1.7. Prove by mathematical induction that

$$s_n = 1 + 2 + 3 + \cdots + n = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

The statement is trivially true for $n = 1$. In fact

$$s_1 = \sum_{i=1}^{1} i = \frac{1(1+1)}{2} = 1$$

We now assume the statement is true for n, that is $s_n = n(n+1)/2$ and want to prove it true for $n+1$, that is $s_{n+1} = (n+1)(n+2)/2$. We have

$$\begin{aligned} s_{n+1} &= 1 + 2 + \cdots + n + n + 1 = s_n + n + 1 \\ &= \frac{n(n+1)}{2} + n + 1 \qquad \text{by the inductive hypothesis} \\ &= \frac{n^2 + n + 2n + 2}{2} = \frac{n^2 + 3n + 2}{2} = \frac{(n+1)(n+2)}{2} \end{aligned}$$

△

Theorem 1.8. *For any $q \neq 1$ we have*

$$\sum_{i=0}^{n} q^i = \frac{q^{n+1} - 1}{q - 1}$$

Proof. The proof is by induction. For $n = 1$ the statement is true

$$\sum_{i=0}^{1} q^i = 1 + q = \frac{q^2 - 1}{q - 1}$$

Now, assume the statement is true for n

$$\sum_{i=0}^{n} q^i = \frac{q^{n+1} - 1}{q - 1}.$$

Then

$$\sum_{i=0}^{n+1} q^i = \sum_{i=0}^{n} q^i + q^{n+1}$$

and therefore

$$\sum_{i=0}^{n+1} q^i = \frac{q^{n+1}-1}{q-1} + q^{n+1} = \frac{q^{n+1}-1+q^{n+2}-q^{n+1}}{q-1} = \frac{q^{n+2}-1}{q-1}$$

and the statement is proved for $n+1$. □

Theorem 1.9 (Strong Mathematical Induction)**.** *To prove $P(n)$ for all positive integers n we complete two steps:*

- *(base step) we verify that $P(0)$ is true;*
- *(inductive step) we show the implication*

$$P(0) \wedge P(1) \wedge \cdots \wedge P(k) \implies P(k+1)$$

is true for all positive integers k.

Because we can use all k statements $P(0), P(1), \ldots, P(k)$ to prove $P(k+1)$, rather than just the $P(k)$ as in a proof by mathematical induction, strong induction is a more flexible proof technique.

It can be proved that Mathematical Induction and Strong Induction are equivalent.

Example 1.10. Show that if $n \in \mathbb{N}$, $n > 1$, then n can be written as the product of primes.

Let $P(n)$ be the proposition that n can be written as the product of primes.

$P(2)$ is true, because 2 can be written as the product of one prime, itself.

The inductive hypothesis is the assumption that $P(j)$ is true for all integers j with $2 \le j \le k$, that is, the assumption that j can be written as the product of primes whenever j is a positive integer at least 2 and not exceeding k. To complete the inductive step, it must be shown that $P(k+1)$ is true, that is, that $k+1$ is the product of primes.

There are two cases to consider, namely, when $k+1$ is prime and when $k+1$ is composite. If $k+1$ is prime, we immediately see that $P(k+1)$ is true. Otherwise, $k+1$ is composite and can be written as the product of two positive integers a and b with $2 \le a \le b < k+1$. Because both a and b are integers at least 2 and not exceeding k, we can use the inductive hypothesis to write both a and b as the product of primes. Thus, if $k+1$ is composite, it can be written as the product of primes, namely, those primes in the factorization of a and those in the factorization of b. △

Also the Well Ordering Principle can be used in proofs.

Example 1.11. In a round-robin tournament every player plays every other player exactly once and each match has a winner and a loser. We say that the players $p_1, p_2, \ldots, p_m$ form a cycle if p_1 beats p_2, p_2 beats p_3, $\ldots p_{m-1}$ beats p_m and p_m beats p_1. Use the well-ordering principle to show that if there is a cycle of length m ($m \ge 3$) among the players in a round-robin tournament, there must be a cycle of three of these players.

We assume that there is no cycle of three players. Because there is at least one cycle in the round-robin tournament, the set of all positive integers n for which there is a cycle of length n is nonempty. By the well-ordering property, this set of positive integers has a least element k, which by assumption must be greater than three. Consequently, there exists a cycle of players $p_1, p_2, \ldots, p_k$ and no shorter cycle exists. Because there is no cycle of three players, we know that $k > 3$. Consider the first three elements of this cycle, p_1, p_2 and p_3. There are two possible outcomes of the match between p_1 and p_3. If p_3 beats p_1, it follows that p_1, p_2, p_3 is a cycle of length three, contradicting our assumption that there is no cycle of three players. Consequently, it must be the case that p_1 beats p_3. This means that we can omit p_2 from the cycle $p_1, p_2, \ldots, p_k$ to obtain the cycle $p_1, p_3, \ldots, p_k$ of length $k-1$, contradicting the assumption that the smallest cycle has length k. We conclude that there must be a cycle of length three. $\triangle$

Exercises on section 1.1

Exercise 1.1. Prove that for every natural number n,

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

Exercise 1.2 (A computer science example)**.** A formal language to be used for simple algebra uses the six symbols,

$$x, y, z, (,), +$$

The words of the language are strings of symbols formed according to the rules:

(1) x, y and z are words.

(2) If A and B are words, so is $(A)(B)$.

(3) If A and B are words, so is $(A) + (B)$.

Nothing is a word other than the words formed according to the rules above. For example, $((x)(z)) + (z)$ is a word, since x and z are words (Rule 1.) and therefore so is $(x)(z)$ (Rule 2.) and hence $((x)(z)) + (z)$ (Rule 3.). However, $(x) + z$ is not a word. Prove that any word in this language has the same number of ")"s and ")"s.

Exercise 1.3. A jigsaw puzzle is put together by successively joining pieces that fit together into blocks. A move is made each time a piece is added to a block, or when two blocks are joined. Use strong induction to prove that no matter how the moves are carried out, exactly $n-1$ moves are required to assemble a puzzle with n pieces.

Exercise 1.4. Prove by mathematical induction that

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

Exercise 1.5. Prove by mathematical induction that

$$1 + 3 + 5 + \cdots + 2n + 1 = (n+1)^2$$

Exercise 1.6. Prove by mathematical induction that

$$2^n > n$$

Exercise 1.7. Prove by mathematical induction that if $x > -1$, $x \neq 0$ and $n \geq 2$

$$(1+x)^n > 1 + nx$$

Exercise 1.8. Prove by mathematical induction that there exists $N \in \mathbb{N}$ such that

$$2^n > n^2$$

for every $n > N$ and find N.

Exercise 1.9. Prove by mathematical induction that any positive integer can be expressed as a sum of distinct powers of 2. [Hint: consider induction on the exponent of 2.]

Exercise 1.10. The following argument proves by mathematical induction that all horses are of the same color.

- Base case: In a set of only one horse, there is only one color.
- Induction step: Assume as induction hypothesis that within any set of n horses, there is only one color. Now look at any set of $n+1$ horses. Number them: 1, 2, 3, $\ldots$, n, $n+1$. Consider the sets $\{1, 2, 3, \ldots, n\}$ and $\{2, 3, 4, \ldots, n+1\}$. Each is a set of only n horses, therefore within each there is only one color. But the two sets overlap, so there must be only one color among all $n+1$ horses.

What's wrong with this argument?

Exercise 1.11. Show that for any $n \geq 1$ the number $8^n - 3^n$ is divisible by 5.

Exercise 1.12. Suppose an ATM machine has only two dollar and five dollar bills. You can type in the amount you want, and it will figure out how to divide things up into the proper number of two's and five's. Prove that the ATM machine can generate any output amount $n \geq 4$

Exercise 1.13. Prove by induction that for every $n > 23$ there are two non negative integers x, y such that $n = 7x + 5y$.

Exercise 1.14. Consider the sum

$$a_n = \frac{1}{2!} + \frac{2}{3!} + \frac{3}{4!} + \cdots + \frac{n}{(n+1)!}$$

(1) Compute the first 4 terms of a_n
(2) Formulate a conjecture about a closed formula for a_n
(3) Prove by induction the above formula

Exercise 1.15. Show by induction that the derivative of x^n is nx^{n-1} using the Leibniz rule about the derivative of the product of two functions.

Exercise 1.16. Given

$$s_n = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n \cdot (n+1)}$$

prove that

$$s_n = \frac{n}{n+1}$$

Exercise 1.17. Let $x \in \mathbb{R}$ and n be an odd integer. Prove or disprove

(1) $\lfloor x - 2 \rfloor = \lfloor x \rfloor - 2$
(2) $\left\lfloor \frac{n^2}{4} + 1 \right\rfloor = \frac{n^2+3}{4}$

1.2. Basic Counting Principles

Theorem 1.12 (Product rule)**.** *Suppose that a procedure can be broken down into a sequence of two tasks. If there are n_1 ways to do the first task and for each of these ways of doing the first task, there are n_2 ways to do the second task, then there are $n_1 n_2$ ways to do the procedure.*

Example 1.13. A new company rents a building with 12 floors. The company plans to install two vending machines on two different floors. How many ways are there to place the machines?

Deciding where to install machine 1 is a task that can be done in 12 ways. When machine 1 is placed, machine 2 can be installed in one of the remaining 11 floors, a task that obviously can be done in 11 ways. By the product rule, there are $12 \cdot 11 = 132$ ways to install the two vending machines. △

Example 1.14. How many different car license plates can be made if each plate contains a sequence of three uppercase English letters followed by three digits? There are 26 English letters and 10 digits. Since the process of creating a plate consists in choosing one by one each element, the overall number of possible license plates is

$$26 \times 26 \times 26 \times 10 \times 10 \times 10 = 17576000$$

△

Example 1.15. Consider a finite set A with n elements. How many subsets does A have? To create a subset of A one needs to go through every element of it and choose whether to include it into the subset or not. We then have n independent choices and each choice has two options. Therefore, by the product rule, we have

$$\underbrace{2 \times 2 \times 2 \times \cdots \times 2}_{n \text{ times}} = 2^n$$

△

Example 1.16. How many passwords can a system generate if they are to be 6 to 8 character long (only lowercase letters and digits) and there must be at least 1 digit? There are 36^6 passwords with 6 characters made with all 36 symbols. Those with only letters are 26^6. Then the number of password of length 6 that contain at least 1 digit is $36^6 - 26^6$. Similarly for length 7 and 8. Therefore, the total number of passwords allowed on the system is

$$36^6 - 26^6 + 36^7 - 26^7 + 36^8 - 26^8 = 2684483063360$$

△

Theorem 1.17 (*)**.** *For all finite sets A and B, the number of distinct functions, $f : A \to B$, mapping A to B is $|B|^{|A|}$*

Proof. Suppose $A = \{a_1, \ldots, a_m\}$. There is a one-to-one correspondence between functions $f : A \to B$ and strings (sequences) of length $m = |A|$ over an alphabet of size $n = |B|$

$$f : A \to B \simeq f(a_1), \ldots, f(a_m)$$

By the product rule, there are n^m such strings of length m. □

The product rule is often phrased in terms of sets in this way: If $A_1, A_2, \ldots, A_m$ are finite sets, then the number of elements in the Cartesian product of these sets is the product of the number of elements in each set. To relate this to the product rule, note that the task of choosing an element in the Cartesian product $A_1 \times A_2 \times \cdots \times A_m$ is done by choosing an element in A_1, an element in A_2, ..., and an element in A_m. By the product rule it follows that

$$|A_1 \times A_2 \times \cdots \times A_m| = |A_1| \times |A_1| \times \cdots \times |A_m|.$$

Theorem 1.18 (Sum rule)**.** *If a task can be done either in one of n_1 ways or in one of n_2 ways, where none of the set of n_1 ways is the same as any of the set of n_2 ways, then there are $n_1 + n_2$ ways to do the task.*

Example 1.19. A college library has 40 textbooks on sociology and 50 textbooks on physics and no book is about both subjects. By the rule of sum, a student at this college can select among $40 + 50 = 90$ textbooks in order to learn more about one or the other of these two subjects. △

The sum rule can be phrased in terms of sets as: If $A_1, A_2, \ldots, A_m$ are pairwise disjoint finite sets, then the number of elements in the union of these sets is the sum of the numbers of elements in the sets. To relate this to our statement of the sum rule, note there are $|A_i|$ ways to choose an element from A_i for $i = 1, 2, \ldots, m$. Because the sets are pairwise disjoint, when we select an element from one of the sets A_i , we do not also select an element from a different set A_j $(i \neq j)$. Consequently, by the sum rule, because we cannot select an element from two of these sets at the same time, the number of ways to choose an element from one of the sets, which is the number of elements in the union, is

$$|A_1 \cup A_2 \cup \cdots \cup A_m| = |A_1| + |A_1| + \cdots + |A_m|$$

when $A_i \cap A_j = \varnothing$ for all $i \neq j$.

Remark 1.20. If a task can be done in either n_1 ways or n_2 ways, then the number of ways to do the task is $n_1 + n_2$ minus the number of ways to do the task that are common to the two different ways.

In terms of sets, this rule is an extension to the sum rule above where A and B are not necessarily disjoint. Formally we have

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

The above remark is true for any collection of sets and is known as the principle of inclusion–exclusion which will be stated and proved later.

Example 1.21. How many binary strings of length 8

- start with 1 or
- end with 00?

The number of strings that start with 1 is $2^7 = 128$, since there are 7 places to be filled, each with one of two symbols. The number of strings that end with 00 is 2^6, similarly as before, since in the starting 6 positions we can choose one of the two symbols. Those that start with 1 and end with 00 are 2^5, for only five places are free. Thus, the number of string that meet all conditions above is

$$128 + 64 - 32 = 160$$

△

Example 1.22. How many bit string of length eight contain either three consecutive 0s or four consecutive 1s?

First let us compute the number of bit strings with 3 consecutive 0's. 3 consecutive 0's can start at position 1, 2, 3, 4, 5, or 6 (as there are total eight positions):

(a) Starting at position 1, strings of type 000xxxxx where the x's are 0s or 1s. The total number is $2^5 = 32$.

(b) Starting at position 2: strings of type 1000xxxx. The first position must contain 1 otherwise we will count twice the strings already considered in the first case. The total number is $2^4 = 16$.

(c) Starting at position 3: strings of type x1000xxx. The second position must contain 1 otherwise we will count twice the strings already considered in previous cases. The total number is $2^4 = 16$.

(d) Starting at position 4: strings of type xx1000xx. As for the previous cases, the total number is $2^4 = 16$.

(e) Starting at position 5: strings of type xxx1000x. Counting similarly to the previous cases, we get $2^4 = 16$ strings but in this case we are considering the strings 00010000 and 00010001, which were already counted in the first case. Thus, the total number is 14.

(f) Starting at position 6: strings of type xxxx1000. Counting similarly to the previous cases, we get $2^4 = 16$ strings but in this case we are considering the strings 00001000, 00011000 and 10001000, which were already counted in the first case (the first two strings) or in the second (the third.) Thus, the total number is 13.

The total number is then $32 + 16 + 16 + 16 + 14 + 13 = 107$.

Calculating the number of strings with 4 consecutive 1's in the same manner, we get

(a) Starting at position 1, strings of type 1111xxxx where the x's are 0s or 1s. The overall number is $2^4 = 16$ but we are considering the strings 11110000, 11110001 and 11111000 which contain 3 0s and were already counted before. So the strings in this case are 13.

(b) Starting at position 2: strings of type 01111xxx. The first position must contain 0 otherwise we will count twice the strings already considered in the first case. The overall number is $2^3 = 8$ but the string 01111000 has been already counted above so the strings are 7

(c) Starting at position 3: strings of type x01111xx. The second position must contain 0 otherwise we will count twice the strings already considered in previous cases. The total number is $2^3 = 8$.

(d) Starting at position 4: strings of type xx01111x. The overall number is $2^3 = 8$ but we are considering also 00011110 and 00011111 which were already counted above. So the number is 6 in this case.

(e) Starting at position 5: strings of type xxx01111. Counting similarly to the previous cases, the overall number is $2^3 = 8$ but in this case we are considering

00001111 and 10001111 which were already counted above. Thus, the total number is 6.

The total number is then $13 + 7 + 8 + 6 + 6 = 40$.

Thus, the total number of bit strings of length eight containing either three consecutive 0's or four consecutive 1's is $107 + 40 = 147$. △

Theorem 1.23 (Pigeonhole principle)**.** *For any positive integer m, if $n > m$ objects (pigeons) are placed in m boxes (pigeonholes), then at least one box contains two or more objects.*

Theorem 1.24 (Generalized pigeonhole principle,*)**.** *If N objects are placed into k boxes, then there is at least one box containing at least $\lceil N/k \rceil$ objects.*

Proof. The proof is by contradiction. Suppose that none of the boxes contains more than $\lceil N/k \rceil - 1$ objects. Then, the total number of objects is at most

$$k\left(\left\lceil \frac{N}{k} \right\rceil - 1\right) < k\left[\left(\frac{N}{k} + 1\right) - 1\right] = N + 1 - 1 = N$$

where the inequality

$$\left\lceil \frac{N}{k} \right\rceil < \frac{N}{k} + 1$$

has been used. This is a contradiction because there are a total of N objects. □

Example 1.25. In a class with 90 students at least 8 students are born in the same month. In fact, there are 12 pigeonholes, one per each month, and 90 pigeons. Each pigeonhole can host 7 pigeons but this allows for $7 \times 12 = 84$ pigeons and at least one pigeonhole must contain more than 7. △

Example 1.26. What is the minimum number of cards to select from a standard deck of 52 cards to guarantee that at least three cards of the same suit are chosen? Selecting 8 cards might get you 2 cards for each of the four suits. The ninth card selected must be of one of the four suits and thus the answer is 9. △

Theorem 1.27. *A function f from a set with $k+1$ or more elements to a set with k elements is not injective.*

Example 1.28. For every integer n there is a multiple of n that has only 0s and 1s in its decimal expansion.

To prove this statement, let n be a positive integer. Consider the $n+1$ integers $1, 11, 111, \ldots, 11\ldots1$ (where the last integer in this list is the integer with $n+1$ 1s in its decimal expansion). Note that there are n possible remainders when an integer is divided by n. Because there are $n + 1$ integers in this list, by the pigeonhole principle there must be two with the same remainder when divided by n. The larger of these integers less the smaller one is a multiple of n, which has a decimal expansion consisting entirely of 0s and 1s. △

Exercises for section 1.2

Exercise 1.18. How many injective functions are there from a set with m elements to one with n elements?

Exercise 1.19 (MC)**.** A multiple-choice test contains 10 questions. There are four possible answers for each question. In how many ways can a student answer the questions on the test if the student answers every question?

a) 40
b) 4^{10}
c) 10^4
d) 4^4

Exercise 1.20 (MC)**.** How many bit strings of length ten both begin and end with a 1?

a) 2^{10}
b) 2^9
c) $2^{10} - 2$
d) 2^8

Exercise 1.21. How many positive integers between 100 and 999 inclusive

(1) are divisible by 7?
(2) are odd?
(3) have the same three decimal digits?
(4) are not divisible by 4?
(5) are divisible by 3 or 4?
(6) are not divisible by either 3 or 4?
(7) are divisible by 3 but not by 4?
(8) are divisible by 3 and 4?

Exercise 1.22 (MC)**.** A multiple-choice test contains 10 questions. There are four possible answers for each question. In how many ways can a student answer the questions on the test if the student can leave answers blank?

a) 50
b) 10^5
c) 5^{10}
d) $4^{10} + 10$

Exercise 1.23. A palindrome is a string whose reversal is identical to the string, like in "madam". How many bit strings of length n are palindromes?

Exercise 1.24. How many bit strings of length 10 either begin with three 0s or end with two 0s?

Exercise 1.25. Pick 5 integers from 1 to 8, inclusive. Show that at least two of them sum to 9.

Exercise 1.26. Show that among $n + 1$ arbitrarily chosen integers, there must exist two whose difference is divisible by n.

Exercise 1.27. Consider the first $2n$ integers, 1, 2, 3,..., $2n$. If more than n of these integers are selected, show that two of the selected integers must be relatively prime. Two numbers are relatively prime if they have no common divisor, except for 1.

Exercise 1.28. If each point of the plane is colored red or blue, prove there are two points of the same color at distance one cm from each other.

1.3. Combinatorial calculus

Definition 1.29. A permutation of a set of n distinct objects is an ordered arrangement of these objects. An ordered arrangement of k elements of a set is called a k–permutation. $\Diamond$

Theorem 1.30 (*). *If S is a set with $|S| = n$ and $k \in \mathbb{N}$ is such that $1 \leq k \leq n$, then the number of k–permutations of S is*

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

Proof. Any permutation is defined by the following process. First, one of the n elements of S is to be chosen for the first place. There are n ways to do so. Then, $n-1$ elements of S remain for the second place. By the product rule, the task of choosing the two first elements can be done in $n(n-1)$ ways. Going on this way, we place all the remaining elements until the k–th one. The final task can be done in $n-k+1$ ways. So we get

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

To get the factorial expression for the permutation, just note that by the definition of the factorial we have

$$n(n-1)(n-2)\cdots(n-k+1)(n-k)! = n!. \qquad \square$$

Remark 1.31. If $k = n$, that is, all elements are to be orderly arranged, then

$$P(n,n) = P(n) = n!$$

Example 1.32. How many ways are there to select a first–prize winner, a second–prize winner and a third–prize winner from 100 different people who have entered a contest? The answer is the number of 3–permutations of a set of 100 elements, that is

$$P(100,3) = 100 \cdot 99 \cdot 98 = \frac{100!}{97!} = 970200$$

$\triangle$

Example 1.33. How many permutations of the letters ABCDEFGH contain the string ABC as a (consecutive) substring? We solve this by noting that this number is equal to the number of permutations of the following six objects: ABC, D, E, F, G, and H. So the answer is $6! = 720$ $\triangle$

Definition 1.34. A k–combination, $C(n.k)$, of a set S is an unordered collection of k elements of S. In other words, it is simply a subset of S of size k. $\Diamond$

Theorem 1.35 (*). *If S is a set with $|S| = n$ and $k \in \mathbb{N}$ is such that $1 \leq k \leq n$, then the number of k–combinations of S is*

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

Proof. The $P(n,k)$ k–permutations of the set can be obtained by forming the $C(n,k)$ k–combinations of the set, and then ordering the elements in each k–combination, which can be done in $P(k,k)$ ways. Consequently, by the product rule,

$$P(n,k) = C(n,k) \cdot P(k,k).$$

This implies that

$$C(n,k) = \frac{P(n,k)}{P(k,k)} = \frac{\frac{n!}{(n-k)!}}{\frac{k!}{(k-k)!}} = \frac{n!}{(n-k)!k!} \qquad \square$$

Definition 1.36. The binomial coefficient is defined as

$$\binom{n}{k} = C(n,k) = \frac{n!}{(n-k)!k!} \qquad \Diamond$$

Theorem 1.37 (*). *For all integers $n \geq 1$, and all integers k , $1 \leq k \leq n$*

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

Proof. Suppose that S is a set with n elements. The function that maps a subset $A \subseteq S$ to its complement A^{C} is a bijection between

subsets of S with k elements

and

subsets with $n-k$ elements.

The identity $C(n,k) = C(n,n-k)$ follows because when there is a bijection between two finite sets, the two sets must have the same number of elements.

Alternatively, one might think of a combination as one of the ways to choose k elements in a set of n. But choosing the elements in the set is logically equivalent to choosing those not in the set, whose number is $n-k$. $\square$

Example 1.38. How many bit strings of length n contain exactly k 1s? The positions of k 1s in a bit string of length n form a k–combination of the set $\{1,2,3,\ldots,n\}$. Hence, there are $C(n,k)$ bit strings of length n that contain exactly k 1s. $\triangle$

Im many applications, elements from a set can be used repeatedly.

Theorem 1.39. *The number of k–permutations of a set of n objects with repetition allowed is n^k.*

Proof. There are n ways to select an element of the set for each of the k positions in the k–permutation when repetition is allowed, because for each choice all n objects are available. Hence, by the product rule there are n^k k–permutations when repetition is allowed. $\square$

Example 1.40. How many four letter words (meaningful or not) can be formed using the letters A, B and C? The answer is the number of 4–permutations of a three element set, that is $3^4 = 81$ △

Example 1.41. Four students are going in exchange in three foreign countries, France, UK and the USA. They are completely free to choose the country to go to. How many exchange parties can be set? Say

$$C = \{\text{France}, \text{UK}, \text{USA}\}.$$

A choice from the students is an arrangement of 4 elements of C. For example

$$\{\text{France}, \text{UK}, \text{France}, \text{USA}\}.$$

meaning the first and third student going to France, the second to the UK and the fourth to the USA. Then, the answer is the number of 4–permutations of a three element set when repetition is allowed, that is $3^4 = 81$.

This can also be seen as finding the number of functions from set $S = \{1, 2, 3, 4\}$, the set of the students, to set C. Since $|S| = 4$ and $|C| = 3$, the number of functions from S to C is 3^4. △

Theorem 1.42 (*)**.** *There are $C(n+k-1, k) = C(n+k-1, n-1)$ k–combinations from a set with n elements when repetition of elements is allowed.*

Proof. To probe the theorem we use the method usually called "stars and bars". Each k–combination of a set with n elements when repetition is allowed can be represented by a list of $n-1$ bars and k stars. The $n-1$ bars are used to mark off n different cells, with the i–th cell containing a star for each time the i–th element of the set occurs in the combination.

For instance, a 6–combination of a set with four elements is represented with three bars and six stars. Here

$$* \quad * \quad | \quad * \quad | \quad | \quad * \quad * \quad *$$

represents the combination containing exactly two of the first element, one of the second element, none of the third element, and three of the fourth element of the set. Considering numbers, the string above represents the combination 112444 (the order is not relevant here since these are combinations.)

The number of lists containing $n-1$ bars and k stars corresponds to a k–combination of the set with n elements, when repetition is allowed. The number of such lists is $C(n-1+k, k)$, because each list corresponds to a choice of the k positions to place the k stars from the $n-1+k$ positions that contain k stars and $n-1$ bars. The number of such lists is also equal to $C(n-1+k, n-1)$, because each list corresponds to a choice of the $n-1$ positions to place the $n-1$ bars. □

Example 1.43. How many ways are there to select four pieces of fruit from a bowl containing apples, oranges, and pears if the order in which the pieces are selected does not matter, only the type of fruit and not the individual piece matters, and there are at least four pieces of each type of fruit in the bowl?

We have $n = 3$ and $k = 4$, so the answer is $C(3-1+4, 4) = C(6, 4) = 15$. The combinations are

AAAA	AAAO	AAAP	AAOO	AAOP
AAPP	AOOO	AOOP	AOPP	APPP
OOOO	OOOP	OOPP	OPPP	PPPP

△

Example 1.44. How many solutions does the equation $x+y+z = 11$ have, where x, y, z are nonnegative integers?

Think of 11 "stars" to be assigned to 3 unknowns, x, y and z. We thus have $k = 11$ and $n = 3$ and therefore the equation has

$$C(3-1+11, 11) = C(13, 11) = 78$$

solutions.

What if all unknowns must be strictly positive? In this case, $x \geq 1$, $y \geq 1$ and $z \geq 1$. This means in the expression $x + y + z$, we have to choose only 8 assignments because 3 are already set by the constraint. Then, the number of solutions is $C(3-1+8, 8) = C(10, 8) = 45$ △

Theorem 1.45. *The number of different permutations of n objects, where there are n_1 indistinguishable objects of type 1, n_2 indistinguishable objects of type 2, ..., and n_k indistinguishable objects of type k, is*

$$\frac{n!}{n_1!n_2!\ldots n_k!}$$

If the objects were distinguishable, then the number of permutation would be $n!$. Consider one ordered arrangement of the n object where n_1 of them are indistinguishable. Then, that arrangement will be repeated $n_1!$ times, each corresponding to one of the $n_1!$ ways to arrange the n_1 objects in their positions. The same can be said for all groups of indistinguishable objects, whence the above statement.

Example 1.46. How many different strings (anagrams) can be made by reordering the letters of the word SUCCESS?

We have seven letters. The three Ss and the two Cs are indistinguishable so there are a total of

$$\frac{7!}{3!2!} = \frac{5040}{12} = 420$$

different anagrams of the word SUCCESS. △

Theorem 1.47 (Principle of Inclusion-Exclusion,*)**.** *Let $A_1, A_2, \ldots, A_n$ be finite sets. Then*

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{1\leq i\leq n} |A_i| - \sum_{1\leq i<j\leq n} |A_i \cap A_j| +$$
$$\sum_{1\leq i<j<k\leq n} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1}|A_1 \cap A_2 \cap \cdots \cap A_n|$$

Proof. Suppose that a is a member of exactly r of the sets A$A_1, A_2, \ldots, A_n$ where $1 \leq r \leq n$. The left–hand side of the equation counts this element only once. We will then prove the formula by showing that the same element is counted exactly once by the right–hand side of the equation.

This element is counted $C(r,1)$ times by $\sum|A_i|$ because there are $C(r,1)$ ways of choosing a set A_i so that is contains a. The same element is counted $C(r,2)$ times by $|A_i \cap A_j|$, again because there are $C(r,2)$ ways of choosing two sets, A_i and A_j among the r that contain a. In general, it is counted $C(r,m)$ times by the summation involving m of the sets A_i. Thus, this element is counted exactly

$$C(r,1) - C(r,2) + C(r,3) - \cdots + (-1)^{r+1}C(r,r)$$

times by the expression on the right–hand side of this equation. By exercise 1.45 we have

$$C(r,0) - C(r,1) + C(r,2) - C(r,3) + \cdots + (-1)^{r+1}C(r,r) = 0$$

Hence,

$$1 = C(r,0) = C(r,1) - C(r,2) + C(r,3) - \cdots + (-1)^{r+1}C(r,r).$$

Therefore, each element in the union is counted exactly once by the expression on the right–hand side of the equation. This proves the theorem. □

Exercises on section 1.3

Exercise 1.29. A group contains n men and n women. How many ways are there to arrange these people in a row if the men and women alternate?

Exercise 1.30. How many permutations of the letters ABCDEFG contain

(1) the string BCD?
(2) the string CFGA?
(3) the strings BA and GF?
(4) the strings ABC and DE?
(5) the strings ABC and CDE?
(6) the strings CBA and BED?

Exercise 1.31. In how many ordered ways can we draw 5 cards from an ordinary deck of 52 cards

(1) with replacement? (that is, putting back the card into the deck after it has been chosen)
(2) without replacement? (that is, leaving the chosen card out of the deck)

Exercise 1.32. Eight people are divided into four pairs to play bridge. In how many ways can this be done?

Exercise 1.33. A domino is an ordered pair $(m; n)$ with $0 \le m \le n \le 6$. How many dominoes are in a set if there is only one of each type?

Exercise 1.34. What is the number of permutations of the set $\{1, 2, \ldots, n\}$ in which 1 precedes 2?

Exercise 1.35. What is the number of permutations of $\{1, 2, \ldots, n\}$ in which the elements 1 and 2 are not next to each other?

Exercise 1.36. What is the number of words of length k in an alphabet of n letters

(1) in which no two consecutive letters are the same
(2) which are palindromes

Exercise 1.37. In how many ways can $\{1, 2, 3, \ldots, n\}$ be partitioned into two disjoint subsets? Assume the order of the two subsets does not matter.

Exercise 1.38. Twenty five girls and twenty five boys sit around a table. Is it always possible to find a person both of whose neighbors are girls?

Exercise 1.39. How many binary $n \times m$ matrices with pairwise different rows are there?

Exercise 1.40. You are given a set U of n elements and a subset $A \subseteq U$ of k elements. Determine the number of subsets $B \subseteq U$ satisfying

(1) $B \subset A$
(2) $B \supset A$
(3) $A \cap B = \emptyset$
(4) $A \cap B \neq \emptyset$

Exercise 1.41. How many ways are there to assign three jobs to five employees if each employee can be given more than one job?

Exercise 1.42. A book publisher has 3000 copies of a discrete mathematics book. How many ways are there to store these books in their three warehouses if the copies of the book are indistinguishable?

Exercise 1.43. How many strings of 10 ternary digits (0, 1, or 2) are there that contain exactly two 0s, three 1s, and five 2s?

Exercise 1.44. How many solutions are there to the equation

$$x_1 + x_2 + x_3 + x_4 + x_5 = 21,$$

where x_i , $i = 1, 2, 3, 4, 5$, is a nonnegative integer such that

(1) $x_1 \geq 1$?
(2) $x_i \geq 2$ for $i = 1, 2, 3, 4, 5$?

1.4. The binomial theorem

Theorem 1.48 (Binomial theorem)**.** *For all $n \geq 0$ and for all $x, y \in \mathbb{R}$*

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

Proof. Since $x + y$ is a homogeneous first degree polynomial in x, y, $(x + y)^n$ is a homogeneous polynomial of degree n. Therefore, its terms are all product of powers of x and y, $x^k y^{n-k}$.

The coefficient of $x^k y^{n-k}$ counts the number of ways k x's and $n-k$ y's can be multiplied to get the final term. Since

$$(x+y)^n = \underbrace{(x+y)(x+y)\dots(x+y)}_{n \text{ times}}$$

k x's can be chosen in $\binom{n}{k}$ ways. This is true for $0 \leq k \leq n$, 0 corresponding to choosing no x's at all, and therefore to the term y^n, and n corresponding to choosing x in every factor, to get x^n. The theorem is thus proved ☐

Theorem 1.49 (*)**.**

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

Proof. An algebraic proof is trivial: setting $x = y = 1$ in the Binomial theorem gets

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

There is also a simple combinatorial proof. Since the binomial coefficient counts the number of k–subsets of a set with n elements, the sum of all binomial coefficients is equal to the total number of subsets which is known to be 2^n. ☐

Theorem 1.50 (Pascal's identity,*)**.** *For all $n \geq 1$,*

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

Proof. Let $S = \{1, 2, \dots, n\}$. Then $\binom{n}{k}$ counts the number of k–subsets if S. Fix an element from the set, say $\{1\}$. The number of k–subsets is the sum of two groups of k–subsets: those that contain $\{1\}$ and those that don't.

The number of k–subsets that don't contain $\{1\}$ is $\binom{n-1}{k}$, because the elements of these subsets are to be chosen among $\{2, 3, \dots, n\}$.

The number of k–subsets that contain $\{1\}$ is $\binom{n-1}{k-1}$, because any of these subsets is the union of $\{1\}$ and a $(k-1)$–subset of $\{2, 3, \dots, n\}$.

This proves the identity ☐

Remark 1.51. Pascal's identity together with the "boundary" conditions

$$\binom{n}{0} = \binom{n}{n} = 1$$

is the rule behind Pascal's triangle.

Theorem 1.52 (Vandermonde's identity,*)**.** *Let $m, n,$ and r be nonnegative integers with r not exceeding either m or n. Then*

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

Proof. Suppose that there are m items in one set and n items in a second set. Then the total number of ways to pick r elements from the union of these sets is

$$\binom{m+n}{r}.$$

Another way to pick r elements from the union is to pick k elements from the second set and then $r-k$ elements from the first set, where k is an integer with $0 \le k \le r$. Because there are

$$\binom{n}{k}$$

ways to choose k elements from the second set, and

$$\binom{m}{r-k}$$

ways to choose $r-k$ elements from the first set, the product rule tells us that this can be done in

$$\binom{n}{k}\binom{m}{r-k}$$

ways. Hence, the total number of ways to pick r elements from the union also equals

$$\sum_{k=0}^{r}\binom{m}{r-k}\binom{n}{k}$$

We have found two expressions for the number of ways to pick r elements from the union of a set with m items and a set with n items. Equating them gives us Vandermonde's identity. □

Theorem 1.53 (*). *Let n, and r be nonnegative integers with $r \le n$. Then*

$$\binom{n+1}{r+1} = \sum_{j=r}^{n}\binom{j}{r}$$

Proof. The left-hand side,

$$\binom{n+1}{r+1}$$

counts the bit strings of length $n+1$ containing $r+1$ ones. We show that the right-hand side counts the same objects by considering the cases corresponding to the possible locations of the final 1 in a string with $r+1$ ones. This final one must occur at position $r+1$, $r+2$,..., or $n+1$. Furthermore, if the last one is the k–th bit there must be r ones among the first $k-1$ positions. Consequently there are

$$\binom{k-1}{r}$$

such bit strings. Summing over k with $r+1 \le k \le n+1$, we find that there are

$$\sum_{k=r+1}^{n+1}\binom{k-1}{r}$$

bit strings of length n containing exactly $r+1$ ones. If we set $j = k-1$ in the last expression we get

$$\sum_{k=r+1}^{n+1} \binom{k-1}{r} = \sum_{j=r}^{n} \binom{j}{r}$$

and the theorem is proved. $\square$

Exercises on section 1.4

Exercise 1.45. Let n be a positive integer. Prove that

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

Exercise 1.46. Let n be a positive integer. Prove that

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

Exercise 1.47. Consider

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

(1) Use Pascal's triangle to find the A_n for $n = 1, 2, 3, 4$
(2) Find the results in the Pascal's triangle and guess a value for A_n
(3) Prove your guess.

Exercise 1.48. Starting from

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

prove that

$$\sum_{k=0}^{n} k \binom{n}{k} = n2^{n-1}$$

1.5. Representation of integers

Theorem 1.54. *Let $b \in \mathbb{N}$ be such that $b > 1$. Every $n \in \mathbb{N}$, $n > 0$ can be expressed uniquely in the form*

$$n = a_k b^k + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0,$$

where k is a non-negative integer, a_0, $a_1, \ldots$, a_k are non-negative integers less than b, and $a_k \neq 0$.

The representation of n given in Theorem 1.54 is called the base b expansion of n. The base b expansion of n is denoted by $(a_k a_{k-1} \ldots a_1 a_0)_b$. For instance, $(245)_8$ represents

$$2 \cdot 8^2 + 4 \cdot 8 + 5 = 165$$

Typically, the subscript 10 is omitted for base 10 expansions of integers because base 10, or decimal expansions, are commonly used to represent integers.

Example 1.55. What is the decimal expansion of the integer that has $(1\ 0101\ 1111)_2$ as its binary expansion? We have

$$\begin{aligned}(1\ 0101\ 1111)_2 &= 1 \cdot 2^8 + 0 \cdot 2^7 + 1 \cdot 2^6 + 0 \cdot 2^5 + 1 \cdot 2^4 + \\ &+ 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 \\ &= 351.\end{aligned}$$

△

Among the most important bases in computer science are base 2, base 8, and base 16. Base 8 expansions are called octal expansions and base 16 expansions are hexadecimal expansions.

Sixteen different digits are required for hexadecimal expansions. Usually, the hexadecimal digits used are 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, and F, where the letters A through F represent the digits corresponding to the numbers 10 through 15 (in decimal notation).

Example 1.56. The decimal expansion of the number with hexadecimal expansion $(2AE0B)_{16}$ is

$$(2AE0B)_{16} = 2 \cdot 16^4 + 10 \cdot 16^3 + 14 \cdot 16^2 + 0 \cdot 16 + 11 = 175627.$$

△

Each hexadecimal digit can be represented using four bits. For instance, we see that $(1110\ 0101)_2 = (E5)_{16}$ because $(1110)_2 = (E)_{16}$ and $(0101)_2 = (5)_{16}$. Bytes, which are bit strings of length eight, can be represented by two hexadecimal digits.

Digits in the binary system are called bits (=BInary digiTS). The largest number one can represent with one bit is 1; with two bits is $11_2 = 3$; with 3 bits is $111_2 = 7$. In general, the largest number with n bits, is $2^n - 1$. Since we start from 0, with n bits we can represent $2^n - 1 + 1 = 2^n$ numbers.

To get the base 10 representation from a base b number, one just computes the powers of n and adds them, as in the example above. To get the base b number from base 10, one starts dividing the number by b: the remainder is kept, and the quotient is then divided again by b, and so on until the quotient is less than b. The last quotient and all the remainders in reverse order are the digits of the number in base b.

Example 1.57. To write 304 in base 6 we have

$$\begin{aligned}304 : 6 &= 50 \text{ remainder } 4 \\ 50 : 6 &= 8, \text{ remainder } 2 \\ 8 : 6 &= 1, \text{ remainder } 2\end{aligned}$$

Then $(304)_{10} = (1224)_6$. To verify $4 + 2 \times 6 + 2 \times 36 + 1 \times 216 = 4 + 12 + 72 + 216 = 304$ △

For fractions, that is numbers in the interval $(0, 1)$, we multiply the number by the base. We take the integer part of the product and this will be the first digit after the decimal point. Then we subtract the integer from the product, getting another fraction and multiply again by the base, repeating the process, until we get 0 or we get a fraction we got before, which means the number is periodic.

For example, to represent 0.5625 in base 4 we proceed as follows:

$$0.5625 \times 4 = 2.25 \text{ integer } 2$$
$$0.25 \times 4 = 1 \text{ integer } 1$$

Since the last product gets an integer, the process stops and $0.5625 = (0.21)_4$. In fact, $2 \times 4^{-1} + 4^{-2} = 0.5 + 0.0625$.

If the fraction is 0.3 we have

$$0.3 \times 4 = 1.2 \text{ integer } 1$$
$$0.2 \times 4 = 0.8 \text{ integer } 0$$
$$0.8 \times 4 = 3.2 \text{ integer } 3$$
$$0.2 \times 4 = 0.8 \text{ integer } 0$$
$$0.8 \times 4 = 3.2 \text{ integer } 3$$
$$\ldots = \ldots$$

so that $0.3 = (0.1030303\ldots)_4$. Numbers that are not periodic in the decimal system could possibly be so in a base b system and vice versa. For example, $1/3 = (0.3333\ldots)_{10}$ in the decimal system but, obviously, $1/3 = (0.1)_3$.

In any base there are numbers that admit different representations. As in the decimal system we have

$$1 = 0.99999\ldots$$

in the base 6 system we have

$$1 = 0.55555\ldots$$

etc.

1.5.1. Representation of numbers inside a computer. The number of bits where to store a number is usually fixed for any given computer. Using binary representation gives us an insufficient range and precision of numbers to do relevant and reliable calculations. To achieve the range of values needed with the same number of bits, we use floating point numbers or float for short.

There are two main kinds of floats: single precision (32 bits) and double precision (64 bits). Most programming language specifications use double precision, but to illustrate the concept of floating point numbers here we use single precision. The mechanism is the same but the numbers are shorter and therefore more manageable.

Instead of using each bit as the coefficient of a power of 2, floats allocate bits to three different parts: the sign indicator, s, which says whether a number is positive or negative; characteristic or exponent, e, which is the power of 2; and the fraction, f, which is the coefficient of the exponent. In the IEEE754 standard for single precision, 1 bit is allocated to the sign indicator, 8 bits are allocated to the exponent, and 23 bits are allocated to the fraction.

With 8 bits allocated to the exponent $2^8 = 256$ values are available. Since we want to be able to make very precise numbers, we want some of these values to represent negative exponents (i.e., to allow numbers that are between 0 and 1). To accomplish this, 127 is subtracted from the exponent to “normalize” it. The value subtracted from the exponent is commonly referred to as the “bias”.

In decimal exponential notation all numbers are written as

$$a_0.a_1a_2a_3\cdots\times 10^k$$

like $3.456\times 10^{-2} = 0.03456$. In any exponential notation, the first digit is always positive (and obviously less than the base). In decimal notation, the digit to the left of the decimal point is between 1 and 9 $(10-1)$.

Since here we are working with powers of 2, the fraction is a number between 1 and 2. In binary, this means that the leading term will always be 1, and, therefore, it is a waste of bits to store it. To save space, the leading 1 is dropped. A float can then be represented as

$$n = (-1)^s 2^{e-127}(1+f).$$

For example, say x is represented (according to standard IEEE754) as

$$1 \quad 10000010 \quad 10000000000000000000000$$

Then the exponent in decimal is

$$1\times 2^7 + 1\times 2^1 - 127 = 3.$$

The fraction is

$$1\times\frac{1}{2^{-1}} + 0\times\frac{1}{2^{-2}} + 0\times\frac{1}{2^{-3}} + \cdots = 0.5.$$

Therefore

$$n = (-1)^1\times 2^3\times(1+0.5) = (-12)_{10}\,.$$

Now, consider 15. Since the number is positive, $s=0$. The largest power of two that is smaller than 15 is 8, so the exponent is 3, making the characteristic $3+127=130=10000010_2$. Then the fraction is

$$\frac{15}{8} - 1 = 0.875 = 1\times 2^{-1} + 1\times 2^{-2} + 1\times 2^{-3} = (0.11100000000000000000000)_2\,.$$

Then, according to IEEE754,

$$15 = 0 \quad 1000010 \quad 11100000000000000000000.$$

Note that the next smallest number is obtained by removing a single bit from the right, that is

$$0 \quad 1000010 \quad 11011111111111111111111 = 14.999999046325684.$$

whilte the next largest number is

$$0 \quad 1000010 \quad 11100000000000000000001 = 15.000000953674316.$$

Therefore, the IEEE754 number 0 1000010 11100000000000000000000 not only represents the number 15, but also all the real numbers halfway between its immediate neighbors, or more specifically, the interval

$$[14.999999523162842, 15.000000476837158].$$

Any computation whose result is within this interval will be assigned 15. We call the distance from one number to the next the gap. Because the fraction is multiplied by 2^{e-127}, the gap grows as the number represented grows. In MATLAB the gap at a given number can be computed using the function `eps`.

What is gained by using IEEE754 versus plain binary? Using 32 bits gives us about 4 billion (2^{32}) numbers. Since the number of bits does not change between binary and IEEE754, IEEE754 must also give us about 4 billion numbers. In binary, numbers have a constant spacing between them. As a result, you cannot have both range (i.e., large distance between minimum and maximum representable numbers) and precision (i.e., small spacing between numbers). Controlling these parameters would depend on where you put the decimal point in your number. IEEE754 overcomes this limitation by using very high precision at small numbers and very low precision at large numbers. This limitation is usually acceptable because the gap at large numbers is still small relatively to the size of the number itself. Therefore, even if the gap is millions large, it is irrelevant to normal calculations if the number under consideration is in the trillions or higher.

Exercises on section 1.5

Exercise 1.49 (MC)**.** How many digits are used in base 7?

a) 6
b) 7
c) 10
d) 8

Exercise 1.50 (MC)**.** The sum of the binary numbers 11001010 and 00111010 is

a) 255
b) 256
c) 260
d) none of the above

Exercise 1.51 (MC)**.** What is the largest decimal number that can be represented with three digits in base 9?

a) 27
b) 999
c) 728
d) none of the above

Exercise 1.52. One's complement representations of integers are used to simplify computer arithmetic. To represent positive and negative integers with absolute value less than 2^{n-1}, a total of n bits is used. The leftmost bit is used to represent the sign. A 0 bit in this position is used for positive integers, and a 1 bit in this position is used for negative integers. For positive integers, the remaining bits are identical to the binary expansion of the integer. For negative integers, the remaining bits are obtained by first finding the binary expansion of the absolute value of the integer, and then taking the complement of each of these bits, where the complement of a 1 is a 0 and the complement of a 0 is a 1.

(1) Find the decimal numbers corresponding to

$$11001, \quad 01101, \quad 10001, \quad 11111.$$

(2) If m is a positive integer less than 2^{n-1}, how is the one's complement representation of $-m$ obtained from the one's complement of m, when bit strings of length n are used?

(3) How is the one's complement representation of the sum of two integers obtained from the one's complement representations of these integers?

Exercise 1.53. $n \in \mathbb{N}$ is such that $0 < n < 10000$. What is the maximum number of digits needed to represent n in base 8?

1.6. Finite and infinite sets

Definition 1.58. If there exists a bijective (1–1) mapping of A onto B, we say that A and B can be put in (1–1) correspondence, or that A and B have the same cadinality and we write $|A| = |B|$. ◊

Definition 1.59. For any positive integer n, let J_n be the set whose elements are the integers $1, 2, \ldots, n$. For any set A, we say:

- A is finite if $|A| = |J_n|$ for some n.
- A is infinite if A is not finite.
- A is countable if $|A| = |\mathbb{N}|$.
- A is uncountable if A is neither finite nor countable.
- A is at most countable if A is finite or countable. ◊

If A is finite and $|A| = |J_n|$ we write $|A| = n$. For two finite sets A and B, we have $|A| = |B|$ if and only if they have the same number of elements. For infinite sets, the above definition tries to catch a similar idea, using a 1–1 correspondence.

Theorem 1.60 (*). *$\mathbb{Z}$ is countable.*

Proof. To prove the statement, it's enough to show a 1–1 correspondence between $\mathbb{N}$ and $\mathbb{Z}$. This is given by:

$$f(n) = (-1)^{n+1} \left\lfloor \frac{n+1}{2} \right\rfloor = \begin{cases} \frac{n}{2} & n \text{ even} \\ -\frac{n-1}{2} & n \text{ odd} \end{cases} \qquad \square$$

The 1–1 correspondence is between a set, $\mathbb{Z}$, and a proper subset of it, $\mathbb{N}$. This is impossible for finite sets: any proper subset has less elements than the entire set. For infinite sets, however, this is indeed an option and in fact it's a characterizing one.

Theorem 1.61 (*). *A countable collection of countable sets is countable.*

Proof. To build the 1–1 correspondence, align the elements in this way

$$\begin{tikzcd}
s_{11} \arrow[r] & s_{12} \arrow[dl] & s_{13} \arrow[r] & s_{14} \arrow[dl] & \cdots \\
s_{21} \arrow[d] & s_{22} \arrow[ur] & s_{23} \arrow[dl] & s_{24} & \cdots \\
s_{31} \arrow[ur] & s_{32} \arrow[dl] & s_{33} & s_{34} & \cdots \\
s_{41} & s_{42} & s_{43} & \cdots &
\end{tikzcd}$$

□

An immediate consequence of the previous theorem is the following.

Theorem 1.62. $\mathbb{Q}$ *is countable.*

Proof. Since every $q \in \mathbb{Q}$ is an integer over another integer, every rational number is a member of a countable collection (the numerators) of a countable set (the denominators) □

Theorem 1.63 (*)**.** $\mathbb{R}$ *is uncountable.*

Proof. Consider the real number in $(0, 1)$. The decimal expansion of each such number is $0.a_1a_2a_3\dots$, where a_i is any digit from the set $\{0, 1, \dots, 9\}$. Now assume, by contradiction, that the real numbers are a countable set. Then the reals in $(0, 1)$ must be countable as well and we can make a list (an infinite list) of all those numbers:

$$\begin{array}{ll}
1 & 0.a_{11}a_{12}a_{13}a_{14}a_{15}\dots \\
2 & 0.a_{21}a_{22}a_{23}a_{24}a_{25}\dots \\
3 & 0.a_{31}a_{32}a_{33}a_{34}a_{35}\dots \\
\dots & \dots
\end{array}$$

Now create another number according to the following rule: for every natural number i, digit i of the new number is 4 if a_{ii} is not 4, and is 7 if a_{ii} is 4. The number created by this rule belongs to $(0, 1)$ but is not in the list because it differs from any number in the list by (at least) one digit. This is a contradiction and therefore the fact the reals are countable must be false. □

Theorem 1.64 (*)**.** *Let A be a set. There is no bijective map between A and $\mathcal{P}(A)$ and hence $|A| < |\mathcal{P}(A)|$*

Proof. We prove this by contradiction. Assume $|A| = |\mathcal{P}(A)|$. Then there is a 1–1 correspondence f between A and $\mathcal{P}(A)$. Let $x \in A$ and let $f(x) \in \mathcal{P}(A)$. $f(x)$ is a subset of A. It might contain x or not. Let

$$B = \{x \in A : x \notin f(x)\}$$

B is a subset of A so it's an element of $\mathcal{P}(A)$ and it must correspond to a $y \in A$, $f(y) = B$. Now, $y \in B$ implies $y \notin f(y) = B$, that is $y \notin B$. On the other hand, $y \notin B$ implies $y \in B$, because that's exactly the way B is defined. Whatever y, we get a contradiction so there's no 1–1 correspondence between A and $\mathcal{P}(A)$ $\square$

The last theorem shows there's no "limit" to infinity, and actually "infintely many elements" is a vague concept. From the computational analysis of problems, the theorem has important consequences.[1]

Every computer program can be expressed as a string drawn from the appropriate alphabet. Thus,

$$\#\text{ of programs} \leq \#\text{ of strings.}$$

The number of strings is strictly less than the number of sets of strings (by the above theorem), that is

$$\#\text{ of strings} < \#\text{ of sets of strings.}$$

Now, let's think about how many problems there are out there that we might want to solve. This really depends on our notion of what a "problem" is, but for now we just focus on one type of problem: deciding whether a string has some property. For example, some strings have even length, some are palindromes, some are legal Java programs, some are mathematical proofs, etc. We can think of a "property" of a string as just a set of strings that happen to share that property. For example, we could say that the property of being a palindrome (reading the same forwards and backwards) could be represented by the set

$$\{a, b, c, \ldots, z, aa, bb, \ldots, zz, aba, aca, ada, \ldots\}.$$

Thus,

$$\#\text{ of sets of strings} = \#\text{ of str. properties.}$$

For each of these properties, we might think about writing a program that could determine whether a string has that given property. In other words, each property of strings (that is, a set of strings) defines a unique problem – determine whether a given string has that property or not. As a result, the total number of problems we might want to solve is at least as big as the number of sets of strings, since there's at least one problem to solve per set of strings. In formula,

$$\#\text{ of str. properties} \leq \#\text{ of problems.}$$

Combining this all together gives the following:

$$\#\text{ of programs} < \#\text{ of problems.}$$

In other words, **there are more problems than there are programs to solve them.** We have just proven that there are problems that cannot possibly be solved by a computer.

[1]What follows is taken from *Mathematical Foundations of Computing* by Keith Schwarz.

Chapter 2

Sequences

2.1. Convergence

Definition 2.1. A sequence is a function $f : \mathbb{N} \to \mathbb{R}$. The values taken by a sequence are called **terms** of the sequence. $\Diamond$

Usually, for sequences we drop the usual notation of functions, $f(n)$, for a_n, where the input is moved to the subscript and the functions are named using the first letters of the alphabet.

Definition 2.2. A real sequence a_n is:

- bounded above if $\exists\ k \in \mathbb{R}$ such that $a_n < k$, $\forall n$
- bounded below if $\exists\ k \in \mathbb{R}$ such that $a_n > k$, $\forall n$
- bounded if it's both bounded above and below
- increasing if $\forall n$ we have $a_{n+1} \geq a_n$ and strictly increasing if $\forall n$ we have $a_{n+1} > a_n$
- decreasing if $\forall n$ we have $a_{n+1} \leq a_n$ and strictly decreasing if $\forall n$ we have $a_{n+1} < a_n$ $\Diamond$

The only sequences that are both increasing and decreasing are constant sequences $a_n = c$, $\forall n$.

Example 2.3. $a_n = a \cdot q^n$ is a real sequence called geometric sequence with first term a and common ratio q.

$$a_n = \begin{cases} a_0 = a \\ a_{n+1} = q \cdot a_n \end{cases} = \left\{a, aq, aq^2, aq^3, \dots\right\} \qquad \triangle$$

Example 2.4. $a_n = a + nb$ is a real sequence called arithmetic sequence with first term a and common difference b.

$$a_n = \begin{cases} a_0 = a \\ a_{n+1} = b + a_n \end{cases} = \{a, a+b, a+2b, a+3b, \dots\} \qquad \triangle$$

Example 2.5. The sequence defined by

$$a_n = \left\{ \frac{n}{n+1} \right\} \qquad n = 0, 1, 2, \ldots$$

is a strictly increasing. In fact

$$a_{n+1} - a_n = \frac{n+1}{n+2} - \frac{n}{n+1} = \frac{1}{(n+1)(n+2)} > 0$$

for all n. a_n is also bounded since for every n we have $0 < a_n < 1$. △

Example 2.6. The sequence defined by $a_n = \{(-1)^n\}$ is bounded since for every n we have $-2 < a_n < 2$. It is neither increasing nor decreasing. △

Definition 2.7. A sequence $\{p_n\}$ is said to converge if there is a point $p \in \mathbb{R}$ with the following property: for every $\varepsilon > 0$ there is an integer N such that $n > N$ implies that $|p_n - p| < \varepsilon$. In this case we also say that $\{p_n\}$ converges to p, or that p is the limit of $\{p_n\}$, and we write

$$\lim_{n \to \infty} p_n = p \qquad \Diamond$$

When $\{p_n\}$ does not converge, we say $\{p_n\}$ is divergent or that the limit does not exist.

Example 2.8. For the sequence $a_n = 1/n$ we have $a_n \to 0$ as $n \to +\infty$. To prove it, for every $\varepsilon > 0$ set $N = \lfloor 1/\varepsilon \rfloor + 1$. Then, for $n > N$, we have

$$0 < a_n = \frac{1}{n} < \frac{1}{\lfloor 1/\varepsilon \rfloor + 1} < \frac{1}{\lfloor 1/\varepsilon \rfloor} < \frac{1}{1/\varepsilon} = \varepsilon$$

which means $d(a_n, 0) = |a_n - 0| < \varepsilon$. △

Theorem 2.9. *Let p_n be a sequence.*

(1) $p_n \to p$ if and only if every neighborhood of p contains all but finitely many of the terms of p_n.

(2) If $p \in \mathbb{R}$, $p' \in \mathbb{R}$, and if p_n converges to p and to p', then $p' = p$.

(3) If p_n converges, then p_n is bounded.

(4) If $E \subseteq \mathbb{R}$ and if p is an accumulation point of E, then there is a sequence p_n in E such that $p = \lim p_n$, $n \to +\infty$

Definition 2.10. If $\forall\ K \in \mathbb{R}$ there exists $n_0 \in \mathbb{N}$ such that $\forall\ n > n_0, n \in \mathbb{N}$ we have $a_n > K$ we say that a_n diverges to $+\infty$, and we write

$$\lim_{n \to +\infty} a_n = +\infty$$

or

$$a_n \to +\infty \qquad n \to +\infty \qquad \Diamond$$

Definition 2.11. Given a sequence $\{p_n\}$, consider a sequence $\{n_k\}$ of positive integers, such that $n_1 < n_2 < n_3 < \ldots$. The sequence $\{p_{n_k}\}$ is called a subsequence of $\{p_n\}$. If $\{p_{n_k}\}$ converges, its limit is called a subsequential limit of $\{p_n\}$. ◊

Example 2.12. $a_n = \{(-1)^n\}$ has two subsequential limits: 1 and -1. △

Theorem 2.13. *$\{p_n\}$ converges to p if and only if every subsequence of $\{p_n\}$ converges to p.*

Theorem 2.14. *Suppose s_n and t_n are real sequences, and*

$$\lim_{n\to+\infty} s_n = s, \qquad \lim_{n\to+\infty} t_n = t.$$

Then

(a) $\lim(s_n + t_n) = s + t$

(b) $\lim(cs_n) = cs$, $\lim(c + s_n) = c + s$, *for any number* c

(c) $\lim(s_n t_n) = st$

(d) $\lim(1/s_n) = 1/s$, *if* $s_n \neq 0$ *for every* n *and* $s \neq 0$.

Theorem 2.15. *Suppose s_n , t_n are real sequences, $s_n \leq t_n$ for $n > N$, where N is fixed, and*

$$\lim_{n\to+\infty} s_n = s, \qquad \lim_{n\to+\infty} t_n = t,$$

then $s \leq t$.

The previous theorem works also in the case of sequences divergent to infinity: if $s = +\infty$, then $t = +\infty$ and if $t = -\infty$, then $s = -\infty$

Theorem 2.16. *If $p > 0$, $\alpha \in \mathbb{R}$ and $|x| < 1$ then*

$$\lim_{n\to\infty} \frac{1}{n^p} = 0, \quad \lim_{n\to\infty} \sqrt[n]{p} = 1, \quad \lim_{n\to\infty} \sqrt[n]{n} = 1, \quad \lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0, \quad \lim_{n\to\infty} x^n = 0$$

Theorem 2.17. *The sequence:*

$$a_n = \left(1 + \frac{1}{n}\right)^n$$

is convergent. Its limit is a transcendental number denoted by e.

Boundedness is, in general, a necessary but not sufficient condition for convergence. However, there is an important case in which convergence is equivalent to boundedness, that of monotonic sequences.

Theorem 2.18 (*). *Suppose $\{a_n\}$ is monotonic. Then $\{a_n\}$ converges if and only if it is bounded.*

Proof. If $\{a_n\}$ converges, then it's bounded. Suppose then that $\{a_n\}$ is increasing and bounded. Let E be the range of $\{a_n\}$. E is bounded so it has a supremum, say α

$$\alpha = \sup_n a_n$$

Fix any $\epsilon > 0$. Then $\alpha - \epsilon$ is not an upper bound of E so that there is an integer N such that

$$\alpha - \epsilon < a_N < \alpha$$

Since a_n is increasing, then this in turn implies

$$\alpha - \epsilon < a_n < \alpha$$

for all $n > N$, which means a_n converges to α. The proof for the case a_n decreasing is similar. $\square$

Theorem 2.19. *If there exists N such that for $n > N$ we have*

$$\left|\frac{a_{n+1}}{a_n}\right| < r < 1$$

then $a_n \to 0$

Proof. If N is such that

$$\left|\frac{a_{n+1}}{a_n}\right| < r < 1$$

holds for $n > N$ then

$$|a_{N+1}| < r|a_N|$$
$$|a_{N+2}| < r|a_{N+1}| < r^2|a_N|$$
$$|a_{N+3}| < r|a_{N+2}| < r^3|a_N|$$

etc. Since $r < 1$, $|a_n| \to 0$ and by Theorem 2.15 $a_n \to 0$ □

Exercises on section 2.1

Exercise 2.1 (MC)**.** $a_n = \begin{cases} \frac{1}{n} & n \text{ even} \\ 3 & n \text{ odd} \end{cases}$

a) is convergent

b) is divergent

c) is bounded

d) is decreasing

Exercise 2.2 (MC)**.** If $a_0 = 4$ and $a_{n+1} = \frac{1}{3}a_n$, $\forall n$, then $\lim_{n\to+\infty} a_n =$

a) 4

b) 0

c) $\frac{1}{3}$

d) none of the others

Exercise 2.3 (MC)**.** $a_n = 2 \cdot \left(\frac{q}{2}\right)^n$, where $q \in \mathbb{R}$, diverges to $+\infty$ if and only if

a) $q \geq 2$

b) $q > 2$

c) $q \geq 1$

d) $q > 2$ or $q < -2$

Exercise 2.4 (TF)**.** Consider $a_n = n^2 - 10n$.

a) a_n is increasing

b) a_n is bounded

c) $a_n \geq 0$ for all n

d) $a_n \to +\infty$

Exercise 2.5 (TF)**.** Consider $a_n = e^{2n} + \sin n$.

a) $a_n \to +\infty$
b) a_n is bounded below
c) a_n is positive
d) a_n is strictly increasing

Exercise 2.6 (TF)**.** If $a_0 = 3$ and $a_{n+1} = -a_n/2$ then

a) $a_n = (-3)^n/2^n$
b) a_n is monotonic
c) a_n does not converge
d) a_n is bounded

Exercise 2.7. Let $n \geq 0$ and let

$$a_n = \begin{cases} 3^{-n} & n \text{ even} \\ \frac{3n-4}{2n} & n \text{ odd} \end{cases}$$

Determine the sup and inf of a_n and specify if they are also max and min.

Exercise 2.8 (TF)**.** Consider a sequence a_n. Find whether the following statements are true or not.

a) if a_n is increasing then $a_n \to +\infty$
b) if a_n is increasing and $b_n = a_n - \frac{1}{n}$ then b_n is strictly increasing
c) if a_n is increasing then a_n^2 is strictly increasing
d) if a_n is not bounded then $|a_n| \to \infty$

Exercise 2.9 (TF)**.** If $a_n < a_{n+1}\ \forall\ n$ then

a) $\lim a_n = +\infty$
b) a_n has no maximum
c) a_n is increasing
d) a_n is not bounded

Exercise 2.10. Let

$$a_0 = k > 1, \qquad a_{n+1} = \frac{a_n + 1}{2}$$

for $n \in \mathbb{N}$. Prove that if $a_n > 1$ then $a_{n+1} > 1$ and that the sequence is monotonic. Prove that the sequence converges and find its limit.

Exercise 2.11. Prove that if $a_n \to L \neq 0$, $L \in \mathbb{R}$, then

$$\lim_{n\to+\infty} \frac{a_{n+1}}{a_n} = 1$$

Find a counterexample showing that the conclusion is false if $L = +\infty$.

Exercise 2.12. Show that every sequence a_n such that $a_n \to +\infty$ has a minimum.

2.2. Asymptotic notation

In the following we consider two generic sequences a_n and b_n. However, in computer science, where sequences represent the computing time of an algorithm and n is the size of the input, a_n and b_n are both non negative and typically divergent to $+\infty$.

Definition 2.20. If

$$\lim_{n\to+\infty} \frac{a_n}{b_n} = 0$$

then we write a_n is $o(b_n)$ and say "a_n is little-o of b_n". $\lozenge$

In the most common case, this roughly means a_n goes to infinity slower than b_n.

Definition 2.21. If there exist $C \in \mathbb{R}$, $C > 0$, and $N \in \mathbb{N}$ such that

$$|a_n| \leq C|b_n|$$

for $n > N$ then we say "a_n is big-O of b_n".

$\lozenge$

In the most common case, this roughly means a_n is "bounded above" by b_n.

Note that once C and N have been found, there are infinitely many constants which can replace them, any $C' > C$ and any $N' > N$.

Example 2.22. To show that $a_n = n^2 + 2n + 1$ is $O(n^2)$ we might note that if $n > 1$ then we have $1 < n^2$ and $n < n^2$ which means

$$n^2 + 2n + 1 \leq n^2 + 2n^2 + n^2 = 4n^2$$

for $n > 1$. That is, $n^2 + 2n + 1$ is $O(n^2)$.

In a similar way, one can prove that $n^2 + 2n + 1$ is $O(n^3)$, because $n^2 < n^3$ for sufficiently large n. $\triangle$

Usually, b_n is chosen as small as possible, and of a "standard" sequence type, like n^k.

Example 2.23. To show that n^2 is not $O(n)$, we must show that no pair of numbers C and N exist such that $n^2 \leq Cn$ whenever $n > N$. We will use a proof by contradiction to show this.

Suppose that there are constants C and N for which $n^2 \leq Cn$ whenever $n > N$. Observe that when $n > 0$ we can divide both sides of the inequality $n^2 \leq Cn$ by n to obtain the equivalent inequality $n \leq C$. However, no matter what C and N are, the inequality $n \leq C$ cannot hold for all n with $n > N$. In particular, once we set a value of N, we see that when n is larger than the maximum of N and C, it is not true that $n \leq C$ even though $n > N$. This contradiction shows that n^2 is not $O(n)$. $\triangle$

Theorem 2.24 (*)**.** *Let $b_n = a_k n^k + a_{k-1} n^{k-1} + \cdots + a_1 n + a_0$, where $a_0, a_1, \ldots, a_k$ are real numbers. Then b_n is $O(n^k)$.*

Proof. Using the triangle inequality, if $n > 1$ we have

$$\begin{aligned} |b_n| &= \left|a_k n^k + a_{k-1}n^{k-1} + \cdots + a_1 n + a_0\right| \\ &\le |a_k| x^k + |a_{k-1}| n^{k-1} + \cdots + |a_1| n + |a_0| \\ &= n^k \left(|a_k| + |a_{k-1}|/n + \cdots + |a_1|/n^{k-1} + |a_0|/n^k\right) \\ &\le n^k \left(|a_k| + |a_{k-1}| + \cdots + |a_1| + |a_0|\right) \end{aligned}$$

This shows that

$$|b_n| \le Cn^k,$$

where

$$C = |a_k| + |a_{k-1}| + \cdots + |a_1| + |a_0|$$

and $N = 1$ □

Example 2.25. The sum of the first n positive integers

$$1 + 2 + 3 + \cdots + n$$

is $O(n^2)$ because every term in the sum is less than n so that

$$1 + 2 + 3 + \cdots + n \le n + n + \cdots + n = n^2$$

△

Example 2.26. $n!$ is $O(n^n)$ because

$$1 \cdot 2 \cdot \cdots \cdot n \le n \cdot n \cdot \cdots \cdot n = n^n.$$

Taking the logarithm we get $\ln(n!)$ is $O(n \ln n)$. △

Example 2.27. We have showed in exercise 1.6 that $n < 2^n$ for every n. Taking logarithms base 2 we get $\log_2 n < n$. Since all logarithms are proportional to one another via the relation

$$\log_a x = \frac{\log_b x}{\log_b a}.$$

we have

$$\log_2 n = \frac{\ln n}{\ln 2}.$$

and therefore $\ln n < (\ln 2)n$, that is $\ln n$ is $O(n)$. △

Theorem 2.28. *Suppose that a_n is $O(b_n)$ and that c_n is $O(d_n)$. Then $a_n + c_n$ is $O(\max(b_n, d_n))$.*

As a consequence we have that if a_n and b_n are both $O(c_n)$ then $a_n + b_n$ is $O(c_n)$.

Theorem 2.29. *Suppose that a_n is $O(b_n)$ and that c_n is $O(d_n)$. Then $a_n c_n$ is $O(b_n d_n)$.*

Example 2.30. To give an estimate of $a_n = 3n \ln(n!) + (n^2 + 3) \ln n$ we note that $\ln(n!)$ is $O(n \ln n)$ and therefore the first term is $O(n^2 \ln n)$. Then, $n^2 + 3$ is $O(n^2)$ and thus $(n^2 + 3) \ln n$ is $O(n^2 \ln n)$. The sequence a_n is the sum of two sequences that are both $O(n^2 \ln n)$ and therefore is $O(n^2 \ln n)$ as well. △

Definition 2.31. If there exist two positive constants $C \in \mathbb{R}$ and $N \in \mathbb{N}$ such that

$$|a_n| \geq C|b_n|$$

for $n > N$ then we say a_n is $\Omega(b_n)$. $\lozenge$

This roughly means a_n is "bounded below" by b_n. We have a_n is $O(b_n)$ if and only if b_n is $\Omega(a_n)$. In fact, if a_n is $O(b_n)$ then there exist C, N such that

$$|a_n| \leq C|b_n|$$

for all $n > N$. But then

$$|b_n| \geq D|a_n|$$

where $D = 1/C$.

Definition 2.32. If both $a_n = O(b_n)$ and $b_n = O(a_n)$ then we write $a_n = \Theta(b_n)$ $\lozenge$

This roughly means a_n and b_n "bind" each other (so none increases faster than the other) or that the two sequence are the same "order of magnitude".

Example 2.33. The sum of the first n positive integers is $\Theta(n^2)$. We already know it's $O(n^2)$ so we need only to prove it's $\Omega(n^2)$, that is we need to find a lower bound for the sum. In the sum

$$1 + 2 + 3 + \cdots + n$$

consider only the first half of the terms. Summing only the terms greater than $\lceil n/2 \rceil$ we find that

$$\begin{aligned} 1 + 2 + 3 + \cdots + n &\geq \lceil n/2 \rceil + (\lceil n/2 \rceil + 1) + \cdots + n \\ &\geq \lceil n/2 \rceil + \lceil n/2 \rceil + \cdots + \lceil n/2 \rceil \\ &= (n - \lceil n/2 \rceil + 1)\lceil n/2 \rceil \\ &\geq (n/2)(n/2) \\ &= n^2/4 \end{aligned}$$

which shows the sum is $\Omega(n^2)$ and therefore also $\Theta(n^2)$ $\triangle$

Remark 2.34. Every k degree polynomial is $\Theta(n^k)$.

Definition 2.35. If

$$\lim_{n \to +\infty} \frac{a_n}{b_n} = 0$$

then we write b_n is $\omega(a_n)$. $\lozenge$

This roughly means b_n goes to infinity faster than a_n.

Definition 2.36. If

$$\lim \frac{a_n}{b_n} = 1$$

we write $a_n \sim b_n$ $\lozenge$

Theorem 2.37. *If $a_n \sim c_n$ and $b_n \sim d_n$ then $a_n b_n \sim c_n d_n$ and $a_n/b_n \sim c_n/d_n$ provided that b_n and d_n are non zero for sufficiently large n.*

Remark 2.38. These are frequent mistakes about the $\sim$ relation.

- $a_n \sim b_n$ does not imply $a_n - b_n \to 0$. For example, $a_n = n$ and $b_n = n + \ln n$

- $a_n \sim b_n, c_n \sim d_n$ does not imply $a_n + c_n \sim b_n + d_n$. For example, $a_n = n + \sqrt{n}$, $b_n = n$, $c_n = -n + \sqrt{n}$, $d_n = -n$.

Theorem 2.39. $a_n + o(a_n) \sim a_n$.

Theorem 2.40. $a_n \sim b_n \iff a_n = b_n + o(b_n)$.

Exercises on sections 2.2

Exercise 2.13 (TF)**.** If $a_n \to 0$ and $b_n \to 0$ then:

a) $a_n \sim b_n$

b) $a_n = o(b_n)$

c) $a_n - b_n \to 0$

d) $a_n + b_n \to 0$

Exercise 2.14 (TF)**.** If $a_n \to 2$ then:

a) $a_n \sim 2$

b) $a_n = O(n)$

c) $a_n = o(1)$

d) $a_n - 2 = O(1)$

Exercise 2.15. Determine whether each of these functions is O(n) or not: $a_n = 10$, $a_n = 3n + 7$, $a_n = n^2 + n + 1$, $a_n = 5 \ln n$ and $a_n = \lceil n/2 \rceil$.

Exercise 2.16. What does it mean for a sequence a_n to be $O(1)$?

Exercise 2.17. What does it mean for a sequence a_n to be $\Theta(1)$?

Exercise 2.18. Show that $\ln(n^2 + 1)$ and $\log_2 n$ are of the same order.

Exercise 2.19. Find the least integer p such that

$$\frac{n^3 + 1}{n^2 + 1}$$

is $O(x^p)$.

2.3. Recurrences

Example 2.41 (Towers of Hanoi)**.** The Towers of Hanoi puzzle consists of three rods and a number of disks of different sizes, which can slide onto any rod. The puzzle starts with the disks in a neat stack in ascending order of size on one rod, the smallest at the top, thus making a conical shape.

The objective of the puzzle is to move the entire stack to another rod, obeying the following simple rules:

- Only one disk can be moved at a time.
- Each move consists of taking the upper disk from one of the stacks and placing it on top of another stack or on an empty rod.
- No larger disk may be placed on top of a smaller disk.

With 1 disk, the puzzle can be (obviously) solved in 1 move. Say there are n disks on rod 1. We define the number of moves needed to solve the puzzle, that is to move the entire stack onto another rod, say rod 3, as T_n. To move n disks to rod 3, a viable strategy would be to:

(1) move $n-1$ disks to rod 2,

(2) move the bottom (largest) disk from rod 1 to rod 3

(3) move the $n-1$ disks on rod 2 to rod 3.

The number of required moves according to this strategy is

(1) T_{n-1}

(2) 1

(3) T_{n-1}.

so in the end we get $T_n = 2T_{n-1} + 1$. Thus we want to solve the recurrence

$$\begin{aligned} T_1 &= 1 \\ T_n &= 2T_{n-1} + 1. \end{aligned}$$

By simple calculations we get $T_2 = 3$, $T_3 = 7$, $T_4 = 15$ and we make a guess: $T_n = 2^n - 1$. To prove the guess is correct, we use mathematical induction. Since a base case is already proved, only the inductive step is to be checked. Thus, we assume $T_n = 2^n - 1$ and try to prove that $T_{n+1} = 2^{n+1} - 1$. Using recursion we have

$$T_{n+1} = 2T_n + 1 = 2\left(2^n - 1\right) + 1 = 2^{n+1} - 2 + 1 = 2^{n+1} - 1$$

The conclusion is proved. $\triangle$

Example 2.42 (Towers of Hanoi 2)**.** Again with the Towers of Hanoi problem

$$\begin{aligned} T_1 &= 1 \\ T_n &= 2T_{n-1} + 1. \end{aligned}$$

Because $T_{n-1} = 2T_{n-2} + 1$, we have

$$T_n = 2T_{n-1} + 1 = 2\left(2T_{n-2} + 1\right) + 1 = 4T_{n-2} + 2 + 1$$

Then, again, $T_{n-2} = 2T_{n-3} + 1$ which leads to

$$T_n = 4T_{n-2} + 2 + 1 = 4\left(2T_{n-3} + 1\right) + 2 + 1 = 8T_{n-3} + 4 + 2 + 1$$

Since $1 + 2 + \cdots + 2^{k-1} = 2^k - 1$, we can now make a guess:

$$T_n = 2^k T_{n-k} + 2^k - 1 \tag{2.1}$$

that must be proved by induction on k. In other words, assuming

$$P(k) : T_n = 2^k T_{n-k} + 2^k - 1$$

is true, we want to prove

$$P(k+1) : T_n = 2^{k+1} T_{n-k-1} + 2^{k+1} - 1$$

is also true. Since

$$T_{n-k} = 2T_{n-k-1} + 1$$

we have

$$\begin{aligned} T_n &= 2^k\left(2T_{n-k-1}+1\right)+2^k-1 \\ &= 2^{k+1}2T_{n-k-1}+2^k+2^k-1 \\ &= 2^{k+1}2T_{n-(k+1)}+2^{k+1}-1 \end{aligned}$$

So the guess is correct. Now, we can go back to the start of the recurrence since we know what's T_1. We get T_1 when $k = n-1$ in the recursion definition. Then

$$\begin{aligned} T_n &= 2^{n-1}T_1+1+2+\cdots+2^{n-2} \\ &= 2^{n-1}+2^{n-1}-1 \\ &= 2^n-1 \end{aligned}$$

which is the final result. △

Example 2.43. Find a recurrence relation and give initial conditions for the number of bit strings of length n that do not have two consecutive 0s. How many such bit strings are there of length five?

Let a_n denote the number of bit strings of length n that do not have two consecutive 0s. To obtain a recurrence relation for a_n, note that by the sum rule, the number of bit strings of length n that do not have two consecutive 0s equals the number of such bit strings ending with a 0 plus the number of such bit strings ending with a 1. We will assume that $n \geq 3$, so that the bit string has at least three bits.

The bit strings of length n ending with 1 that do not have two consecutive 0s are precisely the bit strings of length $n-1$ with no two consecutive 0s with a 1 added at the end. Consequently, there are a_{n-1} such bit strings. Bit strings of length n ending with a 0 that do not have two consecutive 0s must have 1 as their $(n-1)$–st bit; otherwise they would end with a pair of 0s. It follows that the bit strings of length n ending with a 0 that have no two consecutive 0s are precisely the bit strings of length $n-2$ with no two consecutive 0s with 10 added at the end. Consequently, there are a_{n-2} such bit strings. We conclude that

$$a_n = a_{n-1} + a_{n-2}$$

for $n \geq 3$. The initial conditions are $a_1 = 2$, because both bit strings of length one, 0 and 1 do not have consecutive 0s, and $a_2 = 3$, because the valid bit strings of length two are 01, 10, and 11. To obtain a_5, we use the recurrence relation three times to find that $a_5 = 13$. △

Example 2.44. We want to find a recurrence relation for C_n, the number of ways to parenthesize the product of $n+1$ numbers,

$$x_0 \cdot x_1 \cdot x_2 \cdot \cdots \cdot x_n,$$

to specify the order of multiplication. For example, $C_3 = 5$ because there are five ways to parenthesize $x_0 \cdot x_1 \cdot x_2 \cdot x_3$ to determine the order of multiplication:

$$\begin{gathered}((x_0 \cdot x_1) \cdot x_2) \cdot x_3 \\ (x_0 \cdot (x_1 \cdot x_2)) \cdot x_3 \\ (x_0 \cdot x_1) \cdot (x_2 \cdot x_3) \\ x_0 \cdot ((x_1 \cdot x_2) \cdot x_3) \\ x_0 \cdot (x_1 \cdot (x_2 \cdot x_3)).\end{gathered}$$

To develop a recurrence relation for C_n, we note that however we insert parentheses in the product $x_0 \cdot x_1 \cdot x_2 \cdot \cdots \cdot x_n$, one "$\cdot$" operator remains outside all parentheses, namely, the operator for the final multiplication to be performed. For example, in $(x_0 \cdot (x_1 \cdot x_2)) \cdot x_3$, it is the final "$\cdot$", while in $(x_0 \cdot x_1) \cdot (x_2 \cdot x_3)$ it is the second "$\cdot$". This final operator appears between two of the $n+1$ numbers, say, x_k and x_{k+1}. There are $C_k C_{n-k-1}$ ways to insert parentheses to determine the order of the $n+1$ numbers to be multiplied when the final operator appears between x_k and x_{k+1}, because there are C_k ways to insert parentheses in the product $x_0 \cdot x_1 \cdot x_2 \cdot \cdots \cdot x_k$ to determine the order in which these $k+1$ numbers are to be multiplied and C_{n-k-1} ways to insert parentheses in the product $x_{k+1} \cdot x_{k+2} \cdot \cdots \cdot x_n$ to determine the order in which these $n-k$ numbers are to be multiplied. Because this final operator can appear between any two of the $n+1$ numbers, it follows that

$$C_n = C_0 C_{n-1} + C_1 C_{n-2} + \cdots + C_{n-2} C_1 + C_{n-1} C_0 = \sum_{k=0}^{n-1} C_k C_{n-k-1}$$

Note that we have $C_0 = 1$ because with just x_0 there are no parentheses. We also have $C_1 = 1$ because with $x_0 \cdot x_1$ again there are no parentheses. The numbers C_n are called the Catalan numbers.

C_n is also the number of monotonic lattice paths along the edges of a grid with $n \times n$ square cells, which do not pass above the diagonal. A monotonic path is one which starts in the lower left corner, finishes in the upper right corner, and consists entirely of edges pointing rightwards or upwards.

△

Example 2.45 (Merge-Sort)**.** Let T_n be the maximum number of comparisons used while Merge Sorting a list of n numbers. For now, assume that n is a power of 2. This ensures that the input can be divided in half at every stage of the recursion.

If there is only one number in the list, then no comparisons are required, so $T_1 = 0$.

Otherwise, T_n includes comparisons used in sorting the first half (at most $T_{n/2}$), in sorting the second half (also at most $T_{n/2}$), and in merging the two halves. The number of comparisons in the merging step is at most $n-1$. This is because at least one number is emitted after each comparison and one more number is emitted at the end when one list becomes empty. Since n items are emitted in all, there can be at most $n-1$ comparisons.

Therefore, the maximum number of comparisons needed to Merge Sort n items is given by this recurrence:

$$\begin{aligned} &T_1 = 0 \\ &T_n = 2T_{n/2} + n - 1 \qquad \text{for } n \geq 2 \text{ and a power of } 2 \end{aligned}$$

This fully describes the number of comparisons, but not in a very useful way; a closed-form expression would be much more helpful. To get that, we have to solve the recurrence.

Trying to guess is useless:

$$\begin{aligned} T_1 &= 0 \\ T_2 &= 2T_1 + 2 - 1 = 1 \\ T_4 &= 2T_2 + 4 - 1 = 5 \\ T_8 &= 2T_4 + 8 - 1 = 17 \\ T_{16} &= 2T_8 + 16 - 1 = 49 \end{aligned}$$

On the contrary, back substitution seems to work:

$$\begin{aligned} T_n &= 2T_{n/2} + n - 1 \\ &= 2(2T_{n/4} + n/2 - 1) + n - 1 \\ &= 4T_{n/4} + 2n - 3 \\ &= 4(2T_{n/8} + n/4 - 1) + 2n - 3 \\ &= 8T_{n/8} + 3n - 7 \\ &= 8(2T_{n/16} + n/8 - 1) + 3n - 7 \\ &= 16T_{n/16} + 4n - 15 \end{aligned}$$

and it looks as if

$$T_n = 2^k T_{n/2^k} + kn - (2^k - 1)$$

This is in fact true, since taking one more step back we get

$$\begin{aligned} T_n &= 2^k T_{n/2^k} + kn - (2^k - 1) \\ &= 2^k (T_{n/2^{k+1}} + n/2^k - 1) + kn - 2^k + 1 \\ &= 2^{k+1} T_{n/2^{k+1}} + n - 2^k + kn - 2^k + 1 \\ &= 2^{k+1} T_{n/2^{k+1}} + (k+1)n - 2^{k+1} + 1 \end{aligned}$$

which is exactly the above formula with $k+1$ in place of k. Then we can now go back to get a closed form expression for T_n. Taking $k = \log_2 n$ we have $T_{n/2^k} = T_1 = 0$. Then:

$$\begin{aligned} T_n &= 2^k T_{n/2^k} + kn - (2^k - 1) \\ &= 2^{\log_2 n} T_{n/2^{\log_2 n}} + n \log_2 n - (2^{\log_2 n} - 1) \\ &= nT_1 + n \log_2 n - n + 1 \\ &= n \log_2 n - n + 1 \end{aligned}$$

We then see that Merge Sort execution time is $O(n \ln n)$. Note that the gain in performance over a $O(n^2)$ sorting algorithm is about $n/\log_2 n$. With $n = 2^{20} = 1048576$ the ratio is of the order of 10^5. $\triangle$

Exercises on section 2.3

Exercise 2.20. Find a recurrence relation for the number of strictly increasing sequences of positive integers that have 1 as their first term and n as their last term, where n is a positive integer. That is, sequences $a_1, a_2, \ldots, a_k$, where $a_1 = 1$, $a_k = n$, and $a_j < a_{j+1}$ for $j = 1, 2, \ldots, k-1$. What are the initial conditions? How many sequences of the type described above are there when n is an integer with $n \geq 2$?

Exercise 2.21. Find a recurrence relation for the number of bit strings of length n that contain a pair of consecutive 0s. What are the initial conditions? How many bit strings of length seven contain two consecutive 0s?

Exercise 2.22. Find a recurrence relation for the number of bit strings of length n that contain three consecutive 0s. What are the initial conditions? How many bit strings of length seven contain three consecutive 0s?

Exercise 2.23. Find a recurrence relation for the number of bit strings of length n that do not contain three consecutive 0s. What are the initial conditions? How many bit strings of length seven contain do not contain three consecutive 0s?

Exercise 2.24. Find a recurrence relation for the number of ternary strings of length n that do not contain two consecutive 0s. What are the initial conditions? How many bit strings of length seven contain do not contain two consecutive 0s? *(Ternary strings are strings made with 0, 1 and 2.)*

Exercise 2.25. A line separates the plane into two regions. Two intersecting lines separate the plane into four regions. Suppose that we have n lines in general position, that is no two are parallel and no three lines intersect in the same point. Into how many regions do these lines divide the plane?

Exercise 2.26. Using the results in the previous exercises, find the number of bit strings of length eight that contain either three consecutive 0s or four consecutive 1s. This is an alternative solution to 1.22

2.4. Linear recurrences

Definition 2.46. A linear recurrence of degree (or order) k is a recurrence relation of the form

$$a_n = c_1(n)a_{n-1} + c_2(n)a_{n-2} + \cdots + c_k(n)a_{n-k} + g(n),$$

where $c_1(n), c_2(n), \ldots, c_k(n)$ are functions of n, and $c_k \neq 0$. The recurrence relation is homogeneous whenever $g(n) = 0$. ◊

The recurrence relation in the definition is linear because the right-hand side is a sum of previous terms of the sequence each multiplied by a function of n (that is, the recurrence is linear with respect to the a_j). The coefficients might possibly be all constants, rather than functions that depend on n. The degree is k because a_n is expressed in terms of the previous k terms of the sequence. For example, for the Towers of Hanoi problem we have found a recurrence which is linear but not homogeneous for the presence of 1.

Before proving the main theorem for solving linear homogeneous recurrences in the two dimensional case we need an auxiliary result.

Theorem 2.47. *[*] Let A and B be real numbers. A recurrence relation of the form*

$$a_n = Aa_{n-1} + Ba_{n-2},$$

for all integers $n \geq 2$ is satisfied by the sequence

$$1, t, t^2, t^3, \ldots, t^n, \ldots$$

where t is a nonzero real number, if, and only if, t satisfies the equation $t^2 - At - B = 0$.

Proof. Suppose that for some number $t \neq 0$, the sequence

$$1, t, t^2, t^3, \ldots, t^n, \ldots$$

satisfies the recurrence. This means that each term of the sequence equals A times the previous term plus B times the term before that. So for all integers $k \geq 2$

$$t^k = At^{k-1} + Bt^{k-2}.$$

In particular, when $k = 2$, the equation becomes $t^2 = At + B$, or equivalently, $t^2 - At - B = 0$.

On the other side, suppose t is any non zero root of the equation. If we multiply the equation by t^{k-2} we obtain

$$t^{k-2}t^2 = At^{k-2}t + Bt^{k-2}.$$

that is

$$t^k = At^{k-1} + Bt^{k-2}.$$

Hence the sequence $1, t, t^2, t^3, \ldots, t^n, \ldots$ satisfies the recurrence. □

Theorem 2.48 (Two–dimensional version,*)**.** *Let c_1, c_2 be real numbers. Suppose the equation*

$$r^k = c_1 r^{k-1} + c_2 r^{k-2}$$

has 2 distinct roots r_1, r_2. Then a sequence a_n is a solution of the recurrence relation

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

if and only if

$$a_n = \alpha_1 r_1^n + \alpha_2 r_2^n$$

for every n, where α_1, α_2 are constants to be determined using the initial conditions on a_j.

Proof. Suppose that for some real numbers A and B, a sequence a_n satisfies the recurrence relation

$$a_n = Aa_{n-1} + Ba_{n-2},$$

for all integers $n \geq 2$, and suppose the characteristic equation $r^2 - Ar - B = 0$ has two distinct roots r and s. We will show that for all integers $n \geq 0$,

$$a_n = Cr^n + Ds^n,$$

where C and D are numbers such that

$$a_0 = Cr^0 + Ds^0, \quad \text{and} \quad a_1 = Cr^1 + Ds^1.$$

Let $P(n)$ be the equation

$$a_n = Cr^n + Ds^n$$

We use strong mathematical induction to prove that $P(n)$ is true for all integers $n \geq 0$. In the base step, we prove that $P(0)$ and $P(1)$ are true. We do this because in the inductive step we need the equation to hold for $n = 0$ and $n = 1$ in order to prove that it holds for $n = 2$.

The truth of $P(0)$ and $P(1)$ is automatic because C and D are exactly those numbers that make the following equations true

$$a_0 = Cr^0 + Ds^0, \quad \text{and} \quad a_1 = Cr^1 + Ds^1.$$

Suppose that $k \geq 1$ and for all integers i from 0 through k,

$$a_i = Cr^i + Ds^i.$$

We must show that $P(k+1)$ holds:

$$a_{k+1} = Cr^{k+1} + Ds^{k+1}.$$

Now by the inductive hypothesis

$$a_k = Cr^k + Ds^k, \quad \text{and} \quad a_{k-1} = Cr^{k-1} + Ds^{k-1}.$$

so

$$\begin{aligned} a_{k+1} &= Aa_k + Ba_{k-1} \\ &= A(Cr^k + Ds^k) + B(Cr^{k-1} + Ds^{k-1}) \\ &= C(Ar^k + Br^{k-1}) + D(As^k + Bs^{k-1}) \\ &= Cr^{k+1} + Ds^{k+1}). \end{aligned}$$

This is what was to be shown. The reason the last equality follows from Theorem 2.47 is that since r and s satisfy the characteristic equation, the sequences r^n and s^n satisfy the recurrence relation. □

Theorem 2.49 (Multi–dimensional version)**.** *Let $c_1, c_2, \ldots, c_k$ be real numbers. Suppose the equation*

$$r^k = c_1 r^{k-1} + c_2 r^{k-2} + \cdots + c_k$$

has k distinct roots $r_1, \ldots, r_k$. Then a sequence a_n is a solution of the recurrence relation

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k},$$

if and only if

$$a_n = \alpha_1 r_1^n + \alpha_2 r_2^n + \cdots + \alpha_k r_k^n$$

for every n, where $\alpha_1, \ldots, \alpha_k$ are constants. The constants are to be determined using the initial conditions on a_j.

Definition 2.50. The equation

$$r^k = c_1 r^{k-1} + c_2 r^{k-2} + \cdots + c_k$$

is called the characteristic equation of the linear recurrence. ♢

Example 2.51. From the recurrence in example 2.43 we have:

$$\begin{aligned} a_1 &= 2 \\ a_2 &= 3 \\ a_n &= a_{n-1} + a_{n-2} \end{aligned}$$

Setting $a_n = x^n$ and simplifying by x^{n-2} gets $x^2 = x + 1$. Solving for x we get

$$x_1 = \frac{1+\sqrt{5}}{2}, x_2 = \frac{1-\sqrt{5}}{2}.$$

Setting

$$\frac{1+\sqrt{5}}{2} = \phi$$

we get

$$x_2 = \frac{1-\sqrt{5}}{2} = -\frac{1}{\phi}$$

so that the general solution to the recurrence is

$$a_n = c_1\phi^n + c_2\left(-\frac{1}{\phi}\right)^n$$

This only tells the dynamics of the sequence and we still need to find c_1 and c_2 so that we get the correct starting point. We use the initial conditions:

$$a_1 = c_1\phi + c_2\left(-\frac{1}{\phi}\right)$$

and

$$a_2 = c_1\phi^2 + c_2\left(-\frac{1}{\phi}\right)^2.$$

Solving for c_1, c_2 we finally get

$$c_1 = \frac{3\phi+2}{3\phi+1}, c_2 = -\frac{1}{3\phi+1}$$

and the general solution is

$$a_n = \frac{3+\sqrt{5}}{2\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n + \frac{\sqrt{5}-3}{2\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

△

Theorem 2.52. *Let c_1 and c_2 be real numbers. Suppose that $r^2 - c_1r - c_2 = 0$ has only one root r_0. Then the sequence a_n is a solution of the recurrence relation $a_n = c_1a_{n-1} + c_2a_{n-2}$ if and only if $a_n = \alpha_1r^n + \alpha_2nr^n$ for every n, where α_1 and α_2 are constants.*

Example 2.53. To find the solution of the recurrence relation $a_n = 6a_{n-1} - 9a_{n-2}$ with initial conditions $a_0 = 1$ and $a_1 = 6$ we build the characteristic equation

$$r^2 - 6r + 9 = 0.$$

The equation has only one root, $r = 3$. Hence, the solution to this recurrence relation is

$$a_n = \alpha_1 3^n + \alpha_2 n 3^n$$

for some constants α_1 and α_2. Using the initial conditions, it follows that

$$a_0 = \alpha_1 \qquad a_1 = 3\alpha_1 + 3\alpha_2$$

that is, $\alpha_1 = 1$ and $\alpha_2 = 1$. Consequently, the solution to this recurrence relation and the initial conditions is

$$a_n = 3^n + n3^n$$

$\triangle$

This method naturally extends to equations of degree more than two.

Example 2.54. To find the solution of the recurrence relation $a_n = 6a_{n-1} - 11a_{n-2} + 6a_{n-3}$ with initial conditions $a_0 = 2$, $a_1 = 5$ and $a_2 = 15$ we build the characteristic equation

$$r^3 - 6r^2 + 11r - 6 = 0.$$

which has three roots, 1, 2 and 3. Hence, the solution to this recurrence relation is

$$a_n = \alpha_1 1^n + \alpha_2 2^n + +\alpha_3 3^n$$

for some constants α_1, α_2 and α_3. Using the initial conditions, it follows that

$$a_0 = \alpha_1 + \alpha_2 + \alpha_3 \qquad a_1 = \alpha_1 + 2\alpha_2 + 3\alpha_3 \qquad a_2 = \alpha_1 + 4\alpha_2 + 9\alpha_3$$

that is, $\alpha_1 = 1$, $\alpha_2 = -1$ and $\alpha_3 = 2$. Consequently, the solution to this recurrence relation and the initial conditions is

$$a_n = 1 - 2^n + 2 \cdot 3^n$$

$\triangle$

Definition 2.55. The Fibonacci numbers, $f_0, f_1, f_2, \ldots$, are defined by the equations $f_0 = 0, f_1 = 1$, and

$$f_n = f_{n-1} + f_{n-2}$$

for $n = 2, 3, 4, \ldots$ $\Diamond$

The characteristic equation is $r^2 = r + 1$ and the general solution to the Fibonacci recurrence is

$$f_n = c_1 \left(\frac{1+\sqrt{5}}{2}\right)^n + c_2 \left(\frac{1-\sqrt{5}}{2}\right)^n$$

From

$$f_0 = c_1 + c_2, f_1 = c_1 \frac{1+\sqrt{5}}{2} + c_2 \frac{1-\sqrt{5}}{2}$$

and the initial conditions we get

$$c_1 + c_2 = 1, c_1 \frac{1+\sqrt{5}}{2} + c_2 \frac{1-\sqrt{5}}{2} = 1$$

and solving for c_1, c_2 we finally get

$$c_1 = \frac{1+\sqrt{5}}{2\sqrt{5}}, c_2 = \frac{\sqrt{5}-1}{2\sqrt{5}}$$

and the general solution is

$$f_n = \frac{1+\sqrt{5}}{2\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n + \frac{\sqrt{5}-1}{2\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^n$$

Theorem 2.56. *If $a_n^{(p)}$ is a particular solution of the nonhomogeneous linear recurrence relation with constant coefficients*

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} + g(n)$$

then every solution is of the form $a_n^{(p)} + a_n^{(h)}$ where $a_n^{(h)}$ is a solution of the associated homogeneous recurrence relation

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k}$$

and $a_n^{(p)}$ is any solution of the nonhomogeneous recurrence.

In view of this theorem, given the solution to the homogeneous equation, the only problem is to find the particular solution. There are many practical rules about this that go beyond the scope of these notes. In general, this particular solution is usually searched among functions of type similar to that of $g(n)$. For example, if $g(n)$ is a constant term, we search for constant sequences, if $g(n)$ is a polynomial of degree p then we search for polynomials of degree p, etc.

Example 2.57. The Towers of Hanoi problem is a linear non–homogeneous recurrence.

In the Towers of Hanoi, we first consider the homogeneous equation, that is $T_n = 2T_{n-1}$. Using the method above for homogeneous equations we get $T_n = c_1 2^n$. Since the $g(n)$ term is 1, a constant term, we look for a constant solution, that is $T_n = q$. Substituting into the recurrence gets

$$q = 2q + 1$$

whose only solution is $q = -1$. Then the general solution to the recurrence is $T_n = c_1 2^n - 1$. To find c_1 we consider $T_1 = 1$ from the definition. We have $T_1 = 2c_1 - 1 = 1$ which gives $c_1 = 1$. △

Example 2.58. To find all solutions of the recurrence relation $a_n = 3a_{n-1} + 2n$ with $a_1 = 3$ we first find a particular solution.

Because $g(n) = 2n$ is a polynomial in n of degree one, a reasonable trial solution is a linear function in n, say, $p_n = cn+d$, where c and d are constants. To determine whether there are any solutions of this form, suppose that $p_n = cn + d$ is such a solution. Then the equation $a_n = 3a_{n-1}+2n$ becomes $cn+d = 3(c(n-1)+d)+2n$. Simplifying and combining like terms gives $(2 + 2c)n + (2d - 3c) = 0$. It follows that $cn + d$ is a solution if and only if $2 + 2c = 0$ and $2d - 3c = 0$, that is, $cn + d$ is a solution if and only if $c = -1$ and $d = -3/2$. Consequently, $a_n^{(p)} = -n - 3/2$ is a particular solution.

For the solution to the homogeneous equation we note that the recurrence $a_n = 3a_{n-1}$ has $a_n^{(h)} = \alpha \cdot 3^n$ as a solution (α is a constant.) By theorem 2.56 all solutions are of the form

$$a_n^{(p)} + a_n^{(h)} = \alpha \cdot 3^n - n - 3/2$$

where α is a constant. To find the solution with $a_1 = 3$, let $n = 1$ in the formula we obtained for the general solution. We find that $3 = 3\alpha - 1 - 3/2$, which implies that $\alpha = 11/6$. The solution to the original recurrence is then

$$a_n = \frac{11}{6} \cdot 3^n - n - 3/2$$

△

Theorem 2.59 (*). *Let s_n be a sequence that satisfies a recurrence relation of the form*

$$s_{2n} = 2s_n + f(n)$$

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

$$s_{2^m} = 2^m \left(s_1 + \frac{1}{2} \sum_{i=0}^{m-1} \frac{f(2^i)}{2^i} \right)$$

for $m \in \mathbb{N}$. In particular, if

$$s_{2n} = 2s_n + A + Bn$$

for constants A and B, then

$$s_{2^m} = 2^m s_1 + (2^m - 1)A + \frac{B}{2} m 2^m$$

Thus, if $n = 2^m$ in this case, we have

$$s_n = ns_1 + (n-1)A + \frac{B}{2} n \log_2 n$$

Proof. The proof is by induction. We verify that if

$$s_{2^m} = 2^m \left(s_1 + \frac{1}{2} \sum_{i=0}^{m-1} \frac{f(2^i)}{2^i} \right)$$

for some $m \in \mathbb{N}$, then

$$s_{2^{m+1}} = 2^{m+1} \left(s_1 + \frac{1}{2} \sum_{i=0}^{m} \frac{f(2^i)}{2^i} \right)$$

Since $s_{2^0} = 2^0 s_1$, because there are no terms in the summation, the base case is true.

$$\begin{aligned}
s_{2^{m+1}} &= s_{2 \cdot 2^m} \\
&= 2s_{2^m} + f(2^m) \\
&= 2^{m+1} \left(s_1 + \frac{1}{2} \sum_{i=0}^{m-1} \frac{f(2^i)}{2^i} \right) + f(2^m) \\
&= 2^{m+1} \left(s_1 + \frac{1}{2} \sum_{i=0}^{m-1} \frac{f(2^i)}{2^i} + \frac{f(2^m)}{2 \cdot 2^m} \right) \\
&= 2^{m+1} \left(s_1 + \frac{1}{2} \sum_{i=0}^{m} \frac{f(2^i)}{2^i} \right)
\end{aligned}$$

The special case $f(n) = A + Bn$ gives the following summation

$$\begin{aligned} s_{2^m} &= 2^m \left(s_1 + \frac{1}{2} \sum_{i=0}^{m-1} \frac{A + B2^i}{2^i} \right) \\ &= 2^m s_1 + 2^{m-1} \sum_{i=0}^{m-1} \left(\frac{A}{2^i} + B \right) \\ &= 2^m s_1 + \sum_{i=0}^{m-1} A2^{m-1-i} + 2^{m-1} \sum_{i=0}^{m-1} B \\ &= 2^m s_1 + A\frac{2^m - 1}{2 - 1} + 2^{m-1} mB \\ &= 2^m s_1 + A(2^m - 1) + m2^m \frac{B}{2}. \end{aligned}$$

□

Example 2.60. Finding the largest member of a set by dividing and conquering leads to the recurrence

$$T(2n) = 2T(n) + A,$$

where the constant A is the time it takes to compare the winners from the two halves of the set. According to Theorem 2.59 with $B = 0$,

$$T(2^m) = 2^m T(1) + (2^m - 1)A.$$

Here $T(1)$ is the time it takes to find the biggest element in a set with one element. It seems reasonable to regard $T(l)$ as an overhead or handling charge for examining each element, whereas A is the cost of an individual comparison. If $n = 2^m$, we get

$$T(n) = nT(1) + (n - 1)A,$$

so T is $O(n)$. In this algorithm, we examine n elements and make $n-1$ comparisons to find the largest one. Thus, dividing and conquering gives no essential improvement over simply running through the elements one at a time, at each step keeping the largest number seen so far. △

Example 2.61. In the Merge Sort we get to

$$T(2n) = 2T(n) + Bn,$$

since the time to fit together the two ordered halves into one set is proportional to the sizes of the two sets being merged. Theorem 2.59 with $A = 0$ says that

$$T(2^m) = 2^m T(1) + \frac{B}{2} 2^m \cdot m.$$

If $n = 2^m$, we get

$$T(n) = nT(1) + \frac{B}{2} n \log_2 n,$$

so $T(n)$ is $O(n \log_2 n)$. For large n, the biggest part of the cost comes from the merging, not from examining individual elements. This algorithm is actually a reasonably efficient one for many applications. △

Exercises on section 2.4

Exercise 2.27. Solve the recurrence $a_n = a_{n-1} + 2a_{n-2}$, $a_0 = 2$, $a_1 = 7$

Exercise 2.28. Solve the following recurrence

$$a_n = a_{n-1} + 6a_{n-2}$$

with $a_0 = 3$ and $a_1 = 6$

Exercise 2.29. Solve the following recurrence

$$a_n = 7a_{n-1} - 10a_{n-2}$$

with $a_0 = 2$ and $a_1 = 1$

Exercise 2.30. Solve the following recurrence

$$a_n = 2a_{n-1} - a_{n-2}$$

with $a_0 = 4$ and $a_1 = 1$

Exercise 2.31. Solve the following recurrence

$$a_n = 7a_{n-2} + 6a_{n-3}$$

with $a_0 = 9$, $a_1 = 10$ and $a_2 = 32$

Exercise 2.32. Consider the nonhomogeneous linear recurrence relation

$$a_n = 3a_{n-1} + 2^n$$

Show that $a_n = -2^{n+1}$ is a solution of this recurrence relation, find all solutions of this recurrence relation and find the solution with $a_0 = 1$.

Exercise 2.33. In exercise 2.21 we found that the recurrence relation for the number of bit strings of length n that contain a pair of consecutive 0s is

$$a_n = a_{n-1} + a_{n-2} + 2^{n-2}$$

with the initial conditions $a_0 = 0$ and $a_1 = 0$. What is the solution for this recursion?

Exercise 2.34. Solve the recurrence in exercise 2.25

Exercise 2.35. Solve the recurrence

$$s_{2n} = 2s_n + n^2$$

for $n = 2^m$.

2.5. Finite Sums

The sum of the terms $a_m, a_{m+1}, \ldots, a_n$ from the sequence a_n can be written as

$$\sum_{j=m}^{n} a_j$$

Here, the variable j is called the index of summation, and the choice of the letter j as the variable is arbitrary; that is, we could have used any other letter, such as i or k. Here, the index of summation runs through all integers starting with its lower limit m and ending with its upper limit n. A large uppercase Greek letter sigma, Σ, is used to denote summation. The usual laws for arithmetic apply to summations. For example, when $a, b \in \mathbb{R}$ and $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$, we have

$$\sum_{j=1}^{n}(ax_j + by_j) = a\sum_{j=1}^{n} x_j + b\sum_{j=1}^{n} y_j$$

Sometimes it is useful to shift the index of summation in a sum. This is often done when two sums need to be added but their indices of summation do not match. When shifting an index of summation, it is important to make the appropriate changes in the corresponding summand.

Example 2.62. Suppose we have the sum

$$\sum_{j=1}^{5} j^2$$

but want the index of summation to run between 0 and 4 rather than from 1 to 5. To do this, we let $k = j - 1$. Then the new summation index runs from 0 (because $k = 1 - 0 = 0$ when $j = 1$) to 4 (because $k = 5 - 1 = 4$ when $j = 5$), and the term j^2 becomes $(k+1)^2$. Hence,

$$\sum_{j=1}^{5} j^2 = \sum_{k=0}^{4}(k+1)^2 \qquad \triangle$$

Double summations arise in many contexts (as in the analysis of nested loops in computer programs). An example of a double summation is

$$\sum_{i=1}^{4}\sum_{j=1}^{3} ij$$

To evaluate the double sum, consider first the inner summation. The counter is j so i is a constant inside the summation and can be factored out to get

$$\begin{aligned}
\sum_{i=1}^{4}\sum_{j=1}^{3} ij &= \sum_{i=1}^{4} i \sum_{j=1}^{3} j \\
&= \sum_{i=1}^{4} i \frac{3(3+1)}{2} \\
&= 6\sum_{i=1}^{4} i \\
&= 6\frac{4(4+1)}{2} \\
&= 60.
\end{aligned}$$

We can also use summation notation to add all values of a function, or terms of an indexed set, where the index of summation runs over all values in a set. That is, we write

$$\sum_{s \in S} f(s)$$

to represent the sum of the values $f(s)$, for all members s of S.

Example 2.63. To find

$$\sum_{k=50}^{100} k^2$$

first note that because

$$\sum_{k=1}^{100} k^2 = \sum_{k=1}^{49} k^2 + \sum_{k=50}^{100} k^2$$

we have

$$\sum_{k=1}^{100} k^2 - \sum_{k=1}^{49} k^2 = \sum_{k=50}^{100} k^2$$

We know that

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

and therefore

$$\sum_{k=50}^{100} k^2 = \frac{100 \cdot 101 \cdot 201}{6} - \frac{49 \cdot 50 \cdot 99}{6} = 297925 \qquad \triangle$$

If we think of $\sum_{i=0}^{n-1} x^i$ as a function of x we have

$$f(x) = \sum_{i=0}^{n-1} x^i = \frac{1-x^n}{1-x}$$

and we can take the derivative:

$$f'(x) = \sum_{i=0}^{n-1} ix^{i-1} = \left(\frac{1-x^n}{1-x}\right)' = \frac{1-nx^{n-1}+(n-1)x^n}{(x-1)^2}$$

Since the sum starts with 0, we have

$$f'(x) = \sum_{i=1}^{n-1} ix^{i-1} = \frac{1-nx^{n-1}+(n-1)x^n}{(x-1)^2}$$

Multiplying by x we get

$$xf'(x) = \sum_{i=1}^{n-1} ix^i = \frac{x-nx^n+(n-1)x^{n+1}}{(x-1)^2} = x + 2x^2 + 3x^3 + \cdots + (n-1)x^{n-1}$$

When a summation has the form

$$\sum_{k=m}^{n} f(k)$$

where $f(k)$ is a monotonically increasing function, we can approximate it by integrals thanks to the following theorem.

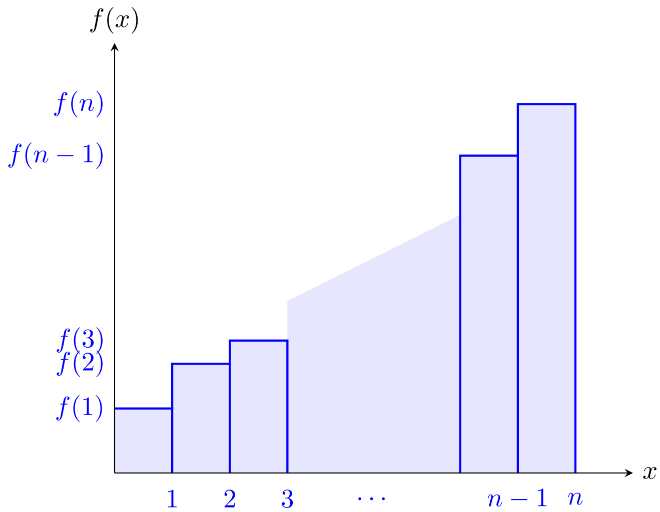

Figure 1. The area of the i–th rectangle is $f(i)$. The shaded region is S.

Theorem 2.64. *[*] Let $f : \mathbb{R} \to \mathbb{R}$ be an increasing function. Define*

$$S = \sum_{i=1}^{n} f(i)$$

and

$$I = \int_{1}^{n} f(x)\,\mathrm{d}x$$

Then

$$I + f(1) \leq S \leq I + f(n).$$

Similarly, if f is decreasing, then

$$I + f(n) \leq S \leq I + f(1).$$

Proof. Suppose f is increasing. The value of the sum S is the sum of the areas of n unit–width rectangles of heights $f(1), f(2), \ldots, f(n)$, the shaded area in figure 1.

The value of

$$I = \int_{1}^{n} f(x)\,\mathrm{d}x$$

is the shaded area under the curve of $f(x)$ from 1 to n shown in Figure 2. Comparing the shaded regions in Figures 1 and 2 shows that S is at least I plus the area of the leftmost rectangle. Hence,

$$S \geq I + f(1)$$

which is the lower bound for S given in the statement. To derive the upper bound for S, we shift the curve of $f(x)$ from 1 to n one unit to the left as shown in Figure

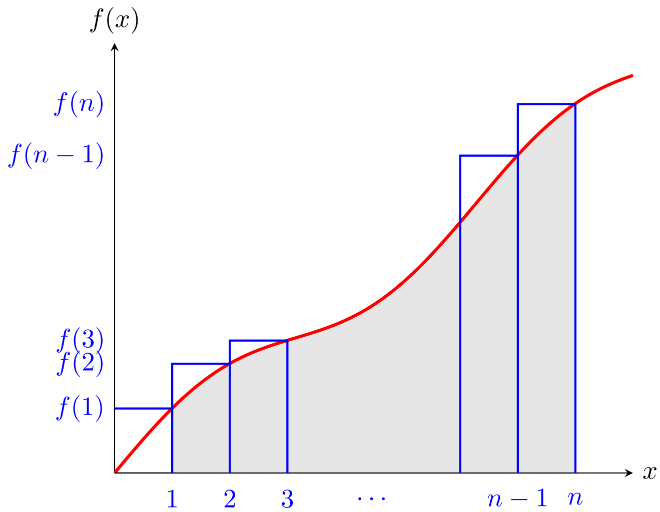

Figure 2. The shaded area under the curve $f(x)$ from 1 to n is $I = \int_1^n f(x)\,\mathrm{d}x$.

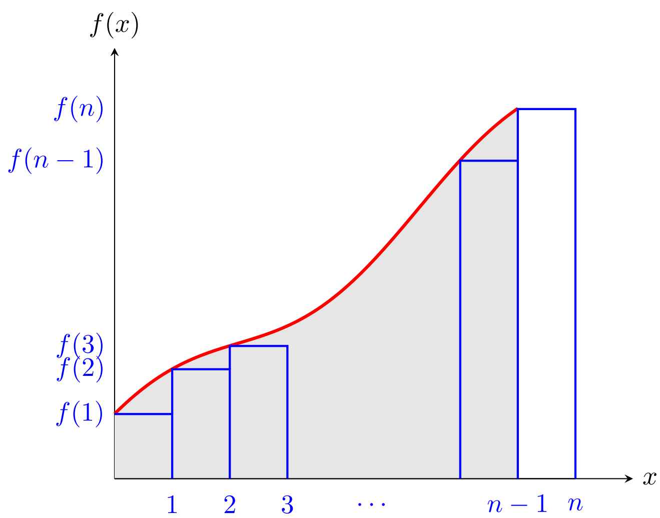

Figure 3. This curve is the same as that in figure 2 shifted left by 1.

3. Comparing the shaded regions in Figures 1 and 3 shows that S is at most I plus the area of the rightmost rectangle. That is,

$$S \le I + f(n)$$

which is the upper bound for S given in the statement. A similar argument proves the decreasing case. $\square$

Definition 2.65. The harmonic numbers are defined by:

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \qquad \Diamond$$

Applying Theorem 2.64 the the n–th harmonic number we can get a very tight estimate of it. Since $1/x$ is decreasing, we obtain

$$\int\limits_1^{n+1} \frac{\mathrm{d}x}{x} + \frac{1}{n} \le \sum_{k=1}^{n} \frac{1}{k} \le \int_1^n \frac{1}{x}\,\mathrm{d}x + \frac{1}{1}$$

that is

$$\ln n + \frac{1}{n} \le H_n \le \ln n + 1$$

which implies

$$H_n = \ln n + O(1)$$

as $n \to +\infty$.

In a similar way, the product of $a_1, \ldots, a_n$ can be denoted by

$$\prod_{k=1}^{n} a_k$$

For products we sometimes can convert them into a sum by taking a logarithm. For example, if

$$P = \prod_{k=1}^{n} f(k)$$

then

$$\ln(P) = \sum_{k=1}^{n} \ln[f(k)]$$

We can then apply our summing tools to find a closed form (or approximate closed form) for $\ln P$ and then exponentiate at the end to undo the logarithm. For example, for the factorial function $n!$ we start by taking the logarithm:

$$\begin{aligned} \ln(n!) &= \ln[1 \cdot 2 \cdot 3 \cdot \ldots (n-1) \cdot n] \\ &= \ln 1 + \ln 2 + \ln 3 + \cdots + \ln(n-1) + \ln(n) \\ &= \sum_{i=1}^{n} \ln(i). \end{aligned}$$

No closed form for this sum is known. However, we can apply Theorem 2.64 to find good closed–form bounds on the sum. To do this, we first compute

$$\int_1^n \ln(x)\,\mathrm{d}x = n \ln n - n + 1$$

and then we plug this into the Theorem's statement to get

$$n \ln n - n + 1 \le \sum_{i=1}^{n} \ln(i) \le n \ln n - n + 1 + \ln n.$$

Exponentiating the above inequalities gets
$$\frac{n^n}{e^{n-1}} \le n! \le \frac{n^{n+1}}{e^{n-1}}$$
which means that $n!$ is within a factor of n of n^n/e^{n-1}. Among the many formulas that provide bounds for the factorial, the most famous is the so called Stirling's formula
$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\epsilon(n)}$$
where
$$\frac{1}{12n+1} \le \epsilon(n) \le \frac{1}{12n}.$$
We note that $\epsilon(n) > 0$ which implies
$$n! > \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$
for all $n \in \mathbb{N}$. Moreover, since $\epsilon(n) \to 0$, we have
$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$
and the estimate provided by this formula is quite good.

Exercises on section 2.5

Exercise 2.36. Prove that
$$\left(\sum_{i=1}^n a_i + \sum_{i=1}^n b_i\right)\left(\sum_{i=1}^n a_i - \sum_{i=1}^n b_i\right) = \sum_{i,j=1}^n (a_i a_j - b_i b_j)$$

Exercise 2.37. Prove that
$$\sum_{k=1}^n H_k = (n+1)H_n - n.$$

2.6. Series

Definition 2.66. Given any a_n, the sequence s_n defined by
$$s_n = \sum_{i=0}^n a_i = a_0 + a_1 + a_2 + \cdots + a_n$$
is the sequence of partial sums of a_n. If s_n converges to S we say the series $\sum\limits_{n=0}^{+\infty} a_n$ converges to S, and we write $\sum\limits_{n=0}^{+\infty} a_n = S$. If s_n diverges to $+\infty$ we say the series $\sum\limits_{n=0}^{+\infty} a_n$ diverges to $+\infty$, and we write $\sum\limits_{n=0}^{+\infty} a_n = +\infty$ $\Diamond$

Remark 2.67. Series' properties:

- one can factor out any common number (that is, a number that does not depend on the index)
- where the sum starts matters for S

- convergence does not depend on the starting index

Theorem 2.68. *[*] If $\sum a_n$ converges then $a_n \to 0$.*

Proof. Consider the sequence of partial sums, s_n. If the series converges, then $s_n \to S$. Then

$$|a_n| = |s_n - s_{n-1}| \leq |s_n - S| + |s_{n-1} - S|$$

Now fix $\varepsilon > 0$. We can find N such that for $n - 1 > N$, $|s_n - S| < \varepsilon$ and $|s_{n-1} - S| < \varepsilon$ which means

$$|a_n| \leq 2\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, $|a_n| \to 0$ and therefore $a_n \to 0$. □

Theorem 2.69. *If $|x| < 1$ then the geometric series converges and its sum is*

$$\sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$$

If $q \geq 1$ the series divergs to $+\infty$. If $q \leq -1$ the series has no limit.

Proof. The partial sums of a geometric series are

$$s_n = \sum_{i=0}^{n-1} x^i = \frac{1-x^n}{1-x}$$

and we get the conclusion by taking the limit of s_n as $n \to +\infty$. □

Remark 2.70. By convention $0^0 = 1$

Remark 2.71. If $|x| < 1$ the geometric series defines a function

$$f(x) = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

The function is differentiable for every x such that $|x| < 1$ and it can be proved that the derivative of f is given also by the series obtained differentiating each term in the series, so that

$$f'(x) = \sum_{k=0}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}$$

We have

$$\sum_{k=0}^{\infty} kx^{k-1} = \sum_{k=1}^{\infty} kx^{k-1}$$

since the first term in the series on the left is 0, and

$$\sum_{k=1}^{\infty} kx^{k-1} = \sum_{j=0}^{\infty} (j+1)x^j$$

where $j = k - 1$. Renaming j to k gets

$$\sum_{k=0}^{\infty} kx^{k-1} = \sum_{k=0}^{\infty} (k+1)x^k$$

and therefore
$$\sum_{k=0}^{\infty}(k+1)x^k = \frac{1}{(1-x)^2}.$$
Since
$$\sum_{k=0}^{\infty}(k+1)x^k = \sum_{k=0}^{\infty}kx^k + \sum_{k=0}^{\infty}x^k$$
substituting the expressions for f and f' in the last equality gives
$$\frac{1}{(1-x)^2} = \sum_{k=0}^{\infty}kx^k + \frac{1}{1-x}$$
whence
$$\sum_{k=0}^{\infty}kx^k = \frac{1}{(1-x)^2} - \frac{1}{1-x} = \frac{x}{(1-x)^2}$$
The last equation gives a practical formula for computing sums that appear frequently in the study of algorithms in computer science.

Theorem 2.68 provides a necessary condition for convergence which is not sufficient.

Example 2.72. The series
$$\sum_{n=1}^{\infty}\frac{1}{n}$$
is called harmonic series and it diverges to $+\infty$. In fact we have
$$\sum_{n=1}^{\infty}\frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{8} + \frac{1}{9} + \cdots + \frac{1}{16} + \ldots$$
Now note that
$$\begin{aligned}\frac{1}{3} + \frac{1}{4} &> \frac{1}{4} + \frac{1}{4} = \frac{1}{2},\\ \frac{1}{5} + \cdots + \frac{1}{8} &> \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{1}{2},\\ \frac{1}{9} + \cdots + \frac{1}{16} &> \frac{1}{16} + \frac{1}{16} + \cdots + \frac{1}{16} = \frac{1}{2},\end{aligned}$$
and so on. Therefore the partial sums increase by $1/2$ with 2, 4, 8 terms ... and s_n is thus bound to diverge to $+\infty$ $\triangle$

Theorem 2.73. *[*] Suppose $a_1 \geq a_2 \geq a_3 \geq \cdots \geq 0$. Then the series*
$$\sum_{n=1}^{+\infty}a_n$$
converges if and only if the series
$$\sum_{k=0}^{+\infty}2^k a_{2^k}$$
converges.

Proof. By Theorem 2.18, it suffices to consider boundedness of the partial sums. Let

$$s_n = a_1 + a_2 + \cdots + a_n, t_k = a_1 + 2a_2 + \cdots + 2^k a_{2^k}.$$

For $n < 2^k$,

$$\begin{aligned} s_n &\le a_1 + (a_2 + a_3) + \cdots + (a_{2^k} + \cdots + a_{2^{k-1}-1}) \\ &\le a_1 + 2a_2 + \cdots + 2^k a_{2^k} \\ &= t_k \end{aligned}$$

so that

$$s_n \le t_k$$

On the other hand, if $n > 2^k$,

$$\begin{aligned} s_n &\ge a_1 + a_2 + (a_3 + a_4) + \cdots + (a_{2^{k-1}+1} + \cdots + a_{2^k}) \\ &\ge \frac{1}{2}a_1 + a_2 + 2a_4 + \cdots + 2^{k-1}a_{2^k} \\ &= \frac{1}{2}t_k \end{aligned}$$

so that

$$2s_n \ge t_k$$

The two inequalities show that the sequences s_n and t_k are either both bounded or both unbounded. This completes the proof. □

Definition 2.74. If $p \in \mathbb{R}$, the series

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

is the generalized harmonic series. ◊

Theorem 2.75 (*)**.** *The generalized harmonic series converges if and only if $p > 1$.*

Proof. If $p < 0$, divergence follows from Theorem 2.68. If $p > 0$ we can apply Theorem 2.73, and we are led to the series

$$\sum_{k=0}^{+\infty} 2^k \frac{1}{2^{kp}} = \sum_{k=0}^{+\infty} 2^{(1-p)k}$$

Now, $2^{1-p} < 1$ if and only if $1 - p < 0$, and the result follows by comparison with the geometric series. □

Theorem 2.76 (Comparison)**.** *If $a_n \ge 0$, $b_n \ge 0$ and $a_n \le b_n$ for each n (or at least for a sufficiently large n), then:*

- *if $\sum a_n$ diverges then $\sum b_n$ diverges*
- *if $\sum b_n$ converges then $\sum a_n$ converges*

Theorem 2.77 (Asymptotic comparison)**.** *If $a_n \ge 0$, $b_n \ge 0$ and $a_n \sim b_n$ then $\sum a_n$ and $\sum b_n$ both converge or diverge to $+\infty$.*

The previous two theorems enable us to determine whether a series is convergent or not, provided we have another series whose behavior is known.

Theorem 2.78 (Ratio test)**.** *The series* $\sum a_n$

- *converges if*

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| < 1$$

- *does not converge if there exists* N *such that*

$$\left|\frac{a_{n+1}}{a_n}\right| > 1$$

  for $n \geq N$*.*

Proof.
- From

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \beta < 1$$

  we know we can find $b < 1$ such that

$$\left|\frac{a_{n+1}}{a_n}\right| < b$$

  for all $n > N$. In particular,

$$\begin{gathered}|a_{N+1}| < b|a_N|,\\ |a_{N+2}| < b|a_{N+1}| < b^2|a_N|,\\ \dots\\ |a_{N+p}| < b^p|a_N|,\end{gathered}$$

  that is

$$|a_n| < b^n b^{-N}|a_N|$$

  for $n > N$. The result follows by comparison with the geometric series $\sum b^n$.
- If the condition hold, one can show similarly to the previous case, that $a_n \not\to 0$, so that the series does not converge.

□

Note that

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = 1$$

implies nothing: consider for example $\sum 1/n$ and $\sum 1/n^2$.

Example 2.79. To find the sum of the series

$$\sum_{n=2}^{+\infty}\frac{1}{n(n-1)}$$

we write the first terms of the partial sums sequence:

$$\begin{aligned}\sum_{n=2}^{+\infty}\frac{1}{n(n-1)} &= \frac{1}{2\cdot 1}+\frac{1}{3\cdot 2}+\frac{1}{4\cdot 3}+\dots\\ &= \left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\dots\\ &= 1-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\frac{1}{3}-\frac{1}{4}+\dots\end{aligned}$$

and it's clear that

$$s_n = 1-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\frac{1}{3}+\dots+\frac{1}{n}-\frac{1}{n-1}=1-\frac{1}{n}$$

Then the series converges to 1. Series for which each term simplifies with the previous one are called telescoping series. △

Exercises on sections 2.6

Exercise 2.38 (MC)**.** The series $\sum_{n=0}^{+\infty}\frac{n+5}{2n+7}$

a) converges to 0
b) converges to $\frac{1}{2}$
c) diverges to $+\infty$
d) does not have a limit

Exercise 2.39 (MC)**.** The series $\sum_{n=0}^{+\infty} q^n$ converges to 6. Then:

a) $q \to 0$
b) $q = \frac{1}{6}$
c) $q = \frac{6}{7}$
d) $q = \frac{5}{6}$

Exercise 2.40 (MC)**.** The sum of the series $\sum_{n=1}^{+\infty}(2/5)^n$ is

a) 5/3
b) 2/5
c) 2/3
d) $+\infty$

Exercise 2.41 (MC)**.** $\sum_{n=1}^{+\infty}\frac{3\ln n+n^2}{2e^{-n}+n^3}$

a) converges to 0
b) converges to $S > 0$
c) diverges to $+\infty$

d) none of the others

Exercise 2.42 (TF)**.** Let $\sum_{n=0}^{+\infty} a_n = 3$ and let s_n be the sequence of its partial sums. Then

a) $s_n \to 3$

b) $s_n \to 0$

c) $a_n \to 0$

d) $s_n = o\,(3)$

Exercise 2.43. Prove that

$$\sum_{k=1}^{n} \frac{1}{2k-1} = \ln\sqrt{n} + O(1)$$

by manipulating the harmonic series

Exercise 2.44. Prove that if $|x| < 1$ then

$$\sum_{k=0}^{\infty} k^2 x^k = \frac{x(1+x)}{(1-x)^3}$$

Exercise 2.45. Show that $0.777777\cdots = 7/9$.

Answers

1.1 The statement is trivially true for $n = 0, 1, 2$. Assume it's true for n. Then we want to prove

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

We have

$$\begin{aligned} 2^0 + 2^1 + \cdots + 2^{n+1} &= 2^0 + 2^1 + \cdots + 2^n + 2^{n+1} \\ &= 2^{n+1} - 1 + 2^{n+1} \\ &= 2 \cdot 2^{n+1} - 1 \\ &= 2^{n+2} - 1 \end{aligned}$$

and the statement is true.

1.2 We use induction on the length of words, that is, the number of symbols in them. A word of length 1 is either x, y or z so the number of ('s and)'s in a word of length 1 is zero. So the result is true for words of length 1.

Assume the result is true for words of lengths less than n. A word of length n is either of the form $(A)(B)$ or of the form $(A) + (B)$, for some words A, B of length less than n. By induction assumption, A and B have the same number of ('s and)'s. Therefore, $(A)(B)$ and $(A) + (B)$ also have the same number of ('s and)'s. So the result is true for words of any length.

1.3 Let $P(n)$ be the statement that exactly $n - 1$ moves are required to assemble a puzzle with n pieces. Now $P(1)$ is trivially true. Assume that $P(j)$ is true for all $j \leq k$, and consider a puzzle with $k + 1$ pieces. The final move must be the joining of two blocks, of size j and $k + 1 - j$ for some integer j with $1 \leq j \leq k$. By the inductive hypothesis, it required $j - 1$ moves to construct the one block, and $k+1-j-1 = k-j$ moves to construct the other. Therefore, $1+(j-1)+(k-j) = k$ moves are required in all, so $P(k + 1)$ is true.

1.4 The statement is true for $n = 1$, $1 = \frac{6}{6} = 1$. Now, we assume

$$1 + 4 + 9 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

and want to show that

$$1+4+9+\cdots+n^2+(n+1)^2 = \frac{(n+1)(n+2)(2(n+1)+1)}{6} = \frac{(n+1)(n+2)(2n+3)}{6}$$

We have

$$\begin{aligned} 1 + 4 + 9 + \cdots + n^2 + (n+1)^2 &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \\ &= \frac{n(n+1)(2n+1) + 6(n+1)^2}{6} \\ &= \frac{(n+1)[n(2n+1) + 6(n+1)]}{6} \\ &= \frac{(n+1)[2n^2 + 7n + 6]}{6} \\ &= \frac{(n+1)(n+2)(2n+3)}{6} \end{aligned}$$

1.5 The statement is true for $n = 1$, $1 + 3 = (1+1)^2 = 4$. Now, assume

$$1 + 3 + 5 + \cdots + 2n + 1 = (n+1)^2$$

and we want to show that

$$1 + 3 + 5 + \cdots + 2n + 1 + 2(n+1) + 1 = (n+2)^2$$

We have

$$\begin{aligned} 1 + 3 + 5 + \cdots + 2n + 1 + 2(n+1) + 1 &= (n+1)^2 + 2(n+1) + 1 \\ &= n^2 + 4n + 1 + 2 + 1 \\ &= n^2 + 4n + 4 = (n+2)^2 \end{aligned}$$

1.6 The statement is true for $n = 1$, since $2 > 1$. Now, assume $2^n > n$. We want to show that $2^{n+1} > n + 1$. We have

$$2^{n+1} = 2 \cdot 2^n > 2n = n + n > n + 1$$

where the last inequality follows from the fact that for the inductive step, $n > 1$.

1.7 The statement is true for $n = 2$ since $x^2 > 0$ and

$$(1+x)^2 = 1 + 2x + x^2 > 1 + 2x.$$

Now, assume $(1+x)^n > 1 + nx$. We want to show that $(1+x)^{n+1} > 1 + (n+1)x$. We have

$$\begin{aligned} (1+x)^{n+1} &= (1+x)^n \cdot (1+x) \\ &> (1+nx)(1+x) \\ &= 1 + nx + x + nx^2 \\ &= 1 + (n+1)x + nx^2 > 1 + (n+1)x \end{aligned}$$

1.8 The statement is false for $n = 1, 2, 3, 4$ and true for $n = 5$, so $n = 5$ could be the base case. Now we assume $2^n > n^2$ and want to prove that $2^{n+1} > (n + 1)^2$. We have

$$2^{n+1} = 2 \cdot 2^n = 2^n + 2^n > 2^n + n^2$$

but we need $n^2 + 2n + 1$ on the right hand side. Then we need to prove that $2^n > 2n + 1$. We prove the last statement by induction. $2^n > 2n + 1$ is true for $n = 3$. Assuming is true for n we have

$$2^{n+1} = 2 \cdot 2^n = 2^n + 2^n > 2n + 1 + 2n + 1 > 2n + 2 + 1 = 2(n + 1) + 1$$

where the last inequality follows from the fact that if $n \geq 3$ then $2n > 1$. This proves $2^n > 2n + 1$ for every $n \geq 3$. Returning to the original question, we now have, for $n \geq 5$

$$2^{n+1} = 2 \cdot 2^n = 2^n + 2^n > n^2 + 2n + 1 = (n + 1)^2$$

and the have proved the original statement. Since $n = 5$ is the base case, we have $N = 4$, that is, the statement is proved for every $n > 4$.

1.9 The statement is true for $n = 1$ since $1 = 2^0$ and $2 = 2^1$ so that any positive number less than or equal to 2^1 can be expressed as the sum of distinct powers of two. Now, assume this is true for any number less than or equal to 2^k. We want to show that any number n such that $2^k < n \leq 2^{k+1}$ can be expressed as the sum of distinct powers of 2.

If $n = 2^{k+1}$ the conclusion is true. If $n < 2^{k+1}$ then $n = 2^k + p$ where p is an integer such that $p < 2^k$. By the inductive hypothesis, p can be expressed as the sum of distinct powers of 2 and none of them can be 2^k since $p < 2^k$. Then n is equal to 2^k plus all the powers of two that sum up to p.

1.10 The base case $n = 1$ is trivial (as any horse is the same color as itself), and the inductive step is correct in all cases $n > 1$. However, the inductive step is incorrect for $n = 1$, because the statement that "the two sets overlap" is false (there are only $n + 1 = 2$ horses prior to either removal, and after removal the sets of one horse each do not overlap). The argument is taken from *A random walk in science* by R. L. Weber.

1.11 The statement is true for $n = 1$. Suppose it's true for k, that is $8^k - 3^k = 5p$ for a $p \in \mathbb{Z}$. Then

$$\begin{aligned} 8^{k+1} - 3^{k+1} &= 8 \cdot 8^k - 3 \cdot 3^k = (5 + 3) \cdot 8^k - 3 \cdot 3^k \\ &= 5 \cdot 8^k + 3 \cdot 8^k - 3 \cdot 3^k = 5 \cdot 8^k + 3(8^k - 3^k) \\ &= 5 \cdot 8^k + 3 \cdot 5p \end{aligned}$$

which is divisible by 5.

1.12 The base case is obvious: to generate 4 dollars, just use 2 twos. Inductive step: assume the ATM can generate any amount up to n. To generate $n+1$ consider the two cases:

- to generate n at least one five dollars bill is used. Then substitute the five with 3 twos and you generate $n + 1$

- to generate n only twos are needed. Then substitute 2 twos with one five and you get $n+1$.

1.13 The base case is necessarily 24. In fact, $24 = 7 \cdot 2 + 5 \cdot 2$. For the inductive step, assume we have $n = 7x + 5y$. Then, to get

- $n+1$, we decrease x by 2 and increase y by 3
- $n+2$, we increase x by 1 and decrease y by 1
- $n+3$, we decrease x by 1 and increase y by 3
- $n+4$, we increase x by 2 and increase y by 2
- $n+5$, we leave x unchanged and increase y by 1

Now, all the changes performed were legal, because both x and y never became negative (at the beginning, we had $x = 2$ and $y = 2$). Moreover, at the end, x is unchanged and y is increased by 1 so the same operations can be performed again, to move forward by another 5 numbers, and so on. Then the process can go on indefinitely and the theorem is proved.

Technically, this is a special form of induction, where the statement is proved separately for some naturals, and then to all $\mathbb{N}$.

1.14 We have

$$a_1 = \frac{1}{2!}, a_2 = \frac{5}{3!}, a_3 = \frac{23}{4!}, a_4 = \frac{119}{5!}$$

which suggests that

$$a_n = \frac{(n+1)! - 1}{(n+1)!}$$

To prove the conjecture, we already have three base cases. We then assume

$$a_n = \frac{(n+1)! - 1}{(n+1)!}$$

and try to prove that

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

From the definition of a_n we have

$$a_{n+1} = \frac{1}{2!} + \frac{2}{3!} + \cdots + \frac{n}{(n+1)!} + \frac{n+1}{(n+2)!}$$

that is

$$a_{n+1} = a_n + \frac{n+1}{(n+2)!}$$

Substituting the formula for a_n and a little algebra gets

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

1.15 The derivative of x is 1 so we have the base case. Assume the derivative of x^n is nx^{n-1}. To find the derivative of x^{n+1} we write $x^{n-1} = x \cdot x^n$. Then, by the product rule we have

$$\begin{aligned}\left(x^{n+1}\right)' &= (x \cdot x^n)' \\ &= 1 \cdot x^n + x \cdot nx^{n-1} \\ &= x^n + nx^n \\ &= (n+1)x^n\end{aligned}$$

1.16 To prove the conjecture via mathematical induction, we first look for the base case. For $n = 1$ we have

$$s_1 = \frac{1}{1+1} = \frac{1}{2}$$

so the base case is true. Now, assume

$$s_n = \frac{n}{n+1}.$$

We want to prove that

$$s_{n+1} = \frac{n+1}{n+2}$$

$$\begin{aligned}s_{n+1} &= s_n + \frac{1}{(n+1)\cdot(n+2)} \\ &= \frac{n}{n+1} + \frac{1}{(n+1)\cdot(n+2)} \\ &= \frac{n(n+2)+1}{(n+1)\cdot(n+2)} \\ &= \frac{n^2+2n+1}{(n+1)\cdot(n+2)} \\ &= \frac{(n+1)^2}{(n+1)\cdot(n+2)} \\ &= \frac{n+1}{n+2}\end{aligned}$$

1.17 Note that for every $x \in \mathbb{R}$ we have $\lfloor x \rfloor = k + r$ for a suitable $k \in \mathbb{Z}$ and a $r \in \mathbb{R}$ such that $0 \leq r < 1$. Moreover, if n is odd, then $n = 2k + 1$ for a suitable $k \in \mathbb{Z}$.

(1) With the above notation,

$$\lfloor x - 2 \rfloor = \lfloor k + r - 2 \rfloor = k - 2 = \lfloor x \rfloor - 2$$

(2) Again with the above notation we have

$$\begin{aligned}\left\lfloor \frac{n^2}{4} + 1 \right\rfloor &= \left\lfloor \frac{4k^2+4k+1}{4} + 1 \right\rfloor = \left\lfloor k^2 + k + \frac{5}{4} \right\rfloor \\ &= k^2 + k + 1 \\ &= \frac{4k^2+4k+4}{4} = \frac{(2k+1)^2+3}{4} \\ &= \frac{n^2+3}{4}\end{aligned}$$

1.18 First note that when $m > n$ there are no one–to–one functions from a set with m elements to a set with n elements. For $m \leq n$ suppose the elements in the domain are $a_1, a_2, \ldots, a_m$. There are n ways to choose the value of the function at a_1. Because the function is injective, the value of the function at a_2 can be picked in $n-1$ ways (because the value used for a_1 cannot be used again). In general, the value of the function at a_k can be chosen in $n-k+1$ ways. By the product rule, there are

$$n(n-1)(n-2)\cdots(n-m+1)$$

injective functions from a set with m elements to one with n elements.

1.19 b).

1.20 d).

1.21 (1) 128

(2) 450

(3) 9

(4) 675

(5) 450

(6) 450

(7) 225

(8) 75

1.22 c).

1.23 If n is odd, then the palindrome string is $(n+1)/2$ bits, followed by the first $(n-1)/2$ bits in reverse order. Then there are $2^{(n+1)/2}$ palindrome strings. If n is even, then the first $n/2$ bit must be repeated in reverse order in the second half of the string, therefore there are $2^{n/2}$ palindrome strings.

1.24 There are 2^7 bit strings that begin with three 0s and 2^8 that end with two 0s. Those that begin with three 0s and end with two 0s are 2^5. Then the answer is

$$2^7 + 2^8 - 2^5 = 352$$

1.25 Group the integers from 1 to 8 into 4 subsets:

$$\{1,8\}, \{2,7\}, \{3,6\}, \{4,5\}$$

Since there are 4 subsets and 5 numbers, by the pigeonhole principle at least two numbers get picked from the same subset. Then those two numbers sum to 9.

1.26 For every number, compute its remainder when divided by n. There are at most n possible remainders, 0, $\ldots$, $n-1$. By the Pigeonhole principle, at least two numbers, say p and q have the same remainder when divided by n. Therefore, $p = an + b$ and $q = cn + b$, where b is the (same) remainder. Then $p - q = an + b - cn - b = (a-c)n$ which shows that the difference of p and q is indeed divisible by n.

1.27 Group the integers from 1 to $2n$ into n subsets:

$$\{1,2\}, \{3,4\}, \{5,6\}, \ldots, \{2n-1, 2n\}$$

Now, associate each of the $2n$ numbers to the subset it belongs to. Since there are $2n$ numbers and n subsets, by the pigeonhole principle at least two numbers get associated with the same subset, which means they differ by 1 and they are therefore relatively prime.

1.28 Consider an equilateral triangle whose side is 1 cm. The vertices are the "pigeons" and they are colored either red or blue, the "holes". At least two vertices must have the same color.

1.29 Suppose the row starts with a woman. Then, the n women can be arranged in $n!$ ways in positions 1, 3, 5, etc.. For each arrangement of women, there are $n!$ arrangements of the n men in positions 2, 4, 6, etc. By the product rule, we have $(n!)^2$ ways of arranging the people. If a man starts the row, then there are $(n!)^2$ ways so that overall the number of ways is $2(n!)^2$.

1.30 (1) $5! = 120$, since the string BCD is to be considered as a single letter.

(2) $4! = 24$

(3) $5! = 120$, because BA and GF are to be treated as a single letter

(4) $4! = 24$

(5) $4! = 6$, because the two conditions are equivalent to ask the string to contain the string ABCDE

(6) 0, because the two conditions in contradiction.

1.31 (1) $52^5 = 380204032$

(2) $P(52, 5) = \frac{52!}{47!} = 311875200$

1.32

$$\binom{8}{2} \cdot \binom{6}{2} \cdot \binom{4}{2}$$

1.33

$$\binom{7}{2} + 7 = 28$$

because dominoes with the same number are not counted in the $\binom{7}{2}$ term.

1.34 There are $n!/2$ such permutations. In fact, in the set of all permutations of n elements, for every permutation π in which 1 comes before 2 there is another permutation π' in which all the elements are in the same place except for 1 and 2 that are swapped and therefore 1 comes after 2. And since every permutation must include both 1 and 2, the set of all couples of such permutations is the same of all $n!$ permutations.

1.35 If 1 and 2 are next to each other, we can consider the two symbols together as a unique symbol, say 0. So, for example, if $n = 5$ permutation 12345 is 0345 and 32145 is 3045. How many permutations are there with $n-1$ symbols? $(n-1)!$. But the "0" in each of them might represent both "12" and "21". So the permutations with 1 and 2 next to each other are $2(n-1)!$. The question asks for the number of permutations where 1 and 2 are not next to each other so the answer is $n! - 2(n-1!)! = (n-1)!(n-2)$

1.36 (1) The first letter can be chosen in n ways, the second in $n-1$ ways, the third in $n-1$ ways, etc. Therefore, the answer is $n(n-1)^{k-1}$.

(2) Let $p = \lfloor \frac{k+1}{2} \rfloor$. The first p letters create an arbitrary word of length p. There are n^p ways to construct such words. The rest of the word is determined by its first part.

1.37 Overall, there are 2^n subsets of S. Choosing one of them, say A, is implicitly choosing the other, which will be $B = S \setminus A$. Since order does not matter, in the list of all possible couples every couple (A, B) is repeated once in (B, A). Therefore, the number of ways to split the set is $2^n/2 = 2^{n-1}$.

1.38 If there is a block of girls (group of girls sitting next to each other) of length at least 3, then the statement is obviously true. So, assume each block of girls consists of at most 2 girls. Hence, there are at least 13 girl blocks, which in turn means there are at least 13 boy blocks. Since the number of boys is 25, there must be a boy block consisting of a single boy, which proves that the answer is yes.

1.39 First solution: The first row can be chosen in 2^m ways, the second row in $2^m - 1$ ways, the third row in $2^m - 2$ ways, etc. Therefore, there are $2^m(2^m - 1)(2^m - 2) \ldots (2^m - n + 1)$ binary $n \times m$ matrices with pairwise different rows.

Second solution: Each row of a matrix must contain m entries from the set $\{0, 1\}$, and so there are 2^m possible distinct rows. We pick n of these, so there are $\binom{2^m}{n}$ ways to select n different rows. These n rows may occur in any order; there are $n!$ ways to arrange n rows. Thus there are $\binom{2^m}{n} n!$ matrices with pairwise different rows.

1.40 (1) $2^k - 1$

(2) B can be represented as $B = A \cup C$ for some nonempty subset $C \subseteq U \setminus A$. There are $2^{n-k} - 1$ nonempty subset $C \subseteq U \setminus A$. Therefore, there are $2^{n-k} - 1$ sets B satisfying $B \supset A$

(3) $A \cap B = \emptyset$ means $B \subset U \setminus A$. There are 2^{n-k} such subsets.

(4) If $A \cap B \neq \emptyset$ then $B = C \cup D$, where C is a nonempty subset of A and D is a subset of $U \setminus A$. C can be chosen in $2^k - 1$ ways and D can be chosen in 2^{n-k} ways. Therefore, B can be chosen in $2^{n-k}(2^k - 1) = 2^n - 2^{n-k}$ ways.

1.41 Say the jobs are 1,2 and 3 and the employees are A, B, C, D and E. Then, an assignment is an ordered arrangement of three letters chosen in a set of five. For example ADD means job 1 is assigned to employee A and jobs 2 and 3 are assigned to employee D. Then, the total number of assignments is the number of 3–permutations of the elements of a set of five with repetitions allowed, that is $5^3 = 125$.

1.42 We are looking for the non negative solutions to the equation

$$x + y + z = 3000$$

According to the stars and bars method, we have $n = 3$ and $k = 3000$ and the number of solutions is

$$C(3 - 1 + 3000, 3000) = C(3002, 3000) = 4504501$$

1.43 The number is
$$\frac{10!}{2!3!5!} = 2520$$

1.44 We use the "stars and bars" method.

(1) If $x_1 \geq 1$ then there are 20 stars and 4 bars so that the number of solutions is $C(20+4,4) = 10626$.

(2) If $x_i \geq 2$, $i = 1,2,3,4,5$, then there are 11 stars (10 are already set) and 4 bars, so tht the number of solutions is $C(11+4,4) = 1365$

1.45 Setting $x = 1$ and $y = -1$ in the Binomial theorem gets
$$(1-1)^n = 0 = \sum_{k=0}^{n} \binom{n}{k}(-1)^k$$
Another way to see this is to say that, for every n,
$$\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \cdots = \binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \ldots$$

1.46 Setting $x = 1$ and $y = 2$ in the Binomial theorem gets
$$(1+2)^n = 3^n = \sum_{k=0}^{n} \binom{n}{k} 2^k$$

1.47 The values for A_n are 1, 2, 6, 20 which are the central values of the even rows of the Pascal's triangle, so the obvious guess is
$$A_n = \binom{2n}{n}$$
To prove that
$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$$
using algebra is extremely time consuming. A combinatorial proof is much easier. The left hand side is
$$\binom{n}{0}\binom{n}{0} + \binom{n}{1}\binom{n}{1} + \cdots + \binom{n}{n}\binom{n}{n}$$
and using the fundamental binomial identity
$$\binom{n}{k}\binom{n}{n-k}$$
we can rewrite it as
$$\binom{n}{0}\binom{n}{n} + \binom{n}{1}\binom{n}{n-1} + \cdots + \binom{n}{n}\binom{n}{0}$$
Now consider the right hand side: $\binom{2n}{n}$ represents the number of subsets of size n that can be formed from a set of $2n$ elements. Name this elements as $1, 2, \ldots, 2n$. It's obvious that any subset of n elements is made from a subset of $\{1, 2, \ldots, n\}$ of size k and a subset of $\{n+1, n+2, \ldots, 2n\}$ of size $n-k$. Which is exactly what the left hand side is counting.

1.48 Let

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

Then

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

and setting $x = 1$ gets

$$n2^{n-1} = \sum_{k=0}^{n} \binom{n}{k} k.$$

1.49 b).

1.50 c).

1.51 c).

1.52 (1) The decimanl numbers are

$$11001 = -6, \quad 01101 = 13, \quad 10001 = -14, \quad 11111 = 0.$$

Note that in one's complement representation, there are two "zeros": the one with all digits equal to 1 and that with all digits equal to 0.

(2) Simply taking the complement of every digit

(3) Just adding the digit and if a carry is present in the leftmost digit (that of the sign) this is used as a carry in the rightmost digit. For example, adding 13 to -6, that is 01101 to 11001 we get 100110 and the first "1" is moved to the last position getting 00111, that is 7.

1.53 Since $8^4 = 4096$ and $8^5 = 32768$, 5 digits will be enough to represent any integer from 0 to 32767.

2.1 c).

2.2 b).

2.3 b).

2.4 *a*) false
b) false
c) false
d) true

2.5 *a*) true
b) true
c) true
d) true

2.6 *a*) true
b) false
c) false
d) true

2.7 The subsequence with even indexes converges to 0 and is monotonic strictly decreasing. Its terms are 1, 3^{-2}, 3^{-4}, ... and therefore the maximum is 1 and there is no minimum. The greatest lower bound (infimum) is 0.

The subsequence with odd indexes can be written as

$$\frac{3n-4}{2n} = \frac{3}{2} - \frac{2}{n},$$

its terms are $-1/2$, $5/6$, $11/10$, ... it's monotonic strictly increasing and it converges to $3/2$. For this subsequence the minimum is $-1/2$ and there is no maximum. The least upper bound (supremum) is $3/2$.

For the sequence as a whole, the supremum is $3/2$ and it's not a maximum while the minimum is $-1/2$ and it's (obviously) the infimum.

2.8 *a*) false
b) true
c) false
d) false

2.9 *a*) false
b) true
c) true
d) false

2.10 If $a_n > 1$ we have

$$a_{n+1} = \frac{a_n + 1}{2} > \frac{1+1}{2} = 1.$$

Comparing a_{n+1} with a_n we have

$$a_{n+1} \leq a_n \iff \frac{a_n + 1}{2} \leq a_n \iff a_n + 1 \leq 2a_n \iff 1 \leq a_n$$

so that $a_n > 1$ implies $a_{n+1} < a_n$ and $a_{n+1} > 1$. Since $a_0 > 1$, we have $1 < a_1 < a_0$ and the sequence is monotonically strictly decreasing. To show that it converges it's enough to prove that it's bounded below. But this is obvious because starting with $a_0 > 1$ all the terms are greater than 1 which is a lower bound for the sequence. Thus the sequence converges.

To find its limit L, note that if $n \to +\infty$ we have $a_n \to L$ and $a_{n+1} \to L$ and L must solve the equation

$$L = \frac{L+1}{2}$$

that is $L = 1$.

2.11 Fix $\varepsilon > 0$. Since $a_n \to L$ then there exists N such that

$$L - \varepsilon < a_n < L + \varepsilon$$

for $n > N$. Therefore, it is also true that

$$L - \varepsilon < a_{n+1} < L + \varepsilon$$

Since $L \neq 0$, both the denominator and the numerator do not vanish when ε is sufficiently close to 0. Then

$$\frac{L-\varepsilon}{L+\varepsilon} \leq \frac{a_{n+1}}{a_n} \leq \frac{L+\varepsilon}{L-\varepsilon}.$$

But

$$1 - \frac{2\varepsilon}{L+\varepsilon} = \frac{L-\varepsilon}{L+\varepsilon} \leq \frac{a_{n+1}}{a_n} \leq \frac{L+\varepsilon}{L-\varepsilon} = 1 + \frac{2\varepsilon}{L-\varepsilon}$$

and the last chain of inequalities shows that indeed the fraction tends to 1 as $n \to +\infty$.

About the case $L = +\infty$, consider $a_n = 2^n$. We have

$$\frac{a_{n+1}}{a_n} = \frac{2^{n+1}}{2^n} = 2$$

for every n which does not tend to 1.

2.12 Since $a_n \to +\infty$, for every M there are finitely many n such that $a_n < M$. Let A_M the set of the a_n such that $a_n < M$:

$$A_M = \{a_n : a_n < M\}$$

So the cardinality of A_M is finite for every M. Choose M such that A_M is not empty (for example, choose $M = a_0 + 1$). Then A_M is a finite set and necessarily has a minimum. This is the minimum of the sequence.

2.13 *a*) false
b) false
c) true
d) true

2.14 *a*) true
b) true
c) false
d) true

2.15 All sequences but $a_n = n^2 + n + 1$ are $O(n)$.

2.16 From the definition, it means there exists C such that $|a_n| < C \cdot 1 = C$. This means a_n is bounded.

2.17 From the definition, it means there exists C_1 such that $|a_n| < C_1 \cdot 1 = C_1$ and that there exists C_2 such that $|a_n| \geq C_2 \cdot 1 = C_2$. This means a_n is bounded but also that it can't be closer to 0 than C_2.

2.18 For $n > 1$ we have $n^2 + 1 < n^3$ (easily proved by induction). Then

$$\ln(n^2+1) \leq \ln(n^3) = 3\ln n = \frac{3\ln 2}{\ln 2}\ln n = 3\ln 2 \log_2 n$$

which shows $\ln(n^2+1)$ is $O(\log_2 n)$. On the other side we have (for $n > 1$)

$$\log_2 n = \frac{\ln n}{\ln 2} < \frac{\ln(n^2+1)}{\ln 2}$$

which shows $\log_2 n$ is $O(\ln(n^2+1))$. Thus the two sequences are of the same order.

2.19 We need to find the least p and the corresponding C and k such that

$$\left|\frac{n^3+1}{n^2+1}\right| \leq Cn^p$$

for $n > k$. The left hand side is positive for every integer n so we can remove the absolute value sign. From algebra we have

$$\frac{n^3+1}{n^2+1} = n + \frac{1-n}{n^2+1}$$

and therefore for $n > 1$

$$\left|\frac{n^3+1}{n^2+1}\right| \leq |n| + \left|\frac{1-n}{n^2+1}\right| = n + \frac{n-1}{n^2+1}$$

Set

$$a_n = \frac{n-1}{n^2+1}.$$

We have $a_n \to 0$ as $n \to +\infty$. Moreover, a_n is monotonically strictly decreasing. In fact

$$\begin{aligned} a_n - a_{n+1} &= \frac{n-1}{n^2+1} - \frac{(n+1)-1}{(n+1)^2+1} \\ &= \frac{n-1}{n^2+1} - \frac{n}{n^2+2n+2} \\ &= \frac{(n-1)(n^2+2n+2) - n(n^2+1)}{(n^2+1)(n^2+2n+2)} \\ &= \frac{n^3+2n^2+2n-n^2-2n-2-n^3-n}{(n^2+1)(n^2+2n+2)} \\ &= \frac{n^2-n-2}{(n^2+1)(n^2+2n+2)} \end{aligned}$$

which is positive for $n > 2$. Then, the max of a_n is one of a_1, a_2, a_3. Since we have $a_1 = 0$, $a_2 = 1/5$ and $a_3 = 1/5$, then $a_n < 1/5$ for all n. Going back to the original sequence,

$$\left|\frac{n^3+1}{n^2+1}\right| \leq n + \frac{n-1}{n^2+1} \leq n+1 < 2n$$

for $n > 1$. Thus $p = 1$, $C = 2$ and $k = 1$ and

$$\left|\frac{n^3+1}{n^2+1}\right| \leq 2n \qquad \text{for } n > 1.$$

2.20 Let s_n be the number of sequences $\{a_1, \ldots, a_k\}$ strictly increasing and such that $a_1 = 1$ and $a_k = n$. We have $s_1 = 1$ because there is only one such sequence, $\{1\}$. We have also $s_2 = 1$, because there is only one $s_2 = \{1, 2\}$. For s_3 we have two sequences: $\{1, 3\}$ and $\{1, 2, 3\}$. For s_4 we have four: $\{1, 4\}$, $\{1, 2, 4\}$, $\{1, 3, 4\}$, $\{1, 2, 3, 4\}$. For s_5 we have eight: $\{1, 5\}$, $\{1, 2, 5\}$, $\{1, 3, 5\}$, $\{1, 4, 5\}$, $\{1, 2, 3, 5\}$, $\{1, 2, 4, 5\}$, $\{1, 3, 4, 5\}$, $\{1, 2, 3, 4, 5\}$. And similarly for $n > 5$.

As is clear from the last two cases, every sequence in s_n is obtained appending n to every sequence in $s_1, \ldots, s_{n-1}$. So

$$s_n = s_1 + s_2 + \cdots + s_{n-1}.$$

With the initial conditions $s_1 = 1$, $s_2 = 1$ we get $s_3 = 2$, $s_4 = 4$, $s_5 = 8$, etc. If we assume $s_n = 2^{n-2}$ for $n \geq 2$ then

$$s_{n+1} = s_1+s_2+\cdots+s_n = s_1+\big(s_2 + \cdots + s_n\big) = 1+\big(1 + 2 + \cdots + 2^{n-2}\big) = 1+\big(2^{n-1} - 1\big) = 2^{n-1}.$$

2.21 Let a_n be number of bit strings of length n that contain a pair of consecutive 0s. We have $a_1 = 0$ and $a_2 = 1$ (there is only one string, “00”). We have $a_3 = 3$, because there are three such strings: “000”, “100” and “001”.

For a_4, we take all such strings of length 3 and add a 1 at the end., getting “0001”, “1001” and “0011”. Then, take all such strings of length 2 and add 10 at the end to get “0010”. Finally, we take *all* strings with length 2 and add 00 at the end. Since there are four strings of length 2, “00”, “10”, “01” and “11”, we get “0000”, “1000”, “0100” and “1100”. These three cases cover all possibile bit strings with two consecutive 0s of length 4.

For the general case, note that every bit string ends either with 1 or with 0 and this is true also for bit strings with two consecutive 0s. In the first case, the two consecutive 0s must be in the previous $n - 1$ positions. In the second case, there are two sub-cases: the string ends with 00 or with 10. If the string ends with 00 then this is a string with the property, no matter what the previous $n - 2$ bits are. If, finally, the string ends with 10, then the two consecutive zeros must appear in the first $n - 2$ positions. This argument proves that

$$a_n = a_{n-1} + a_{n-2} + 2^{n-2}$$

with the initial conditions $a_1 = 0$ and $a_2 = 1$. From the recursion we get

$$a_3 = 3, a_4 = 8, a_5 = 19, a_6 = 43, a_7 = 94$$

2.22 Let a_n be number of bit strings of length n that contain three consecutive 0s. We have $a_0 = 0$, $a_1 = 0$, $a_2 = 0$ and $a_3 = 1$ (there is only one string, “000”).

Following an argument similar to that of exercise 2.21, note that every bit string ends either with 1 or with 0 and this is true also for bit strings with three consecutive 0s. In the first case, the three consecutive 0s must be in the previous $n - 1$ positions. In the second case, there are three sub-cases: the string ends with 10, with 100 or with 000. If the string ends with 10, then the three consecutive zeros must appear in the first $n - 2$ positions. If the string ends with 100, then the three consecutive zeros must appear in the first $n - 3$ positions. If the string ends with 000 then this is a string with the property, no matter what the previous $n - 3$ bits are. This argument proves that

$$a_n = a_{n-1} + a_{n-2} + a_{n-3} + 2^{n-3}$$

with the initial conditions $a_0 = 0$, $a_1 = 0$ and $a_2 = 0$. From the recursion we get

$$a_3 = 1, a_4 = 3, a_5 = 8, a_6 = 20, a_7 = 47$$

2.23 Let b_n be number of bit strings of length n that do not contain three consecutive 0s. We have $b_0 = 1$, $b_1 = 2$, $b_2 = 4$ and $b_3 = 7$ (all strings but “000”).

A bit string with no three consecutive zeros of length n either ends in 1 or 0. The number of strings with no three zeros that end in 1 is equal to that of the number of bit strings of length $n - 1$ with the same property, b_{n-1}. If a bit string with no three consecutive zeros ends in 0, then it either ends in 10 or in 100. The number of strings that end in 10 is b_{n-2} and that of those that end in 100 is b_{n-3}. No two such sets overlap with one another so we have

$$b_n = b_{n-1} + b_{n-2} + b_{n-3}$$

Another way to solve the problem is to find a recursion for the number of bit strings of length n and then subtract this from the total number of strings. From exercise 2.22, if a_n is the number of strings with 000 we have

$$a_n = a_{n-1} + a_{n-2} + a_{n-3} + 2^{n-3}$$

with the initial conditions $a_0 = 0$, $a_1 = 0$ and $a_2 = 0$. Let b_n be the number of bit strings of length n with no three consecutive zeros. Then, for every n, we have

$$a_n = 2^n - b_n$$

and therefore

$$2^n - b_n = 2^{n-1} - b_{n-1} + 2^{n-2} - b_{n-2} + 2^{n-3} - b_{n-3} + 2^{n-3}.$$

Since $2^{n-3} + 2^{n-3} = 2^{n-2}$, $2^{n-2} + 2^{n-2} = 2^{n-1}$ and $2^{n-1} + 2^{n-1} = 2^n$, we get

$$-b_n = -b_{n-1} - b_{n-2} - b_{n-3}$$

that is

$$b_n = b_{n-1} + b_{n-2} + b_{n-3}$$

and the initial conditions are $b_0 = 1$ (the empty string), $b_1 = 2$ and $b_2 = 4$. We then have

$$b_3 = 7, b_4 = 13, a_5 = 24, a_6 = 44, a_7 = 81$$

2.24 Let b_n be number of bit strings of length n that do not contain two consecutive 0s. We have $b_0 = 1$ and $b_1 = 3$ ("0", "1" and "2"). There are 9 ternary strings of length 2 and only one of them, "00" contains two zeros so $b_2 = 8$.

For the general case, note that a ternary string ends in 2, 1 or 0. The number of strings with no two consecutive zeros that end in 1 is equal to that of the number of ternary strings of length $n-1$ with the same property, b_{n-1}. The same is true for strings that end in 2. If a ternary string ends with 0 and has no two consecutive zeros, then it either ends in 10 or in 20. The number of strings that end in 10 is b_{n-2} and that of those that end in 20 is b_{n-2}. No two such sets overlap with one another so we have

$$b_n = 2b_{n-1} + 2b_{n-2}$$

with the initial conditions $b_0 = 1$, $b_1 = 3$. We then have

$$b_2 = 8, b_3 = 22, b_4 = 60, a_5 = 164, a_6 = 448.$$

2.25 To answer this question let a_n be the appropriate number of regions. We have already seen that $a_1 = 2$ and $a_2 = 4$. Sketching a simple drawing shows that $a_3 = 7$.

In general suppose that we add a line ℓ_{n+1} to already existing lines $\ell_1, \ldots, \ell_n$. The existing lines split ℓ_{n+1} into $n+1$ segments, each of which splits an existing region into two parts. Hence we have

$$a_{n+1} = a_n + (n+1)$$

The initial condition is $a_1 = 2$. The recurrence is easily solved:

$$\begin{aligned} a_{n+1} &= a_n + (n+1) \\ &= a_{n-1} + n + (n+1) \\ &= a_{n-2} + (n-1) + n + (n+1) \\ &= \dots \\ &= a_1 + 2 + \cdots + (n+1). \end{aligned}$$

Since $a_1 = 2 = 1 + 1$ we have

$$a_{n+1} = 1 + 1 + 2 + \cdots + (n+1) = 1 + \frac{(n+1)(n+2)}{2}$$

or, setting $n = n + 1$,

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

For example,

$$a_2 = 4, a_3 = 7, a_4 = 11, a_5 = 16, \dots$$

2.26 The number of bit strings of length n containing three consecutive 0s was found in 2.22 and is given by

$$a_n = a_{n-1} + a_{n-2} + a_{n-3} + 2^{n-3}$$

with the initial conditions $a_0 = 0$, $a_1 = 0$ and $a_2 = 0$. From the recursion we get

$$a_3 = 1, a_4 = 3, a_5 = 8, a_6 = 20, a_7 = 47, a_8 = 107.$$

In a similar way, one can find that the number of bit strings of length n containing four consecutive 1s is given by

$$b_n = b_{n-1} + b_{n-2} + b_{n-3} + b_{n-4} + 2^{n-4}$$

with the initial conditions $b_0 = 0$, $b_1 = 0$, $b_2 = 0$, $b_3 = 0$ and $b_4 = 1$. From the recursion we get

$$b_5 = 3, b_6 = 8, b_7 = 20, b_8 = 48.$$

Now, again, we have double counted the strings containing both three consecutive 0's and four consecutive 1's. Those are 11110000, 11110001, 11111000, 01111000, 00011110, 00011111, 00001111, 10001111 i.e. 8 strings. Thus, the total number of bit strings of length eight containing either three consecutive 0's or four consecutive 1's is $107 + 48 - 8 = 147$.

2.27 The characteristic equation is $x^2 = x + 2$ and its solution are -1 and 2. So the general solution is

$$a_n = c_1(-1)^n + c_2 2^n$$

From the initial conditions we get $c_1 = -1$ and $c_2 = 3$ and finally

$$a_n = -(-1)^n + 3 \cdot 2^n$$

2.28 The recurrence is homogeneous and linear. The characteristic equation is $r^2 - r - 6 = 0$ with roots -2 and 3. The general solution is

$$\alpha(-2)^n + \beta 3^n$$

The initial conditions imply $\alpha = 3/5$ adn $\beta = 12/5$ and the solution to the original recurrence is

$$\frac{3}{5}(-2)^n + \frac{12}{5}3^n$$

2.29 The recurrence if homogeneous and linear. The characteristic equation is $r^2 - 7r - 10 = 0$ with roots 2 and 5. The general solution is

$$\alpha 5^n + \beta 2^n$$

The initial conditions imply $\alpha = -1$ adn $\beta = 3$ and the solution to the original recurrence is

$$-5^n + 3 \cdot 2^n$$

2.30 The recurrence if homogeneous and linear. The characteristic equation is $r^2 - 2r + 1 = 0$ with unique root 1. The general solution is

$$\alpha 1^n + \beta n 1^n = \alpha + \beta n$$

The initial conditions imply $\alpha = 4$ and $\beta = -3$ and the solution to the original recurrence is

$$4 - 3n$$

2.31 The recurrence if homogeneous and linear. The characteristic equation is $r^3 - 7r - 6 = 0$ with roots -1, -2 and 3. The general solution is

$$\alpha(-1)^n + \beta 3^n + \gamma(-2)^n$$

The initial conditions imply $\alpha = 8$, $\beta = 4$ and $\gamma = -3$and the solution to the original recurrence is

$$8(-1)^n + 4 \cdot 3^n - 3 \cdot (-2)^n$$

2.32 To verify that -2^{n+1} is a solution we substitute into the equation:

$$-2^{n+1} = 3(-2^n) + 2^n = -3 \cdot 2^n + 2^n = -2 \cdot 2^n = -2^{n+1}$$

and the equality is verified. To find all solution we need to find the solutions to the corresponding homogeneous equation, that is

$$a_n = 3a_{n-1}$$

whose solutions are $a_n = \alpha 3^n$. Therefore, the solutions to the original recurrence are

$$a_n = \alpha 3^n - 2^{n+1}$$

and since $a_0 = \alpha - 2$ should be 1, we have $\alpha = 3$. Then, the solution with $a_0 = 1$ is

$$a_n = 3^{n+1} - 2^{n+1}$$

2.33 The recursion is a linear non homogeneous recursion. We first get a particula solution. Since the non homogeneous term is 2^{n-2} we look for a solution of type $a_n = c2^n$. Substituting into the equation gives:

$$c2^n = c2^{n-1} + c2^{n-2} + 2^{n-2}$$

which divided by 2^{n-2} gets

$$4c = 2c + c + 1$$

whose only solution is $c = 1$. So $a_n = 2^n$ is a particular solution of the given recurrence.

The homogeneous equation is

$$a_n = a_{n-1} + a_{n-2}$$

whose solution is

$$c_1\left(\frac{1+\sqrt{5}}{2}\right)^n + c_2\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Setting

$$\phi = \frac{1+\sqrt{5}}{2}$$

we get

$$-\frac{1}{\phi} = \frac{1-\sqrt{5}}{2}.$$

The general solution is then

$$a_n = c_1\phi^n + c_2\left(-\frac{1}{\phi}\right)^n + 2^n.$$

From the initial conditions we get

$$a_0 = c_1 + c_2 + 1 = 0, a_1 = c_1\phi + c_2\left(-\frac{1}{\phi}\right) + 2 = 0.$$

whence

$$c_1 = -\frac{2\phi+1}{\phi+2}, c_2 = \frac{\phi-1}{\phi+2}$$

and the general solution is

$$a_n = -\frac{2\phi+1}{\phi+2}\phi^n + \frac{\phi-1}{\phi+2}\left(-\frac{1}{\phi}\right)^n + 2^n.$$

2.34 We need to solve

$$a_{n+1} = a_n + n + 1, a_1 = 2.$$

This is a linear non homogeneous recurrence. Its characteristic equation is $r = 1$ and therefore the solution is $a_n = 1^n = 1$. What remains to be found is a particular solution. Since the non homogeneous term is a first degree polynomial we look for a similar polynomial, say $an + b$. Substituting into the recurrence gets

$$a(n+1) + b = an + b + n + 1$$

that is

$$a = n + 1$$

and a is not a constant term. If we try with $an^2 + bn + c$ we get

$$a(n+1)^2 + b(n+1) + c = an^n + bn + c + n + 1$$

and, after a bit of algebra,

$$2an + a + b = n + 1$$

whence $2a = 1$ and $a + b = 1$, whose solution is $a = b = 1/2$. Thus the general solution to the recurrence is

$$a_n = \frac{1}{2}n^2 + \frac{1}{2}n + k$$

and since $a_1 = 2$ we have $k = 1$. Thus the solution to the original problem is

$$a_n = \frac{1}{2}n^2 + \frac{1}{2}n + 1.$$

2.35 Applying Theorem 2.59 we get

$$s_{2^m} = 2^m \left(s_1 + \frac{1}{2} \sum_{i=0}^{m-1} \frac{2^{2i}}{2^i} \right)$$

which simplifies to

$$s_{2^m} = 2^m \left(s_1 + \frac{2^m - 1}{2} \right).$$

2.36 From the algebraic identity $(a+b)(a-b) = a^2 - b^2$ we get

$$\left(\sum_{i=1}^n a_i + \sum_{i=1}^n b_i \right) \left(\sum_{i=1}^n a_i - \sum_{i=1}^n b_i \right) = \left(\sum_{i=1}^n a_i \right)^2 - \left(\sum_{i=1}^n b_i \right)^2 = \sum_{i,j=1}^n (a_i a_j - b_i b_j)$$

2.37 The statement can be proved by induction. The statement is true for $n = 1$, because $1 = 2 - 1$, and $n = 2$, because

$$H_1 + H_2 = 1 + \left(1 + \frac{1}{2} \right) = 3H_2 - 2 = 3 \cdot \frac{3}{2} - 2.$$

Suppose the statement is true for n. For $n + 1$ the statement is

$$\sum_{k=1}^{n+1} H_k = (n+2)H_{n+1} - (n+1).$$

We have

$$\begin{aligned}
\sum_{k=1}^{n+1} H_k &= \sum_{k=1}^{n} H_k + H_{n+1} \\
&= (n+1)H_n - n + H_{n+1} \\
&= (n+1)\left(H_{n+1} - \frac{1}{n+1} \right) - n + H_{n+1} \\
&= (n+2)H_{n+1} - \frac{n+1}{n+1} - n \\
&= (n+2)H_{n+1} - (n+1)
\end{aligned}$$

The statement can also be proved using double sums.

$$\begin{aligned}
\sum_{k=1}^{n} H_k &= \sum_{k=1}^{n}\sum_{j=1}^{k}\frac{1}{j} \\
&= \sum_{j=1}^{n}\sum_{k=j}^{n}\frac{1}{j} \\
&= \sum_{j=1}^{n}\frac{1}{j}\sum_{k=j}^{n}1 \\
&= \sum_{j=1}^{n}\frac{1}{j}(n-j+1) \\
&= \sum_{j=1}^{n}\left(\frac{n+1}{j}-1\right) \\
&= (n+1)\sum_{j=1}^{n}\frac{1}{j}-n \\
&= (n+1)H_n-n
\end{aligned}$$

2.38 c).

2.39 d).

2.40 c).

2.41 c).

2.42 *a*) true

b) false

c) true

d) false

2.43 We have

$$\sum_{k=1}^{n}\frac{1}{2k-1} = 1+\frac{1}{3}+\frac{1}{5}+\cdots+\frac{1}{2n-1}$$

and therefore

$$\begin{aligned}
\sum_{k=1}^{n} \frac{1}{2k-1} &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{2n-1} - \frac{1}{2} - \frac{1}{4} - \cdots - \frac{1}{2n-2} \\
&= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{2n-1} - \frac{1}{2}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n-1}\right) \\
&= \sum_{k=1}^{2n-1} \frac{1}{k} - \frac{1}{2} \sum_{k=1}^{n-1} \frac{1}{k} \\
&= \ln(2n-1) + O(1) - \frac{1}{2}\ln(n-1) + O(1) \\
&= \ln 2 + \ln(n - \frac{1}{2}) + O(1) - \frac{1}{2}\ln(n-1) + O(1) \\
&= \frac{1}{2}\ln n + O(1)
\end{aligned}$$

2.44 Since $|x| < 1$ we have

$$f(x) = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

Differentiating once as in remark 2.71 we get

$$\sum_{k=0}^{\infty} kx^k = \frac{x}{(1-x)^2}$$

Differentiating once more we get

$$\frac{2}{(1-x)^3} = \sum_{k=0}^{\infty} k^2 x^k + 3\sum_{k=0}^{\infty} kx^k + 2\sum_{k=0}^{\infty} x^k$$

whence

$$\sum_{k=0}^{\infty} k^2 x^k = \frac{2}{(1-x)^3} - 3\sum_{k=0}^{\infty} kx^k - 2\sum_{k=0}^{\infty} x^k = \frac{2}{(1-x)^3} - 3\frac{x}{(1-x)^2} - 2\frac{1}{1-x}$$

and finally

$$\sum_{k=0}^{\infty} k^2 x^k = \frac{x(1+x)}{(1-x)^3}$$

2.45 We have

$$\begin{aligned}
0.77777\ldots &= \frac{7}{10} + \frac{7}{100} + \frac{7}{1000} + \ldots \\
&= \frac{7}{10} + \frac{7}{10^2} + \frac{7}{10^3} + \ldots \\
&= \sum_{n=1}^{+\infty} \frac{7}{10^n} \\
&= 7\sum_{n=1}^{+\infty} \frac{1}{10^n}
\end{aligned}$$

But

$$\sum_{n=1}^{+\infty} \frac{1}{10^n} = \frac{1}{10} \sum_{n=0}^{+\infty} \frac{1}{10^n} = \frac{1}{10} \frac{1}{1 - \frac{1}{10}} = \frac{1}{9}$$

so that

$$0.77777 \cdots = 7 \sum_{n=1}^{+\infty} \frac{1}{10^n} = 7 \cdot \frac{1}{9} = \frac{7}{9}.$$