3.2. Surjections, Injections, Bijections, Sequences

Note. In this section, we define the terms in the title of the section and give examples. We discuss the "marriage problem" in the context of the existence of an injection. We recursively demonstrate a bijection between $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$ (see Example 3.18).

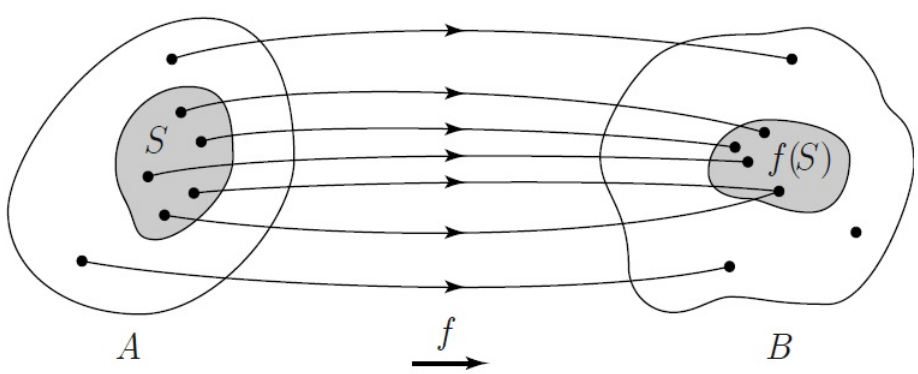
Definition. Consider function f where $f : A \to B$. The subset of B consisting of all those elements of B that are images of elements of A is the *range* or *image* of f, denoted $f(A)$ or $\text{im}(f)$. Symbolically,

$$f(A) = \text{im}(f) = \{f(a) \mid a \in A\}.$$

Note 3.2.A. Since we have define a function as a set of ordered pairs, we have

$$\text{im}(f) = \{b \in B \mid (a, b) \in f \text{ for some } a \in A\} = \{b \in B \mid f(a) = b \text{ for some } a \in A\}.$$

If $S \subseteq A$ where A is the domain of f, then we write $f(S) = \{f(x) \mid x \in S\}$. Graphically, we have (see page 118):



Definition. Consider function $f : A \to B$. Function f is *surjective* or *onto* if $f(A) = B$, in which case f is a *surjection*. In terms of elements, this is equivalent to the condition that for all $b \in B$ there exists an element $a \in A$ such that $f(a) = b$.

Example 3.9. **(a)** The function $f : \mathbb{R} \to \mathbb{R}$ defined as $f(x) = x^2$ is not surjective (onto) because for every $y < 0$ there is no $x \in \mathbb{R}$ such that $f(x) = x^2 = y < 0$ (though we need to only find one such y to show that f is not surjective).

(b) The function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x + 1$ is surjective, because for any $y \in \mathbb{R}$ we have for $x = y - 1 \in \mathbb{R}$ that $g(x) = g(y - 1) = (y - 1) + 1 = y$, as needed to show surjective.

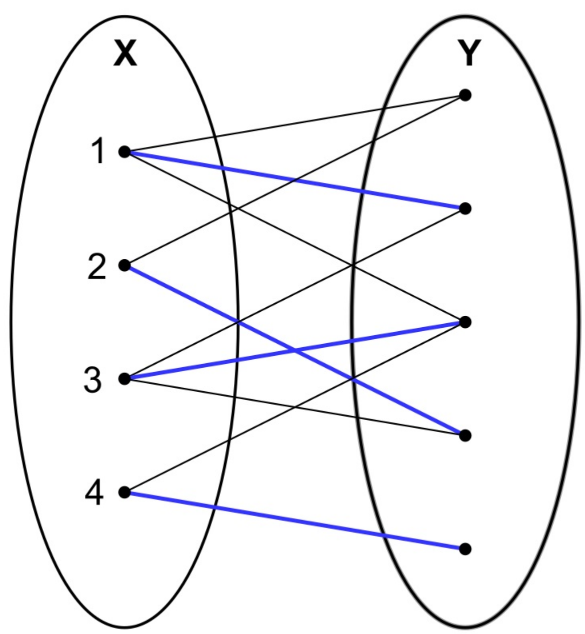
(d) We discussed indexed sets in Section 2.6. Indexed Sets. If S and I are sets and $f : I \to S$ is a surjective function, then we can index set S using set I and function f. The index $i \in I$ associated with $s \in S$ is the i such that $f(i) = s$, in which case we denote s as s_i.

Definition 3.10. A function $f : A \to B$ is said to be *one-to-one* or *injective* if the following implication is true for every $a_1, a_2 \in A$: $f(a_1) = f(a_2) \Rightarrow a_1 = a_2$.

Note 3.2.B. In showing that a function is injective, it is sometimes convenient to consider the contrapositive of Definition 3.10: for all $a_1, a_2 \in A$, $a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)$. That is, different elements in the domain have different images.

Example 3.11(a). The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$ is injective.

Note. We now describe a classical problem which we use to motivate some of the topics we will soon encounter. We follow the traditional (and non-progressive) version of the "marriage problem." Let set $X = \{x_1, x_2, \ldots, x_n\}$ and Y be sets of unmarried women and men, respectively. Suppose that each woman x_i determines the set $M_i \subseteq Y$ of all men that she considers acceptable for marriage. The marriage problem is: Is every woman able to marry an acceptable man (with the condition that only one woman can marry any particular man)? We model this with an injective function and an indexing set. Let $I = \{1, 2, \ldots, n\}$ and suppose we are given a family $\{M_i\}_{i\in I}$ of nonempty sets. We look for an injective function $f : I \to \cup_{i\in I} M_i$ such that $f(i) \in M_i$ for $i \in I$. The one-to-oneness of the function function models the fact that two different women cannot marry the same man. The condition $f(i) \in M_i$ guarantees that each woman marries a man acceptable to her. The set $\{f(1), f(2), \ldots, f(n)\}$ is a *system of distinct representatives* for the family of sets $\{M_i\}_{i\in I}$. If we consider the first k women, where $1 \le k \le n$, then there must be at least k men available in their sets of "acceptable men," $\cup_{i=1}^k M_i$. Quantitatively, this is expressed as: $\#\left(\cup_{i=1}^k M_i\right) \le k$ for all $1 \le k \le n$ (where, for finite set S, $\#S$ denotes the number of elements in set S; this is explored more in Section 4.1. Cardinality; Fundamental Counting Principles). Though not at all obvious, it turns out that this condition is sufficient for the existence of the desired function f. This was proved by Philip Hall in "On Representatives of Subsets," *Journal of the London Mathematical Society*, **10**(1), 26-30 (1935). The first page of this paper can be viewed on the Journal of the London Mathematical Society webpage (accessed 1/31/2022). For this reason, this result is called "Hall's Marriage Theorem." Gerstein declares: "The proof is beyond the scope of our present treatment. . . ." We can also describe the result in terms of graph theory. We can consider the bipartite graph with partite sets X and Y, and (x, y) as an edge of the graph if and only if woman x finds man y as acceptable for marriage. The existence of the desired function f is then equivalent to finding an X-saturating matching of the bipartite graph. We will introduce some graph theory terminology in Section 5.9. Graphs. Many more details are given in ETSU's senior/graduate level Introduction to Graph Theory (MATH 4347/5347 I have some online notes for this class). In ETSU's graduate level Graph Theory 2 (MATH 5450), a proof of Halls Marriage Theorem is given (see Section 16.2. Matchings in Bipartite Graphs and Theorem 16.4 "Hall's Theorem"). An illustration of the graph and the matching (in blue) is given below in an image from the Wikipedia Hall's marriage theorem page (accessed 1/31/2022).



Example 3.12. Find a system of distinct representatives for the following family of sets: $\{1\}$, $\{1, 2, 3, 8\}$, $\{3, 4\}$, $\{2, 4, 5\}$, $\{3, 6\}$, $\{1, 4, 7\}$, $\{6\}$.

Solution. Since there are seven sets, then we take the indexing set as $I = \{1, 2, \ldots, 7\}$ and we look for an injection from I to $\cup_{i \in I} M_i = \{1, 2, \ldots, 8\}$ such that $f(i) \in M_i$ for all $i \in I$. We apply a "greedy algorithm" which is (informally) a technique where make a choice at each step which seems to be best at the time, but may later require some backtracking and modification of the choice. Here, we go through the sets from first to last and choose a as a representative the least number compatible with the choices already made (if possible). This gives the first six representatives as 1, 2, 3, 4, 6, 7, but then we cannot choose a representative from the last set. So we backtrack to fix this to get the modified choice:

$$\begin{array}{cccccc} 1, & 2, & 3, & 4, & \not 6, & 7 \\ & & & & 3 & \end{array}$$

But this is also a problem, so we backtrack again to get:

$$\begin{array}{cccccc} 1, & 2, & \not 3, & 4, & \not 6, & 7 \\ & & 4 & & 3 & \end{array}$$

but this repeats representative 4, so again we modify to get:

$$\begin{array}{cccccc} 1, & 2, & \not 3, & \not 4, & \not 6, & 7 \\ & & 4 & 5 & 3 & \end{array}$$

So we have the choice of distinct representatives: 1, 2, 4, 5, 3, 7, 6 (respectively). By Hall's Marriage Theorem, a solution exists. We see that this greedy algorithm finds a solution. Notice that it is not a unique solution, since we could also take the distinct representatives: 1, 8, 4, 2, 3, 7, 6 (respectively).

Definition 3.13. A function $f : A \to B$ is *bijective* or a *one-to-one correspondence* if it is both injective (i.e., one to one) and surjective (i.e., onto).

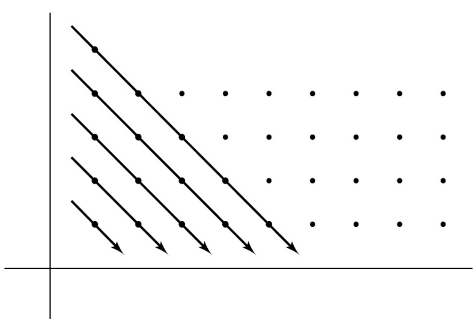
Note 3.2.C. Since a function is ultimately a set of ordered pairs (by Definition 3.2), we can classify a bijection f in terms of the set of ordered pairs: Function $f : A \to B$ is a bijection if and only if each $a \in A$ is the first coordinate of exactly one pair belonging to f, and each $b \in B$ is the second coordinate of exactly one pair belonging to f. Notice that every injection is a bijection between its domain and its *range*.

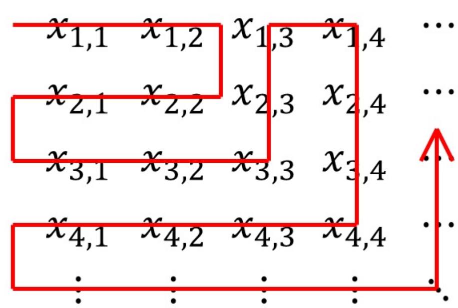
Definition 3.16. Let A be a set and let $\mathbb{N}_n = \{x \in \mathbb{N} \mid 1 \leq x \leq n\} = \{1, 2, \ldots, n\}$. A *finite sequence* in A is a function $f : \mathbb{N}_n \to A$ for some $n \in \mathbb{N}$. The *length* of this sequence is n. An *infinite sequence* in A is a function $f : \mathbb{N} \to A$.

Note. In this class, we will use the term "sequence" to indicate either a finite or an infinite sequence. This may be different from your use of the term in calculus, where the term "sequence" mean "infinite sequence" (and you were concerned with convergence or divergence of the sequence). However, notice that the definition of infinite sequence given above is exactly the same as that given in Calculus 2 (with set $A = \mathbb{R}$; see my online Calculus 2 notes on Section 10.1. Sequences).

Note/Definition. In keeping with the notation for sequences given in Calculus 2, if we have infinite sequence f in set A where $f(1) = a_i$, $f(2) = a_2$, etc., then we denote the sequence as $a_1, a_2, a_3, \ldots$. If f is a finite sequence of length n with similar values, then we denote the sequence as $(a_1, a_2, \ldots, a_n)$, called an *n-tuple* with a_i as its ith *coordinate*. In any such sequence, we call a_i the ith *term*. We use various notation to indicate sequences: $\{a_i\}_{1 \le i \le n}$, $\{a_n\}_{n \in \mathbb{N}}$, $\{a_n\}_{n \ge 1}$.

Example 3.18. Consider the set $\mathbb{N} \times \mathbb{N}$. Graphically, we can think of this as all points in the first quadrant of the Cartesian plane with integer coefficients. See the figure below (left). We wish to find a bijective function $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$; that is, we wish to make a list of these point by picking a first one, a second one, etc. Notice that we cannot start at point $(1, 1)$ and then run along the bottom row of points to the end, then start on the second row from the bottom at point $(1, 2)$,...since there *is no end* to the bottom row (or any of the other rows).





In this figure (left), we see a strategy to follow in constructing the desired bijection. We go along diagonals that run upper left-lower right one at a time. These leads to the sequence $(1, 1)$; $(1, 2)$, $(2, 1)$; $(1, 3)$, $(2, 2)$, $(3, 1)$; $(1, 4)$, $(2, 3)$, $(3, 2)$, $(4, 1)$; $(1, 5)$, etc. We now give an explicit formula for function f. Notice that points (a, b) and (c, d) are on the same diagonal in the figure if and only if $a + b = c + d$ (and this sum increases by 1 each time we go "up" to the next diagonal), and on such a diagonal point (a, b) precedes point (c, d) on their common diagonal if and only if $a < c$. The point (a, b) is on an "earlier" diagonal than point (c, d) if and only if $a + b < c + d$. So we have the rule: point (a, b) precedes (c, d) on the list if and only if $a + b < c + d$ (i.e., (a, b) is on a lower diagonal than (c, d)) or $a + b = c + d$ and $a < c$ (i.e., (a, b) and (c, d) are on the same diagonal with (a, b) earlier on the diagonal than (c, d)). Now we give a recursive definition of the function based on this strategy. Define $f(1) = (1, 1)$. For $k \geq 1$ and with $f(k) = (a, b)$, define

$$f(k+1) = \begin{cases} (a+1, b-1) & \text{if } b > 1 \\ (1, a+1) & \text{if } b = 1. \end{cases}$$

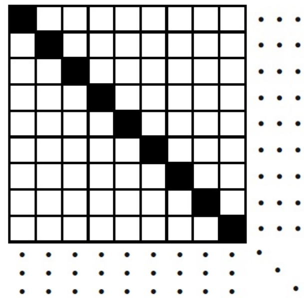
Notice that when $b = 1$ we have reached the bottom of a diagonal, so we start over in the first column of the next diagonal. When $b > 1$ and we are *not* at the bottom of a row, we simply go over one (to $a+1$) and down one (to $b-1$) so that we go from (a, b) to $(a+1, b-1)$. It is intuitively clear that f is a bijection. None-the-less, we now give an argument that it is surjective (onto). ASSUME that f is not surjective. Then among the points (x, y) not in $f(\mathbb{N})$, consider those whose coordinate sum $x + y$ is minimal (so this is the lowest diagonal with an omitted point), and among those choose (a, b) as the point with the smallest first coordinate. If $a \neq 1$ (so that the point is not in the first column) then $(a-1, b+1)$ precedes (a, b) on the same diagonal (since $a-1 < a$ and $(a-1)+(b+1) = a+b$) and by the choice of (a, b) we must have $(a-1, b+1) \in f(\mathbb{N})$; that is, $f(k) = (a-1, b+1)$ for some $k \in \mathbb{N}$. Then $f(k+1) = ((a-1)+1, (b+1)-1) = (a, b)$, CONTRADICTING the assumption that $(a, b) \notin f(\mathbb{N})$ when $a \neq 1$. If $a = 1$ (so that (a, b) is in the first column; also we take $b \neq 1$ so that we are not considering point $(1, 1)$) then $(a, b) = (1, b)$ and $(b-1, 1)$ is at the bottom of the preceding diagonal and so by the choice of (a, b) (as an omitted point on the lowest diagonal) we have $(b-1, 1) = f(k)$ for some

$k \in \mathbb{N}$. Then $f(k+1) = (1, (b-1)+1) = (1, b) = (a, b)$, CONTRADICTING the assumption that $(a, b) \notin f(\mathbb{N})$ when $a = 1$. Therefore the assumption is false and f is, in fact, surjective. To show that f is injective, "the details are technical and we omit them" (Gerstein, page 126). The figure above (right) hints that there are other ways to "weave" through the array of points in the first quadrant, and so other ways to construct bijection f. This figure is from my online notes for Analysis 1 (MATH 4217/5217) for Section 1.3. The Completeness Axiom in which it is being shown that the countable union of a countable collection of countable sets is countable (see Theorem 1-19).

Note. In the next example we show that some sets cannot be put in a one-to-one correspondence with $\mathbb{N}$. When we address cardinalities of infinite sets in Section 4.3. Countable and Uncountable Sets, we'll see that this implies that some infinite sets are "bigger" than others!

Example 3.19. Let F denote the set of all functions from $\mathbb{N}$ to $\mathbb{N}$. Can the members of f be listed as a sequence? That is, is there a surjection $g : \mathbb{N} \to F$?

Note. The technique used in the solution to Example 3.19 is called the "Cantor diagonalization argument." You will see this argument again in Analysis 1 (MATH 4217/5217); see my online supplemental notes for Section 1.3. The Completeness Axiom and notice the proof of Theorem 1-20: "The real numbers in $(0, 1)$ form an uncountable set." In this class, we address these ideas in Section 4.3. Countable and Uncountable Sets. Gerstein illustrates the Cantor diagonalization argument with the following image (form page 127):



Cantor's gameboard

Each function $g_n : \mathbb{N} \to \mathbb{N}$ in the solution/proof of Example 3.19 is a sequence of natural numbers. We can imagine laying out the terms of the sequences in the squares of the checkerboard, so that the first row contains the first sequence, and so forth. Then by creating a new sequence that results from taking as its first term something different from the natural number in the first darkened square (by adding 1 to it, for example, as is done in Example 3.19), we know that the newly created sequence will not equal the original first sequence (since the first term is different). Next, we move to the natural number in the second darkened square on the diagonal and make the second term in the new sequence something different from this entry, so that the new sequence is different from the second sequence because the second terms are different. Similarly, move down the diagonal picking numbers different than those in the black boxes as the entries of the new sequence. For example, if m is the entry in the kth diagonal black box, then make the kth entry of the new sequence $m+1$ (as in the Example 3.19), so that the new sequence is different for kth sequence (since they differ in the kth coordinate). We'll use this idea again in Chapter 4, "Finite and Infinite Sets."

Revised: 2/1/2022