Chapter 4

Mappings

Starting with this discussion, we begin the study of the theory of groups in earnest. To that end, one of the very first minimum viable concepts that we should become familiar with is that of *a mathematical mapping* or just *a mapping*, which is a concept that travels well throughout and is highly portable across the rest of the grown-up mathematics.

Mathematicians frequently take the everyday, spoken language, words and assign certain meanings to them that are completely foreign to the original. For example, the meanings of the following words:

- *group*
- *ring*
- *field*
- *face*
- *pencil*
- *residue*
- *pole*
- *root*

in the everyday speech are completely different from the meanings of these words in mathematics.

Sometimes mathematicians invent new words that seem to be shipped to us from a parallel universe.

For example, exactly what kind of sauce do we eat these animals with:

- *holomorphic*
- *homomorphic*
- *homeomorphic*
- *nilpotent*

and what do these words even mean?

Luckily, the meaning of the word *mapping* and the overall *concept* that sits behind the associated notion as far as mathematicians are concerned does not really betray our everyday intuition.

Indeed, when we set out to make a good old *geographic map* of a certain region, we usually do so at a reduced scale in the following way:

- if we see this bridge over there then we *correlate* that bridge to its *image* on the map
- if we see this river over there then we correlate that river to *its* image on the map
- if we see this wheat field over there then we correlate that wheat field to its image on the map
- and if we see this village over there then we correlate that village to its image on the map

Let us now take a technical survey of what just transpired.

I. During the above map-making process, *all* the original physical objects of interest were somehow depicted on the map. In other words, *to each* original physical object we assigned its image on the map, which, by the presence of the quantifier *all*, means that not a single such object was not depicted in the said process.

Exercise 4.0.1: answer on your own, what will happen to the map of this region if we were to not depict the wheat field on it?

Solution: if we were to not depict the wheat field on our map then we would have a blank spot on it, which would translate into the absence of the needed information. By that virtue, such a map will be not complete and not very useful. □

II. During the above map-making process, each original physical object of interest was depicted on the map *once and only once*.

In other words, on our geographic map we do not really see two, three or five copies of the same bridge sprinkled at different locations. On the contrary, we see exactly *one , and only one* image of this bridge and that image is exactly where it belongs.

Exercise 4.0.2: answer on your won, what will happen if our map did contain, for example, five different images of the bridge on it?

Solution: if our map did contain five different images of the bridge on it then it would introduce a phenomenon that is not welcome in mathematics and that is known as *ambiguity* because by looking at such a map, we would not be able to know with absolute certainty whether this image of the physical bridge does depict the real thing.

In other words, if our map did contain five different images of the bridge on it then we would not be able to answer the following questions with absolute certainty:

Where in reality is that bridge located? Is that bridge located here or in any one of the other four places?

In that case we would have to guess and wonder: is the actual bridge really there where *this* image says that it is? Such a map will not be very useful also. □

III. As a result of our map-making process, different original objects were shown as different images on the map. That is, *the wheat field* is distinctly different from *the river* that runs near by and *the bridge* over this river is not *the village*. Thus, when we look at the map of this region, we see distinctly different images of the real objects: one image that depicts the wheat field, for example, and a different image that depicts the river, for another example.

IV. Also as a result of our map-making process, every image on the map has its physical or real and tangible parent. Intuitively, we may say that a geographic map has no *extras* that are not accounted for in reality. In other words, when we look at our map and see just *four* objects depicted on it:

- the bridge
- the river
- the wheat field and
- the village

Then, when we actually go and visit that region, we expect that all that we will see there will be exactly four *real, life-size* entities:

- the Memorial Bridge (say)
- the Fast River (say)
- the Wide Field (say) and
- the Grouppleton Village (say)

and nothing else.

If, for whatever reason and all of a sudden, in the middle of the wheat field we see a large hill that was not depicted on the map then we frown in disappointment because we realize that the map that we are holding in our hands *is not accurate*.

Conversely, if there is an image of a large hill in the middle of the wheat field marked on the map but we do not really see that hill when we visit the region then, again, we frown in disappointment because we realize that the map that we are holding in our hands is also *not accurate*.

In summary, the process of drawing such geographic maps is now called *map-making* and the immediate and the instantaneous process of assigning or correlating this bridge to its image on that map is referred to as *mapping*.

The word *mapping* as a noun simply means *the result* of the above *process* and it may also mean *a rule* according to which the above process was carried out.

In geographic map-making such a rule boils down to getting both *the scale* and *the relative placement* of the images right. For even if we manage to correctly depict the proportions and the relative sizes of the images of the wheat field and of the village, we can still mess things up by placing the village on the wrong side of the wheat field.

Conversely, even if we place the image of the village on the correct side of the image of the wheat field, we can still mess things up if we do not capture the proportions and the relative sizes of these objects correctly.

Only when we nail both, the proportions and the relative placements of various objects, do we get a proper geographic map of this or that region.

Note that we did not really need any special mathematical training to make the above observations because such observations were driven by our common sense and the specific task at hand.

In order to formalize and mathematize such a process, we step thus.

Replace our proverbial geographic region with certain collection of elements known as *a set* and symbolize such a set with A.

A set, being a notion that in mathematics is fundamental or basic, is not given a formal definition. Intuitively, a set is a collection of distinct *elements* whose order is irrelevant and from our **Prerequisites** chapter we remember that such elements can be any kind of things: submarines, bananas, molecules, planets, forces, other sets and so on.

One, but not the only, way to record a set and its contents is as follows:

$$A = \{\text{bridge}, \text{river}, \text{wheat field}, \text{village}\}$$

Definition 1: the number of elements of a set S is called *the cardinality* of the set S and is symbolized as $|S|$.

A synonym for *the cardinality* of a set is *the size* of this set.

Definition 2: a set that has a finite number of elements is called *a finite set*.

For example, the cardinality or the size or the number of elements of the finite set $S = \{A, B, C\}$ is equal to three:

$$|S| = 3$$

Definition 3: a set that is not finite is called *an infinite set*.

The set of all integers, negative, zero and positive whole numbers, denoted $\mathbb{Z}$, is an example of an infinite set.

We will be looking at some infinite sets here and there but we will be *working* mostly with sets that have a finite number of elements.

Next, replace our geographic map with another set symbolized as B.

Relax some of the restrictions that were imposed on the way in which to the elements of the set A, formerly known as *a river*, *a village* and *a bridge*, we assigned the elements of the set B, formerly known as *the image of a bridge* and so on, via the following definition.

Definition 4: *a mapping* of a set A to a set B is a rule according to which to each element a of the set A there assigned a particular element b of the set B.

Such an element b is called *the image* of the element a under a given mapping and such an element a is called *a preimage* of the element b under a given mapping.

The name of a mapping is usually symbolized with a lower case letter, f being a popular choice, for example, but any other like letter will do.

A mapping can be defined either formulaically, formally and symbolically, or descriptively, in words.

The *from-to* direction of a mapping is symbolized with a left-to-right arrow:

$$f : A \to B$$

and is read as follows: *a map* f *from* A to B.

The set, A, that provides *the source* of the elements that are being mapped by a particular mapping is called *the domain* of the mapping, while the set, B, that provides the destination or the elements that are being mapped to by a particular mapping is called *the codomain* of the mapping.

Once a mapping f is defined, a synonym for the image b can be symbolized in the so-called *functional notation* as follows:

$$b = f(a)$$

For example, if we let $A = \{a, g, l, o\}$ and let $B = \{1, 2, 3, 4, 5, 6\}$ then we can invent the following mapping f of A to B.

Let $f : A \to B$ be a mapping according to which each element of A is assigned its sequential number in the word *algorithm*. Using the said functional notation, we, thus, can write:

$$f(a) = 1,\ f(g) = 3,\ f(l) = 2$$

and so on.

Note that we defined the above map f *verbally* and that is perfectly fine. A mathematical map can also be defined formulaically, which is an approach to defining mathematical maps that most of our readers should be familiar with. For example, the following formulaic definition of the map $f : \mathbb{R} \to \mathbb{R}$:

$$f(x) = x^2$$

can also be rendered verbally as follows: the map f of $\mathbb{R}$ to $\mathbb{R}$ correlates every real number, x, with its square, x^2.

In order to aid our comprehension of what this or that mathematical map does or how it behaves, such a map can be represented geometrically or visually in a number of different ways and it is important to understand right away that a mathematical map and any one of its many possible geometric representations are two distinct mathematical objects.

For example, geometrically the elements of the above sets A and B can be represented with *points*. Moreover, the geometric representations of the elements of the set A can be drawn as equally spaced points on a horizontal straight line, while the geometric representations of the elements of the set B can be drawn as equally spaced points in a perpendicular or vertical direction.

The heart and soul of a mapping $f : A \to B$, then, can be captured as follows.

In the chosen plane the parallel vertical straight lines v can pass through the points that represent the elements a of the mapping's domain, the set A.

In the same plane the parallel horizontal straight line h can pass through the points that represent the elements b of the mapping's codomain, the set B.

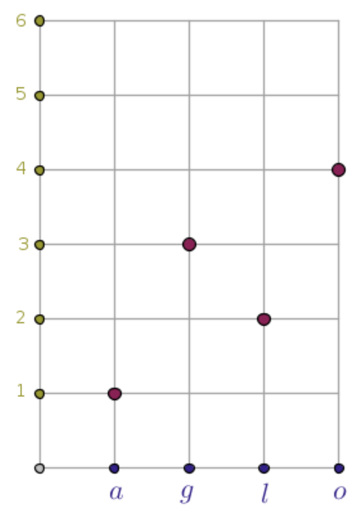
The said straight lines v and h thus constructed will form a unit square lattice or a grid.

The nodes of that grid, the points where the above straight lines v and h intersect, can then be used in order to mark the points where the verticals v that pass through the points that represent the particular preimages of the mapping f intersect with the horizontals h that pass through the points that represent the corresponding images of the mapping f.

A construct thus produced will be *a graph* of the map $f : A \to B$.

For example let the map f from the set $A = \{a, g, l, o\}$ to the set $B = \{1, 2, 3, 4, 5, 6\}$ correlate the element a to the element 1, the element g to the element 3, the element l to the element 2 and the element o to the element 4.

Then the (complete) graph of the map $f : A \to B$ may be rendered as shown below (Figure 4.0.1):



In that graph the blue points represent the elements of the set A, the green points represent the elements of the set B and the red points portray the essence of the given map f by showing which preimage of f have which image under f.

If either of the sets or both sets of a given map $f : A \to B$ are infinite then the best that we can hope for is *a partial graph* of the map f.

Once that idea is understood, we can agree to use the phrases *geometric images of* or *geometric representations of* and *elements of* a set interchangeably.

Note that in the above geometric representation of the mapping f no vertical straight line, associated with its element of the set A, goes without a point on it, as it should be with any mapping.

However, not all *horizontal* straight lines, associated with their elements of the set B, carry a point on them. That is perfectly fine.

Let us now highlight the important fine details of the **Definition 4** by paying a special attention to the word *particular* and to the universal quantifier *to each* in it.

Key points of a mapping

I. If the direction of a mapping is from a set A to a set B then no element a of the set A can have more than one image b in the set B.

The word *particular* in the **Definition 4** means that if a rule f wants to be *a mapping* then such a rule cannot allow some element(s) a of the set A to have, say, two distinct images $b_1 \neq b_2$ in the set B. Moreover, such a rule cannot allow for the said element a to have three, four, five and so on distinct images from the set B.

If a rule f wants to be a legitimate mathematical mapping then it has to assign *one and only one* image b from the codomain set, B, to a given element a of the domain set, A.

For example, if a set A consists of three people, a set B consists of, say, five rooms and a certain rule f says:

$$\text{send people to their rooms somehow}$$

then if such a rule *is* a mapping then it cannot ask a given person to go to two, or three, or four, or five different rooms. Under such a rule, any given person can go visit one room and one room only.

II. If the direction of a mapping is from a set A to a set B then it is perfectly fine for an element b of the mapping's codomain, the set B, to have more than one preimage a in the mapping's domain, the set A.

In other words, in this case the fact that an element b of the set B has two, or three, or four, or five elements of the set A assigned to it simply means that it so happens that this particular rule correlates or *maps* multiple distinct elements a of the set A to just one element b of the set B.

Reusing the above example, if our rule f is a mapping then it can send multiple, even all, people to the same room just fine.

III. Recalling the significance of the universal quantifies *any*, *each*, *every* and *all*, we would like to stress the fact that the Definition 4 *requires* that *each and every* element a of the domain, the set A, is assigned *an* image from the codomain, the set B, and it should be understood that a requirement is not an option, but a demand that must be met and satisfied by the candidate rule if that rule wants to be a mapping.

Thus, from that definition of a mathematical mapping it follows that if a certain rule, that pairs the elements a of a set A with the elements b of a set B, leaves at least one measly element a of the set A not paired up with its buddy from the set B then such a rule does not qualify to wear *a mapping* uniform and is dismissed from our consideration.

Such a rule is not a mapping, it is something else and we are not interested in it.
Reusing the above example again, if the proposed rule sends just *two* people to their respective rooms but does not send the third person anywhere then such a rule is not a mapping.

IV. If the direction of a mapping is from a set A to a set B then there does not exist any a priory restriction that the set B must be different from the set A and it is perfectly fine if the role of the *destination,* or the *to* set, or the codomain, B, is played by *the source*, or the *from* set, or the domain, A itself.

Putting the cart before the horse somewhat, in the theory of groups we will be interested in this case, when a set A is mapped to itself in a certain way, very often.

Exercise 4.0.3: given two arbitrary finite sets A and B with the cardinalities n_a and n_b respectively, compute the number of different mappings of the set A to the set B.

Solution: we will solve this problem by the tactical means of *a table*, which is a mathematical object that can be interpreted as a vague sort of yet another type of *a two-dimensional graph* of a mapping.

Namely. Let us symbolize the n_a elements of the set A as:

$$a_1, a_2, \ldots, a_{n_a}$$

and let us symbolize the n_b elements of the set B as:

$$b_1, b_2, \ldots, b_{n_b}$$

Next, because the requested mappings will be from the set A to the set B, in the top row, or the proverbial x-axis, of our table we will record the elements of the set A in a highly specific way of putting only one element of that set into one cell of such a row (Figure 4.0.2):

a_1	a_2	$\ldots$	a_{n_a}

How many columns or cells in such a row will there be?

There will be exactly n_a columns or cells in such a row.

In the bottom row, or the proverbial y-axis, of our table we will record the mystery mapping f that somehow assigns to each element a of the set A an element b of the set B.

Since in a mapping a given element b of the set B can very well have multiple preimages a from the set A , it follows that in our table's second row we may very well have *duplicate entries*.

In other words, in the second row of our table the symbols $b_1, b_2, \ldots b_{n_a}$ are not necessarily distinct, since we remember that a mapping cannot leave even a single element a of A not assigned to or not paired up with its buddy element b of B.

In other words, no single cell of the second row of our table can be left blank and not filled in with some symbol b.

Put differently yet, a mystery mapping $f : A \to B$ can very well pair all the elements of the set A with one and the same element of the set B, such as the element b_1, for example, and so on (Figure 4.0.3):

a_1	a_2	$\ldots$	a_{n_a}
b_1	b_2	$\ldots$	b_{n_a}

In each such table the contents of its top row are fixed forever and ever and will not change as we waltz our way from one mystery mapping to another. The only thing that will change from one such table to another are the contents of the second row. Thus, the number of different mappings of the set A to the set B will be equal to the number of the above tables.

Which is to say further that the number of mapping sought-after will be equal to the number of ways in which the second row of our table can be filled out with the elements of the set B.

More precisely, every cell of the second row of our table can be filled with any element of the set B. Therefore, every one of the n_a cells of the second row of our table can be filled in n_b different ways regardless of how the other cells in that row are populated.

In more detail, the first cell of the second row of our table, the cell that sits directly under the a_1 cell, can be filled in by the symbols b from B in n_b different ways.

Regardless of the contents of that first cell of the second row, the *second* cell of the second row, the cell that sits directly under the a_2 cell, can also be filled in by the symbols b from B in n_b different ways and so on.

Hence, by the Multiplication Counting Principle, there will be exactly:

$$\underbrace{n_b \cdot n_b \cdot \ldots \cdot n_b}_{n_a} = n_b^{n_a}$$

different a, b-tables and, therefore, mappings of the set A to the B.

In a slightly different but equivalent notation, there are:

$$|B|^{|A|}$$

different mappings of A to B.

In a grown-up and a somewhat loose notation that will, however, be understood by all mathematicians, there are:

$$B^A$$

different mappings of a A to a set B. $\square$

For example, if a set A has 2 elements and a set B has 3 elements then the total number of mappings of A to B is equal to $3^2 = 9$.

For example, if a set A has 3 elements and a set B has 4 elements then the total number of mappings of A to B is equal to $4^3 = 64$.
For example, if a set A has 5 elements and a set B has 7 elements then the total number of mappings of A to B is equal to $7^5 = 16,807$ and so on.

Note that intuitively speaking we managed to reduce the solution of the above problem to a situation that we are familiar with from daily life.

Consider a good old *combination lock* (Figure 4.0.4):

This particular combination lock is equipped with 5 numerical rings, each of which is engraved with 10 different digits, 0 through 9.

Each such ring also populates one preferred *position* that can be filled with *any* of the above ten digits *independently* of what digits will be dialed in in the other similar positions on the remaining rings.

A string of five such preferred positions constitutes this lock's *password*.
Therefore, by the same Multiplication Counting Principle, this particular lock can have:

$$10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 10^5 = 100,000$$

different passwords, one of which will unlock *this* lock.

We now see that:

- one numerical ring on a combination lock corresponds to our set B
- the number of digits on each such ring corresponds to the number of elements of our set B
- the number of the numerical rings on a combination lock corresponds to the number of elements of our set A
- each specific password dialed in on such a combination lock corresponds to a mystery mapping $f : A \to B$
- the number of passwords of such a combination lock corresponds to the number of possible such $f : A \to B$ mappings

Intuitively speaking, we dial a mystery mappings $f : A \to B$ with the numerical rings of a combination lock, which should really be called *a permutation* lock.

The ability to recognize and uncover a connection and a similarity between the seemingly different things in mathematics is a part of the discussed earlier *abstraction* program and that ability will be developed with time.

Introducing and playing with concrete toy models of a given problem in order to solve that problem is a perfectly fine approach. Experimenting with small and simple numbers and scenarios may actually help us hone in on a useful pattern that may allow us to unlock a solution sought-after.

We now make the *to* language in the phrase *mapping of A to B* much more pedantic and much more precise, as we limit our attention to just three important *classes* of mappings:

- *a surjection*
- *an injection* and
- *a bijection*

that are used heavily not only in the theory of groups but just about everywhere else in mathematics.

4.1 Surjections

In order to hone in on a formal definition of a surjective mapping, let us look at a specific example that should highlight the key property of that type of mappings. Imagine that there are 20 students and 17 different chairs in a classroom. Let us, further, agree that according to a certain rule, whose exact details are irrelevant at the moment, we managed to seat all 20 students in such a way that:

- not a single student is left standing
- absolutely all 17 chairs are occupied by at least one student
- no student can occupy more than one chair and
- no single chair is empty

That is it. The above rule constitutes a mapping of the students to the chairs that is:

- *surjective* or
- *a surjection* or
- *onto*

In a more mathematically refined sense, if A is the set of 20 students and B is the set of 17 different chairs in the classroom then a rule f according to which each chair in the classroom has at least one student sitting on it constitutes *a surjection* or is *onto*.

More pedantically, if A is the set of 20 students and B is the set of 17 different chairs in the classroom then a rule f according to which each chair in the classroom has at least one student sitting on it maps the set A *onto* the set B.

Note that here and now the word *onto* gains a very precise mathematical meaning.

Another example.

Assume that we have 5 different plates and 9 different apples.

Then, if we come up with a rule f according to which:

- each and every apple is assigned to a plate
- each plate has at least one apple in it
- no apple can be cut into multiple pieces and occupy multiple plates and
- no plate is left empty, without an apple in it

then such a rule will map the apples to the plates *surjectively* or will map the apples *onto* the plates.

More pedantically, if A is the set of 9 different apples and B is the set of 5 different plates then a rule f according to which each plate has at least one apple in it maps the set A *onto* the set B.

Yet another example.

Assume that we have 13 different jewelry rings and we want to beautify all 10 fingers of our both hands with them.

Then, if we come up with a rule f according to which:

- each and every jewelry ring is assigned to a finger
- each finger of our both hands is adorned with a jewelry ring
- no ring can be put on two or more fingers and
- no finger is left without a ring on it

then such a rule will map the jewelry rings *onto* the fingers.

More pedantically, if A is the set of 13 different jewelry rings and B is the set of 10 fingers on our both hands then a rule f according to which each finger has at least one jewelry ring on it maps the set A *onto* the set B.

Officially.

Definition 5: a mapping $f : A \to B$ is *surjective,* or is *a surjection,* or is *onto* if for every element b of the set B there exists such an element a of the set A that $f(a) = b$.

We remind our readers about the significance and the meaning of the universal and the existential quantifiers in mathematics. In the above definition of a surjective mapping no element b of the mapping's codomain, the set B, goes without a preimage from the mapping's codomain, the set A, and such a definition does not require that the number of such preimages is limited to just one, see below.

As we remember from our earlier discussion, the clause *there exists an element* a is perfectly satisfied if there exist two, or three, or four or more such elements a.

Intuitively speaking, the existential quantifier only establishes the smallest possible threshold for the number of elements that have a certain property, namely *just one*, and for an existential claim to be true one such element *must* exist.

However, if *multiple* elements with the required property exist then it is perfectly fine and the existential claim or requirement will hold or will be satisfied.

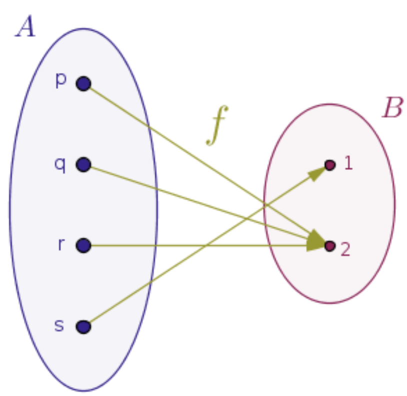
From our earlier examples:

- if we have two students sitting on one chair or
- if we have four apples on one plate or
- if we have three jewelry rings on one finger

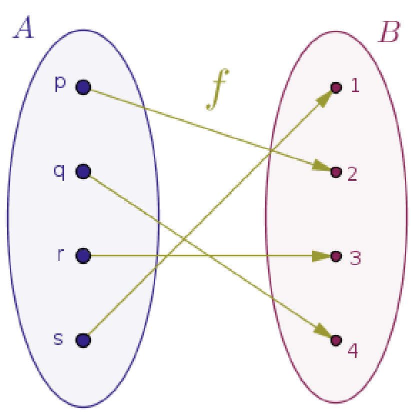
and the rest of the requirements of a surjective mapping are satisfied then the mapping under consideration *will be* surjective.

We already know that geometrically or visually the elements of a given (finite) set can be represented with points.

If geometrically or visually we depict a rule f that is *a surjective* mapping of a finite set A to a finite set B with the respective arrows, then a graph of such a mapping will have at least one arrow pointing at *each and every element* b of the set B (Figure 4.1.1):



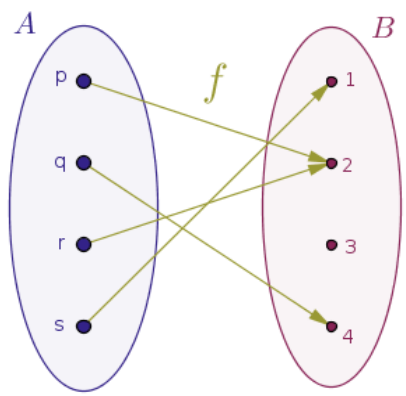
As an illustration of the above geometric requirement, the following graph depicts a mapping f that is also *onto* or *surjective* (Figure 4.1.2):



That sample mapping correlates the elements p, q, r, s of the domain, the set A, to the elements $1, 2, 3, 4$ of the codomain, the set B, in such a way that each and every element of the codomain, represented with a red point, has at least one green arrow pointing to it from the elements of the domain, which are represented by the blue points.

Consequently, since every image of the mapping f has a corresponding preimage, it follows that the mapping f depicted in the Figure 4.1.2 is surjective.

However, the next graph depicts a mapping f that is *not* onto or is *not* surjective (Figure 4.1.3):



Indeed, that mapping f correlates the elements p, q, r, s of the domain, the set A, to the elements $1, 2, 3, 4$ of the codomain, the set B, in such a way that the element 3 of the destination set, the set B, does not have a preimage at all and, by that virtue, violates the definition of a surjective mapping, **Definition 5**.

Consequently, the map f shown in Figure 4.1.3 is not surjective.

We, thus, arrive at the following key take-away points of a mapping that is *surjective*:

I. If a mapping $f : A \to B$ is surjective then it is not required that the preimage a in **Definition 5** must be unique, because under a mapping that is surjective an element b of the set B may very well have *multiple* preimages.

II. A surjective mapping of a finite set A to a set B can exist only if the set B is also *finite* and if the number of elements in the set B is not larger than the number of elements in the set A:

$$|A| \geqslant |B| \tag{1}$$

Using our illustrative examples above:

- if we have only 15 students and 17 chairs or

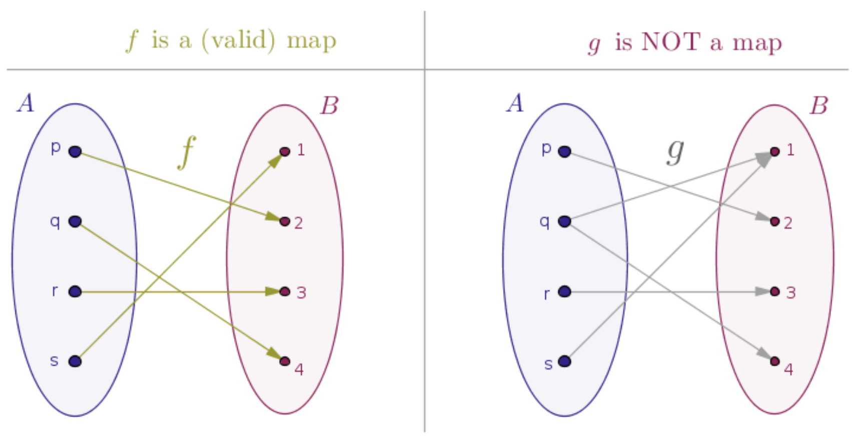
- if we have only 4 apples and 5 plates or
- if we have just 9 jewelry rings and 10 fingers

then it will be impossible to establish a surjective mapping from the sets of students/apples/jewelry rings *onto* the sets of chairs/plates/fingers respectively.

If we make, say, the above 17 chairs play the role of pigeons and if we make the above 15 students play the role of boxes then, because $17 > 15$, when we distribute 17 pigeons across 15 boxes then, according to the Dirichlet's Box Principle, at least one box will have more than one pigeon in it.

That last fact will be interpreted in this case as at least one student being assigned to sit on more than one chair, which, in turn, means that at least one preimage a of the set A will have more than one image b in the set B, which violates the definition of a mapping shown in **Definition 4**, according to which no preimage a of a set A can have more than one image b of a set B.

In order to make it perfectly clear, the rule f, depicted on the left side in the diagram below (Figure 4.1.4):



is a (valid) mathematical map because that rule correlates each and every element of its domain, the set A, to one and only one element in its codomain, the set B.

However, the rule g, depicted on the right side in the diagram above is *not* a map because that rule violates the requirement, according to which no element of the domain of a map can have more than one image in the codomain.

Namely, the rule g assigns more than one image from the set B, the elements of that set named as 1 and 4 , to one and the same element of the set A named as q.

We note in passing that even though elsewhere in mathematics, in the study of Complex Analysis, for example, the notion of *a function* is generalized into *a single-valued function* and *a multi-valued function*, in this text we will work exclusively with single-valued functions or just *functions* or *maps*.

Thus, in our running example, it is impossible to map 15 students to 17 chairs surjectively and a similar reasoning applies to the other illustrative examples.

Exercise 4.1.1: answer the following question on your own.

If A is the set of all real numbers $\mathbb{R}$, B is the set of all nonnegative real numbers $\mathbb{R}^+$ and if the mapping f of $A = \mathbb{R}$ to $B = \mathbb{R}^+$ is given by:

$$f(x) = x^2$$

for all elements x of $\mathbb{R}$ then will such a mapping be surjective?

Solution: in order to solve this problem and answer the question posed, we have to check whether the given mapping f honors the definition of *a surjective* mapping given in **Definition 5**.

We quickly see that yes, such a mapping *will* be surjective because for every element y of $\mathbb{R}^+$ there exists at least one element x of $\mathbb{R}$ such that $f(x) = y$.

Indeed, the operation of $\sqrt{y}$ is well-defined for all the elements y of $\mathbb{R}^+$ and it is sufficient to choose an element x from $\mathbb{R}$ such that:

$$x = \sqrt{y}$$

Even more so, for every element y of $\mathbb{R}^+$ there exist exactly *two* preimages under the given surjective mapping:

$$+\sqrt{y},\ -\sqrt{y}$$

In other words, in this example every element y of $\mathbb{R}^+$ has a finite number of preimages. $\square$

Exercise 4.1.2: answer the following question on your own.

If A is the set of all right triangles in the Euclidean plane T, B is the is set of all nonnegative real numbers $\mathbb{R}^+$ and if the mapping f of $A = T$ to $B = \mathbb{R}^+$ assigns to each right triangle t of T a number that constitutes *the area* of such a triangle under a fixed unit of measurement then will such a mapping be surjective?

Solution: again, in order to solve this problem and answer the question posed we have to check whether the given mapping f honors the definition of *a surjective* mapping given in **Definition 5**.

We, thus, quickly see that yes, such a mapping *will* be surjective because for every element x of $\mathbb{R}^+$ there exists a planar right triangle t the lengths a, b of whose sides are:

$$a = \sqrt{x}, \quad b = 2\sqrt{x}$$

and whose area A_t is exactly x:

$$A_t = \frac{a \cdot b}{2} = \frac{\sqrt{x} \cdot 2\sqrt{x}}{2} = x$$

Moreover, there exist infinitely many such right triangles.

For example, the right triangles the lengths a, b of whose sides are given by:

$$a = \frac{\sqrt{x}}{k}, \quad b = 2k\sqrt{x}, \quad k = 1, 2, 3, \ldots$$

all have x as their area and when $x = 0$ then a degenerate triangle with the side lengths $a = b = 0$ is a perfectly valid planar triangle.

In other words, in this example every element x of $\mathbb{R}^+$ has an infinitely many preimages t of T. $\square$

Exercise 4.1.3: answer the following question on your own.

If A is the set of all three-digit prime numbers, B is the set of all the base-ten digits and if the mapping f of A to B assigns to each three-digit prime number its middle digit then will such a mapping be surjective?

Solution: again, in order to solve this problem and answer the question posed we have to check whether the given mapping f honors the definition of *a surjective* mapping given in **Definition 5**.

Does it?

Well, in this case the set of images B is not only finite but has only ten elements in it:

$$B = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

Thus, in order to solve this problem and answer the question posed, all we have to do is check to see if for every base-ten digit d of B we can produce at least one three-digit prime number whose middle digit is exactly d.

To that end, we pull up a table of all three-digit prime numbers and find that:

$$f(101) = 0, \quad f(113) = 1, \quad f(223) = 2, \quad f(337) = 3, \quad f(443) = 4$$

and that:

$$f(557) = 5, \quad f(661) = 6, \quad f(773) = 7, \quad f(881) = 8, \quad f(997) = 9$$

Hm. Would you look at that? We managed to show, via explicit examples of three-digit prime numbers, that for every base-ten digit d there exists at least one three-digit prime number whose middle digit is exactly d.

Conclusion: the suggested mapping f *is* surjective. $\square$

Exercise 4.1.4: answer the following question on your own.

If A is the set of all real numbers $\mathbb{R}$, B is the set of all nonnegative real numbers $\mathbb{R}^+$ and if the mapping f of $A = \mathbb{R}$ to $B = \mathbb{R}^+$ is given by:

$$f(x) = x^2 + 1$$

for all elements x of $\mathbb{R}$ then will such a mapping be surjective?

Solution: again, in order to solve this problem and answer the question posed we have to check whether the given mapping f honors the definition of *a surjective* mapping given in **Definition 5**.

We quickly see that, while for all the elements y of $\mathbb{R}^+$ that are 1 *or larger* there exist exactly two preimages from $\mathbb{R}$, namely the elements $+\sqrt{y-1}$ and $-\sqrt{y-1}$, it is no longer the case that such preimages exist for *all* elements y of $\mathbb{R}^+$.

Indeed, from our earlier discussion of various quantifiers we remember that just one specimen that does not possess a required property is all that it takes to refute or invalidate a universal claim.

Here is such a specimen, also known as *a counterexample*: for the element $y = 0$ of $\mathbb{R}^+$ there does not exist an element x of $\mathbb{R}$ such that $0 = x^2 + 1$.

In general, over the field of complex numbers, the equation $z^2 + 1 = 0$ does admit two distinct solutions, namely the numbers $z_1 = +i$ and $z_2 = -i$, but these numbers are not real.

Hence, the suggested mapping f is *not* surjective. $\square$

After we get some experience and confidence with computing the number of possible *injections* and *bijections* of one finite set to another finite set, later on, we will compute the number possible *surjections* between such sets, in the advanced **Exercise 4.3.6**.

4.2 Injections

In order to hone in on a formal definition of a surjective mapping, let us look at a specific example that should highlight the key property of that type of mappings. Imagine that there are 20 students and 23 different chairs in a classroom. Let us, further, agree that according to a certain rule, whose exact details are irrelevant at the moment, we managed to seat all 20 students in such a way that:

- not a single student is left standing
- each and every student sits on *a separate* chair
- an occupied chair has exactly and only one student sitting in it
- no student can occupy more than one chair and
- some chairs may very well be empty

That is it. The above rule constitutes a mapping of the students to the chairs that is:

- *injective* or
- *an injection* or
- *one-to-one*

In a more mathematically refined sense, if A is the set of 20 students and B is the set of 23 different chairs in the classroom then a rule f according to which each and every student sits on a separate chair all by her/himself constitutes *an injection* or is *one-to-one*.

More pedantically, if A is the set of 20 students and B is the set of 23 different chairs in the classroom then a rule f according to which each and every student sits on a separate chair all by her/himself maps the set A to the set B *injectively* or *one-to-one*.

Another example.

Say, we have 9 different plates and 5 different apples.

If we come up with some rule f according to which:

- not a single apple is left without a plate
- each and every apple is placed into its own, separate, plate all by itself
- an occupied plate has exactly and only one apple in it
- no apple can be cut into multiple pieces and occupy multiple plates and

- some plates may be empty

then such a rule will map the apples to the plates *injectively* or will map the apples to the plates *one-to-one*.

In this particular case we will have some plates that are empty. This is perfectly fine.

More pedantically, if A is the set of 5 different apples and B is the set of 9 different plates then a rule f according to which each and every apple is placed into its own, separate, plate all by itself maps the set A to the set B *injectively* or *one-to-one*.

Yet another example.

Say, we have 7 different jewelry rings and we wish to beautify the 10 fingers on our both hands with them.

If we come up with some rule f according to which:

- not a single jewelry ring is not assigned to a finger
- each and every such ring is on a separate finger all by itself
- a finger adorned with a ring has exactly and only one ring on it
- no ring can be put on two or more fingers and
- some fingers may not have a ring on them

then such a rule will map the jewelry rings to our fingers *injectively* or *one-to-one*.

In this particular case we will also have some fingers without a ring on them. This is perfectly fine.

More pedantically, if A is the set of 7 different jewelry rings and B is the set of 10 fingers on our both hands then a rule f according to which each and every jewelry ring is put on its own, separate, finger all by itself maps the set A to the set B *injectively* or *one-to-one*.

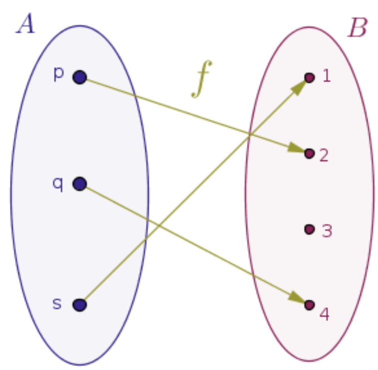
Officially.

Definition 6: a mapping $f : A \to B$ is *injective,* or is *an injection,* or is *one-to-one* if distinct elements of the set A are paired up by that mapping with distinct elements of the set B or if for all elements a_1, a_2 of the set A from the fact that $a_1 \neq a_2$ it follows that $b_1 = f(a_1) \neq f(a_2) = b_2$, where b_1 and b_2 are the elements of the set B that are the images of the elements a_1, a_2 under f.

Note that the above definition of an injective mapping does *not* insist on the fact that *all* elements of the domain have at least one preimage, as was the case with a surjective mapping.

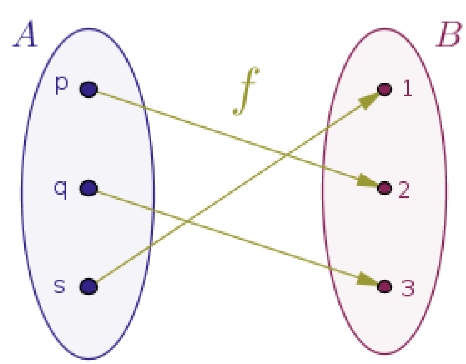
In that respect, all that the above definition does require is that *different* preimages are taken by a proposed mapping into *different* images.

Thus, the graph shown in the Figure 4.2.1 below depicts a perfectly valid injective mapping f of the set $A = \{p, q, s\}$ to the set $B = \{1, 2, 3, 4\}$ even though the element 3 of the codomain, the set B, does not have a preimage at all (Figure 4.2.1):



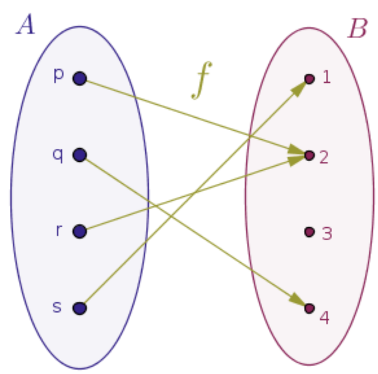
Note that in the above example the cardinality of the mapping's domain, the set A, that has 3 elements in it, is strictly smaller than the cardinality of the mapping's codomian, the set B, that has 4 elements in it. An injective mapping, however, may very well be constructed between the sets that have the same cardinality.

As an example, the graph shown in the Figure 4.2.2 below also depicts a perfectly valid injective mapping f of the set $A = \{p, q, s\}$ to the set $B = \{1, 2, 3\}$ (Figure 4.2.2):



We see that in this particular case the cardinality of the domain, the set A, and the cardinality of the codomain, the set B, is one and the same: the set A has 3 elements in it and the set B has 3 elements in it as well.

However, in the Figure 4.2.3 below the mapping f of the set $A = \{p, q, r, s\}$ to the set $B = \{1, 2, 3, 4\}$ is *not* injective because two different elements, p and r, of the set A have one and the same image, 2, in the set B, which violates the **Definition 6** of an injective mapping (Figure 4.2.3):



In other words, in this case it is possible to exhibit a pair of elements from the mapping's domain that violates the distinctness of corresponding images under that mapping: from the fact that $p \neq r$ it does *not* at all follow that $f(p) \neq f(r)$, as a mapping that is *injective* requires, because $f(p) = 2 = f(r)$.

By analogy with our discussion of *surjective* mappings, if we depict a rule $f : A \to B$ that is *an injective* mapping of a finite set A to a finite set B with the respective arrows, then in a graph of such a mapping no element b of B can have more than one arrow pointing to it and if such an element does have an arrow pointing to it then that arrow must be exactly and only one.

Recalling the mechanics of a mathematical implication IF P THEN Q, we see that it is possible to construct an alternative and equivalent definition of an injection that is based on the contrapositive of the implication used in the **Definition 6**.

To that end, in the above implication we swap its negated premise with its negated conclusion.

Definition 6.1: a mapping $f : A \to B$ is *injective,* or is *an injection*, or is *one-to-one* if for all elements a_1, a_2 of the set A from the fact that $f(a_1) = f(a_2)$ it follows that $a_1 = a_2$.

When and why should one definition of an injective map be used instead of the other?

The particular definition of an injective map is used when it is more convenient. For example, in certain situations it may be easier to show that when the images of a map are different then the preimages of that map must be different as well. In which case the **Definition 6.1** of an injective map should be used.

We, thus, arrive at the following key take-away points of an injective mapping.

I. If a mapping $f : A \to B$ is injective then it is not required that *every* element of its codomain, the set B, has a preimage from it domain, the set A. Under a mapping that is injective an element b of B may very well not have a preimage at all but if an element b of B does have a preimage than that preimage must be exactly and only one preimage.

II. If the sets A and B are both finite and there exists an injection of A to B then, clearly, the number of elements in the set A cannot be larger than the number of elements in the set B:

$$|A| \leqslant |B| \tag{2}$$

Using our illustrative examples above:

- if we have 20 students and 17 chairs or
- if we have 9 apples and 5 plates or
- if we have 13 jewelry rings and 10 fingers

then it will be impossible to establish an injective mapping from the sets of students/apples/jewelry rings to the sets of chairs/plates/fingers respectively.

If we make, say, the above 5 plates play the role of boxes and if we make the above 9 apples play the role of pigeons then, because $9 > 5$, when we distribute 9 pigeons across 5 boxes, according to the Dirichlet's Box Principle, at least one box will have more than one pigeon in it.

That last fact will be interpreted in this case as at least one plate having more than one apple in it, which, in turn, means that to at least one image b, or a certain plate, of the set B, or the set of 5 plates, there will be assigned more than one preimage a, an apple, of the set A, or the set of 9 apples, which violates the **Definition 6** of an injective mapping, according to which no two distinct preimages $a_{1,2}$ of a set A can be assigned to the same image b of the set B.

Thus, it is impossible to injectively map 9 apples to 5 plates.

The similar reasoning applies to the other illustrative examples.

Exercise 4.2.1: if A is the set of all integers $\mathbb{Z}$, B is the set of all *even* integers and if the mapping f of $A = \mathbb{Z}$ to B is given by:

$$f(k) = 6k$$

for all integers k of $\mathbb{Z}$ then will such a mapping be injective?

Solution: in order to determine whether the given map is injective or not, it is convenient to trawl that mapping through the **Definition 6**.

Doing so, we see that in this case for all integers k_1 and k_2 from the fact that $k_1 \neq k_2$ it does follow that:

$$f(k_1) = 6k_1 \neq 6k_2 = f(k_2)$$

Hence, *yes*, the given map *is* injective. □

Exercise 4.2.2: if A is the set of all real numbers $\mathbb{R}$, B is the set of all nonnegative real numbers $\mathbb{R}^+$ and if the mapping f of $A = \mathbb{R}$ to $B = \mathbb{R}^+$ is given by:

$$f(x) = x^2$$

for all x of $\mathbb{R}$ then will such a mapping be injective?

Solution: well, we quickly see that for $x_1 = -3 \neq 3 = x_2$, for example, it follows that:

$$f(-3) = (-3)^2 = 9 = 3^2 = f(3)$$

implying that the images of these two different preimages are, in fact, one and the same.

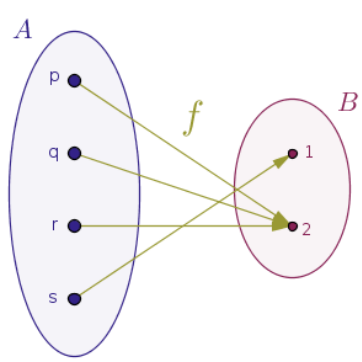
Since the requirement of an injective mapping to take different preimages into different images is violated, it follows that the suggested mapping f is *not* an injection or is *not* one-to-one. □

Comparison

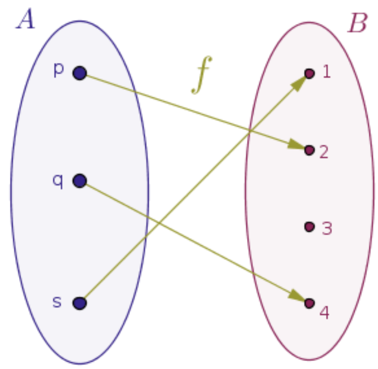
Just by studying the sample graphs of surjective and injective mappings we conclude that:

- a surjective mapping does not at all have to be injective and that
- an injective mapping does not at all have to be surjective

More formally, there exist surjective mappings that are not injective (Figure 4.1.1):



and there exist injective mappings that are not surjective (Figure 4.2.1):



We invite our readers to explain on their own why this is so in both cases.

Exercise 4.2.3: think of your own everyday example of a mapping that is injective but is not surjective and think of your own formal example of a mapping that is injective but is not surjective.

Solution: for an everyday example of a mapping that is injective but is not surjective consider the set P of 183 passengers on a flight from New-York to Los-Angeles in a Boeing 767 airplane that has 216 numbered seats as the set S.

The mapping of the set P to the set S for that flight, that was constructed by an airline company that operated that jet on that day via the process of ticket sales, is certainly *in*jective because for every two elements $p_{1,2}$ of the set P, or every two passengers, from the fact that the elements p_1 and p_2, or two passengers, are distinct, that is, from the fact that $p_1 \neq p_2$, it does follow that their images $s_{1,2}$ from the set S, or the numbered seats, under the given mapping are indeed different, $s_1 \neq s_2$, since different passengers are assigned different seats.

However, such a mapping is certainly not *sur*jective because it is not the case that every element s of S, or every seat, has a preimage p from the set P, or a passenger in it, since that flight is not filled to its capacity.

If, on the other hand, that flight was overbooked and 230 passengers with 230 tickets sold by this airline showed up for the journey from New-York to Los-Angeles in a 216-seat airplane then such a mapping of the passengers to the airplane seats will be *sur*jective but not *in*jective. Explain why.

For a formal example of a mapping that is injective but is not surjective consider the mapping f from the set A of natural numbers $\mathbb{N}$ to the set B of natural numbers $\mathbb{N}$, according to which each element N of A is taken into the element $f(n) = n + 1$ of B.

Such a mapping is certainly *in*jective because for every two elements m, n of A from the fact that $m \neq n$ it follows that $f(m) = m + 1 \neq n + 1 = f(n)$.

However, such a mapping is certainly not *sur*jective because it is not the case that every element b of B has its preimage in A. For example, the very first element of B, say, the element 0, does not have its preimage in A, because no element x of A is a solution of the equation $0 = x + 1$. $\square$

Exercise 4.2.4: given a finite set A of cardinality n_a:

$$A = \{a_1, a_2, \ldots, a_{n_a}\}$$

and a finite set B of cardinality n_b:

$$B = \{b_1, b_2, \ldots, b_{n_b}\}$$

compute the number of possible injections from the set A to the set B.

Solution: from the **Definition 6** of an injective mapping, it follows that if the number of elements n_b of the set B is less than the number of elements n_a of the set A:

$$n_a > n_b$$

then no injections from A to B exist and the answer to the question posed is *zero*.

We, thus, assume now that the number of elements of the set B is not less than the number of elements of the set A:

$$n_a \leqslant n_b$$

In that case we will use the same tactical device that we have used in the solution of the **Exercise 4.0.3**, a table with just two rows and exactly n_a columns, where n_a is the number of elements of the set A.

In the first row of such a table we will record all the elements a of the finite set A once and for all, populating exactly one column of the table with exactly one element a of A.

In the second row of our table we will place n_a distinct symbols, each of which represents an element b of the set B, in exactly the same fashion of one symbol per cell (Figure 4.2.4):

a	a_1	a_2	$\ldots$	a_{n_a}
$f(a)$	b_1	b_2	$\ldots$	b_{n_a}

which means that the number of different elements b in the second row of our table may be less than the number of elements n_b in the set B.

However, do realize that in general from the set of n_b distinct symbols b we will have to pick n_a such symbols.

For example, while mapping the smaller set $A = \{1, 2, 3, 4\}$ to the larger set $B = \{a, b, c, d, e, f, g\}$ injectively, we may assign the elements of the set B to the elements of the set A injectively as follows (Figure 4.2.5):

a	1	2	3	4
$f(a)$	a	b	c	d

and we see that that particular map left the elements e, f, g of the set B out.

Consequently, one such table, from the way it was constructed, will represent one possible injection of the set A to the set B.

Thus, the number of different injections from the set A to the set B sought-after will be equal to the number of different tables constructed as shown above.

Which is to say further that the number of such injections will be equal to the number of ways in which the second row of our table can be filled out with the elements b of the set B.

However, if in the earlier **Exercise 4.0.3** it was possible to fill any given cell of the second row of our table with *any* element b of B *regardless* of how the other cells in that row were filled in, in the case of an injection that freedom is no longer there!

Why?

Try to answer this question on your own, before reading on.

Because, by the definition of a injection, all the entries in the second row of our table must be *distinct* or, in negative terms, no duplicate symbols b of the set B in the second row are allowed.

But we also remember the fundamental requirement or the fundamental mandate of a mapping to leave no preimage a of the set A without its buddy image from the set B. In other words, no cell in the second row of our table can go empty.

Reconciling these two observations against what has to be done, we reason thus.

We have exactly n_b symbols b of the set B to choose from in order to fill in the first, leftmost, cell of the second row of our table.

Once such a choice has been made and regardless of what actual symbol b of the set B has been picked, there will be only $(n_b - 1)$ symbols b left to choose from for the candidates with which the second, or the next leftmost, cell in the second row of our table can be filled in.

That elimination of the chosen symbol b from the menu or from the pool of the consequent choices ensures that that symbol will not be repeated anywhere else, which is precisely what is needed for a mapping that is injective.

Thus, by the Multiplication Counting Principle, the first *two* leftmost cells of the second row of our table can be filled in in $n_b \cdot (n_b - 1)$ different ways.

Reasoning in the same fashion, we see that once the first two leftmost cells of the second row of our table are filled in and regardless of the actual symbols b that were used for that purpose, there will be $(n_b - 2)$ choices left for the symbols b with which the third leftmost cell of the second row of our table can be filled in.

Therefore, by the same counting principle, there will be $n_b \cdot (n_b - 1) \cdot (n_b - 2)$ different ways to fill in the first *three* leftmost cells in the second row of our table.

And so on. Thus, there will be only $(n_b - (n_a - 1))$ symbols b with which the last, rightmost, cell of the second row of our table can be filled in.

Hence, the number of different ways to fill out the second row of our table and, thus, the number of different such tables will be given by:

$$n_b \cdot (n_b - 1) \cdot (n_b - 2) \cdot \ldots \cdot (n_b - n_a + 1) \tag{3}$$

and if we multiply and divide the above result by the factorial of the remaining number $(n_b - n_a)$ then the above result can be written in a shape:

$$\frac{n_b!}{(n_b - n_a)!} \tag{4}$$

that is equivalent to the original formula shown in (**3**).

Thus, the answer to the question posed is. The number of different injections from a finite set A of cardinality n_a to a finite set B of cardinality n_b is:

- *zero* if $n_a > n_b$
- shown in (**3**) and (**4**) otherwise

The numbers shown in (**3**) and (**4**) also correspond to the number of the so-called *linear permutations* of n_b distinct items taken n_a at a time. □

As several concrete examples, if a set A has 2 elements and a set B has 3 elements then the total number of injections from A to B is equal to $3!/(3-2)! = 6$.

If a set A has 3 elements and a set B has 4 elements then the total number of injections from A to B is equal to $4!/(4-3)! = 24$.

If a set A has 5 elements and a set B has 7 elements then the total number of injections from A to B is equal to $7!/(7-5)! = 2520$. For comparison purposes, the total number of *all possible* mappings of A to B, as we computed previously, in the **Exercise 4.0.3**, will be $7^5 = 16807$.

4.3 Bijections

Continuing with our tradition of looking at simple and everyday life honing examples, imagine that we have 20 students and 20 different chairs in a classroom. Let us agree that according to a certain rule, whose exact details are irrelevant, we seated all 20 students in such a way that:

- not a single student is left standing
- each and every student sits on a separate chair all by her/himself
- an occupied chair has exactly and only one student in it
- no student can occupy more than one chair and
- not a single chair is empty

That is it. The above rule constitutes a mapping of students to chairs that is:

- *bijective* or
- *a bjection* or
- *a one-to-one correspondence*

In a more mathematically refined sense, if A is the set of 20 students and B is the set of 20 different, perhaps *numbered* with unique numbers or letters, chairs in the classroom then a rule f according to which each students sits on a separate chair all by her/himself, an occupied chair has exactly and only one student in it and not a single chair is empty constitutes *a bijection* or is *a one-to-one correspondence*.

More pedantically, if A is the set of 20 students and B is the set of 20 different chairs in the classroom then a rule f according to which each student sits on a separate chair all by her/himself, an occupied chair has exactly and only one student in it and not a single chair is empty maps the set A to the set B *bijectively* or establishes *a one-to-one correspondence* between the sets A and B.

Another example.

Say, we have 9 different plates and 9 different apples.

If we come up with a certain rule f according to which:

- not a single apple is left not on a plate
- each apple is placed into its own, separate, plate all by itself
- an occupied plate has exactly and only one apple in it
- no apple can be cut into multiple pieces and occupy multiple plates and
- not a single plate is empty

then such a rule will map the apples to the plates *bijectively* or will establish *a one-to-one correspondence* between the apples and the plates.

More pedantically, if A is the set of 9 different apples and B is the set of 9 different plates then a rule f according to which each apple is placed into its own, separate, plate all by itself, an occupied plate has exactly and only one apple in it and not a single plate is empty maps the set A to the set B *bijectively* or establishes *a one-to-one correspondence* between the sets A and B.

Yet another example.

Say, we have 10 different jewelry rings and we wish to beautify all 10 fingers on our both hands.

If we come up with a certain rule f according to which:

- no single ring is not on a finger
- each and every ring is assigned to a separate finger
- a finger with a ring on it has exactly and only one ring on it
- no ring can be put on two or more fingers and
- no finger goes without a ring on it

then such a rule will map the jewelry rings to our fingers *bijectively* or such a mapping will establish *a one-to-one correspondence* between the jewelry rings and the fingers of our two hands.

More pedantically, if A is the set of 10 different jewelry rings and B is the set of 10 fingers on our both hands then a rule f according to which each jewelry ring is put on its own, separate, finger all by itself, a finger with a ring on it has exactly and only one ring on it and not a single finger is left without a ring maps the set A to the set B *bijectively* or establishes *a one-to-one correspondence* between the sets A and B.

Officially.

Definition 7: a mapping $f : A \to B$ is *bijective* or is *a bijection* or is *a one-to-one correspondence* between the set A and the set B if it is both surjective and injective.

Another way of putting that a given mapping is bijective is to say that it is both *one-to-one and onto*.

Now let us think together. If there exists a bijective mapping between a finite set A and a finite set B then, because such a mapping is *surjective*, the number of elements of the set B cannot be larger than the number of elements of the set A:

$$|A| \geqslant |B|$$

By the same token, because our proposed mapping is also *injective*, the number of elements of the set A cannot be larger than the number of elements of the set B:

$$|A| \leqslant |B|$$

But since a mapping under consideration is both surjective *and* injective, it follows that a bijection between two finite sets A and B can exist if and only if the cardinalities of these sets are, in fact, equal one another:

$$|A| = |B|$$

which is exactly the case in all our illustrative examples, in each of which we establish a one-to-one correspondence between:

the set of 20 students and the set of 20 chairs
the set of 9 apples and the set of 9 plates and
the set of 10 jewelry rings and the set of 10 fingers on our both hands

---

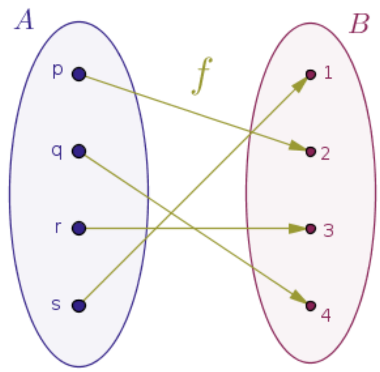
Comparison

From our earlier comparison of surjective and injective mappings, we remember that there exist surjective mappings that are not injective and that there exist injective mappings that are not surjective due to the definitions of such mappings that leave the door for such scenarios wide open.

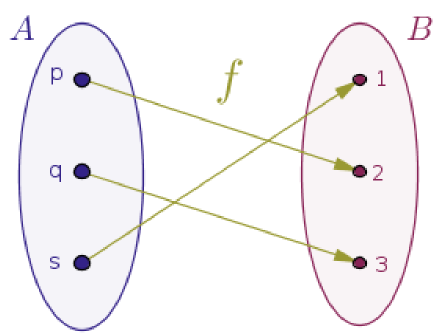
However, that does not mean that there cannot exist mappings that are both injective and surjective at the same time.

On the contrary, there are plenty of mapping that *are* both injective and surjective.

For example, in the Figure 4.1.2, shown below for easier reference, we already saw a surjective mapping that is also injective but we did not know that back then (Figure 4.1.2):

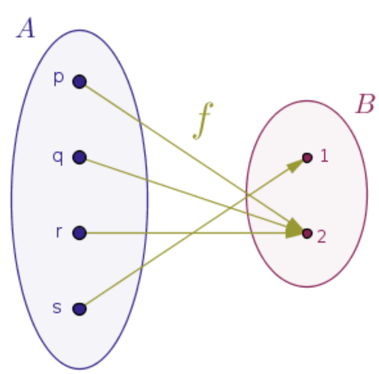


Likewise, in the Figure 4.2.2, shown below for easier reference, we already saw an injective mapping that is also surjective and even though we did not explicitly state that observation back then, we could have already deduced it on our own (Figure 4.2.2):



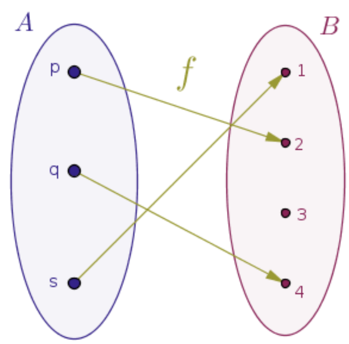
Now, however, we understand that both diagrams shown above constitute the mappings that are injective and surjective at the same time and, hence, are bijective.

In order to make it perfectly clear, our original diagram in Figure 4.1.1:



depicts a mapping that is *only surjective* but is not injective and, hence, is not bijective.

And our original diagram in Figure 4.2.1:



depicts a mapping that is *only injective* but is not surjective and, hence, is not bijective.

We invite our readers to explain *why* in both cases, on their own.

Exercise 4.3.1: if A is the set of all integers $\mathbb{Z}$, B is the set of all even integers and if the mapping f of $A = \mathbb{Z}$ to B is given by:

$$f(k) = 2k$$

for all k of $\mathbb{Z}$ then will such a mapping be bijective?

Solution: in order to solve this problem, we have to trawl the given mapping and the sets that it acts on through the **Definition 7**.

We see that the mapping f is certainly *surjective* because for every element $b = 2k$ of the set of all even integers B there exists at least one element $a = k$ of the set A such that $f(a = k) = 2k = b$.

We also see that the mapping f is certainly *injective* because for all integers k_1 and k_2 of the set A from the fact that $k_1 \neq k_2$ it follows that:

$$f(k_1) = 2k_1 \neq 2k_2 = f(k_2)$$

Thus, we conclude that the suggested mapping f *is* bijective. □

Exercise 4.3.2: if A is the set of all real numbers $\mathbb{R}$, B is the set of all nonnegative real numbers $\mathbb{R}^+$ and if the mapping f of $A = \mathbb{R}$ to $B = \mathbb{R}^+$ is given by:

$$f(x) = x^2$$

for all x of $\mathbb{R}$ then will such a mapping be bijective?

Solution: in the **Exercise 4.1.1** we have shown already that such a mapping is certainly *surjective*.

However, in the **Exercise 4.2.2** we have shown that such a mapping is *not* injective.

It follows, then, that the given mapping f is *not* bijective. $\square$

Exercise 4.3.3: can the **Exercise 4.3.2** be salvaged? In other words, is it possible to shake that exercise in such a way that its mapping *will* be bijective?

Solution: yes, it is possible to salvage the **Exercise 4.3.2** by, intuitively speaking, *narrowing* its domain or the set A.

Namely, if A is the set of all *nonnegative* real numbers, B is the set of all nonnegative real numbers and if the mapping f of $A = \mathbb{R}+$ to $B = \mathbb{R}^+$ is given by:

$$f(x) = x^2$$

for all x of A= $\mathbb{R}^+$ then that mapping f will be bijective.

Such a mapping will certainly be *surjective*, explain *why*.

Such a mapping will now be *in*jective as well for, since we have eliminated all the *negative* real numbers from the mappings domain, the set A, it will now be the case that for all nonnegative real numbers x_1 and x_2 from the fact that $x_1 \neq x_2$ it follows that:

$$f(x_1) = x_1^2 \neq x_2^2 = f(x_2)$$

As such, under the conditions spelled out above, the mapping f will be bijective. $\square$

Exercise 4.3.4: given an arbitrary finite set A:

$$A = \{a_1, a_2, \ldots, a_n\}$$

and an arbitrary finite set B of the same cardinality as that of the set A:

$$B = \{b_1, b_2, \ldots, b_n\}$$

compute the number of possible bijections from the set A to the set B.

(if the cardinalities of the sets A and B are different, that is if $|A| \neq |B|$, then no bijections from A to B exist)

Solution: in order to solve this problem we will use the same tactical device that we have met in the solution of the **Exercise 4.0.3** and the **Exercise4.2.4**, a table with just two rows and exactly n columns, where n is the number of elements of the set $A: \ n = |A|$.

In the first row of such a table we will record all the elements a of the finite set A once and for all, populating exactly one column of the table with exactly one element a of A.

In the second row of our table we will place n distinct symbols, each of which represents an element b of the set B, in exactly the same fashion of one symbol per cell (Figure 4.3.1):

a	a_1	a_2	$\dots$	a_n
$f(a)$	b_1	b_2	$\dots$	b_n

The leftmost column in the above table that carries the symbols a and $f(a)$ are just a convenient reminder that in the first row of such a table there live the elements a_k of the domain of the future bijections and that in the second row of such a table there live the elements b_k of the codomain of such mappings $f(a)$.

Consequently, one such table, from the way it was constructed, will represent one possible bijection from the set A to the set B.

Thus, the number of *different* bijections from the set A to the set B sought-after will be equal to the number of different tables, each of which is constructed as explained.

Which is to say further that the number of required bijections will be equal to the number of different ways in which the second row of our table can be filled out.

However, if in the **Exercise 4.0.3** it was possible to fill any given cell of the second row of our table with *any* element b of the set B *regardless* of how the other cells in that row were filled in, in the case of a bijection that freedom is no longer there!

Why.

Because, by the definition of a bijection, all the entries in the second row of our table must be distinct or, in negative negative terms, no duplicate symbols b of the set B in the second row are allowed.

But we also remember the fundamental requirement or the fundamental mandate of a mapping to leave no preimage a of the domain A without its buddy image from the mapping's codomain, the set B. In other words, no cell in the second row of our table can go empty.

Reconciling these two observations against what has to be done, we reason thus.

We have exactly n different symbols b of B to choose from in order to fill in the first, leftmost, cell of the second row of our table.

Once such a choice has been made and regardless of what actual symbol b of B has been picked, there will be only $(n-1)$ different symbols b left to choose from for the candidates with which we will fill in the second, or the next leftmost, cell in the second row of our table.

That elimination of the chosen symbol b from the menu of the consequent choices ensures that that symbol will not be repeated anywhere else in the second row of our table.

Thus, by the Multiplication Counting Principle, the first two leftmost cells of the second row of our table can be filled in in $n \cdot (n-1)$ different ways.

Reasoning in the same fashion, we see that once the first two leftmost cells of the second row of our table are filled in and regardless of the actual symbols b that were used for that purpose, there will be $(n-2)$ choices left for the symbols b with which the third leftmost cell of the second row of our table can be filled in.

Therefore, by the same counting principle, there will be $n \cdot (n-1) \cdot (n-2)$ different ways to fill in the first three leftmost cells in the second row of our table.

And so on. It stands to reason that there will be only one symbol b with which the last, rightmost, cell of the second row of our table can be filled in.

Hence, the total number of different ways to fill out the second row of our table and, thus, the number of different such tables will be given by:

$$n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1 = n!$$

and that is the answer to the question posed: the number of different bijections from a finite set A to a finite set B of the same cardinality $n = |A| = |B|$ is equal to $n!$ or n factorial. $\square$

The values of such a factorial grow rather aggressively as the number n grows.

For example, if our sets have just *two* elements each then the number of possible bijections from A to B will be just $2! = 2$.

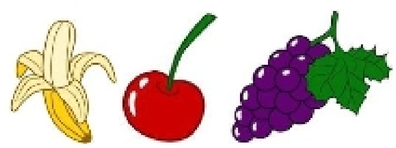
If $|A| = 4 = |B|$ then the number of possible bijections from A to B will be $4! = 24$. For comparison purposes, the total number of *all possible* mappings from A to B will be $4^4 = 256$, which is more than 10 times larger than the number of possible bijections from A to B.

If $|A| = 6 = |B|$ then the number of possible bijections from A to B will be $6! = 720$. For comparison purposes, the total number of *all possible* mappings from A to B will be $6^6 = 46656$, which is more than 60 times larger than the number of possible bijections from A to B.

If $|A| = 10 = |B|$ then the number of possible bijections from A to B will be $10! = 3628800$. For comparison purposes, the total number of *all possible* mappings from A to B will be 10^{10}, which is more than 2700 times larger than the number of possible bijections from A to B, and so on.

Remembering that any kind of things can be the elements of a set, let us do the following

Exercise 4.3.5: given the set A with the cardinality of 3 (Figure 4.3.2):



and the corresponding set B of the same cardinality (Figure 4.3.3):

write down, explicitly, all the possible bijections from the set A to the set B.

Solution: the solution of the previous exercise tells us that the number of different bijections from a finite set A of cardinality n to a set B of the same cardinality is equal to $n!$.

But $n!$ is also the number of ways in which n distinct things can be arranged in a straight line in different order, where any one such order-sensitive arrangement is called *a linear permutation* of n things.

As we already observed at the end of the solution of the previous exercise, with just two sample items a and b, the linear arrangement of these items ab is already different from the linear arrangement ba.

We also remember that in any given bijection from the (finite) set A to the (finite) set B the top row of the table that we used in the previous exercises will always be fixed.

Thus, we reduce the problem of generating all the possible bijections from the set A to the set B to the problem of generating all the possible linear permutations of $n = 3$ distinct elements b of B.

It turns out that there exist different ways of varying efficiency that can accomplish that task in general, for any finite whole positive number n.

One such way is known as *the Johnson-Trotter algorithm* and another such way is known as *the Narayana's algorithm*.

Between the two such algorithms, the Johnson-Trotter algorithm is considered to be more efficient in the sense that in order to generate the next linear permutation from the previous one, that algorithm goes through the smallest number of rearrangements possible.

We describe the inner workings of both the Johnson-Trotter and the Narayana's algorithms in depth in the Appendix A.

In our sample solution we will use the Johnson-Trotter algorithm.

Let us now agree to dispense with the elaborate table format of constructing the bijections and simply line up the elements of a set A with the elements of the corresponding set B vertically.

As such, according to the Johnson-Trotter algorithm, we write down our first bijection from the set A to the set B any which way we want it and we name the row of the elements of the set B as *the current* (linear) permutation.

We pick the following linear arrangement of the given elements (Figure 4.3.4):

which is the first of the total of $n! = 3! = 6$ bijections sought-after, but our readers should experiment with different choices.

We now trawl the rightmost element of the set B, the drum, leftward, one swap with its left neighbor (Figure 4.3.5):

at a time (Figure 4.3.6):

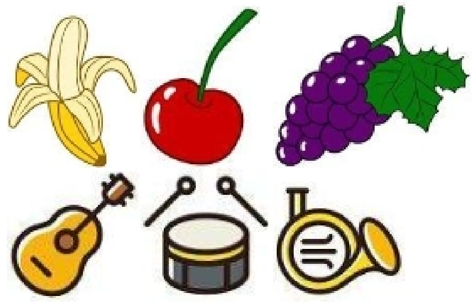
until our drum hits the left invisible *wall* and cannot go any further.

At which point we do the same thing to the new leftmost item in the bottom row, the guitar, swapping it once with the French horn (Figure 4.3.7):

and then we change the direction in which our drum travels to the opposite.

That is, since initially the drum moved in the direction *from right to left*, we now change that direction to *left-to-right* and swap the drum with whatever is on its *right*.

During the first such swap, the drum trades places with the guitar (Figure 4.3.8):



and so on until the said drum hits the right invisible wall and cannot go any further (Figure 4.3.9):

at which point all 6 different bijections from the set A to the set B have been generated.

We recommend that our readers who are not shy with a keyboard copy, compile and run the C code that we show in our discussion of the Narayana's permutations generating algorithm - for the *at will* experimentation purposes. □

Cultivating the ground for the upcoming study of permutations, we observe that in order to make it abundantly clear where the elements of a set begin and where the elements of a set end in the table-like structures that we used in the previous problem, it is customary to surround such a two-row table or a two-row matrix with parentheses like so (Figure 4.3.10):

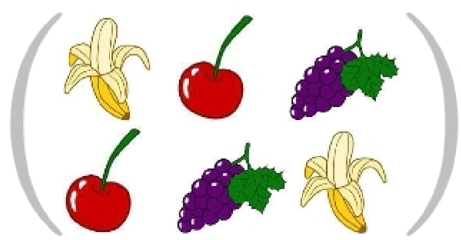
With the sample elements that we chose for the last exercise the above parentheses may look a bit unwieldy but with numbers and single-letter symbols such a parenthesized table looks quite neat.

Bijecting a Set Onto Itself

In our introduction to the theory of groups, as we already mentioned earlier, we will be often interested in the bijections of a finite set onto itself - our definition of a bijection in general and the current discussion of the properties of a bijection in particular do not at all forbid the case when in a bijective mapping $f : A \to B$ the sets A and B are actually one and the same set.

In other words, a bijective mapping $f : A \to A$ is perfectly legal.

Thus, when we consider a bijective mapping of a finite set onto itself then one way to represent one such bijection visually is to simply reuse all and only the symbols a of the set A for the bottom row of our table-like construct (Figure 4.3.11):



In order to make it perfectly clear, in a visual representation of a bijection of a finite set onto itself shown in Figure 4.3.11 above:

- the number of elements in the bottom row must be equal to the number of elements in the top row
- each and every element from the top row must be placed somewhere in the bottom row
- no other elements in the bottom row can be present

It should now be clear that drawing such elaborate symbols that represent the elements of a set all the time is rather time consuming - we might just as well agree to use the good old small consecutive natural numbers for that specific purpose with nothing to lose and everything to gain (Figure 4.3.12):

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

Some number of discussions later we will take a deep dive study of the objects depicted above - the objects that have a dedicated name in mathematics and that were earlier known as *substitutions* and are now known as *permutations*.

The top row in such a permutation is known as *the domain* of a permutation and the bottom row of such a permutation is known as *the codomain* of a permutation and so on.

The mappings that we looked at briefly in our discussions have many more interesting properties and some of the properties of these mapping that we have stated in an informal way can be proved more rigorously.

The purpose of these discussions was to make our readers aware of the fact that such mappings exist and to expose our readers to the minimum viable material that should be digested before a study of the theory of groups is undertaken.

So far, we have computed the number of possible *in*jections and *bi*jections from one finite set to another finite set.

For completeness, below we compute the number of possible *sur*jections between such sets as an exercise that is marked as *advanced* (with an asterisk) for two reasons:

- while we promised our readers earlier that we will shy away from the raw symbolic manipulations in this space, this will be the time to do so and
- a solution of this exercise requires the familiarity with the so-called *Inclusion Exclusion Counting Principle*

Exercise 4.3.6*: given a finite set A of cardinality n_a:

$$A = \{a_1, a_2, \ldots, a_{n_a}\}$$

and a finite set B of cardinality n_b:

$$B = \{b_1, b_2, \ldots, b_{n_b}\}$$

compute the number of possible *sur*jections from the set A to the set B.

Solution: from the **Definition 5** of a surjective mapping it follows that if the number of elements of the set A is less than the number of elements of the set B, meaning that if $n_a < n_b$, then no surjections from A to B exist and the answer to the question posed is *zero*.

If $n_a \geqslant n_b$ then such surjections do exist and in order to count them all we will use the Inclusion Exclusion Principle by counting the mappings that are *easy to count*.

Namely.

First, we compute the number of *all possible* mappings of A to B.

Since we already did just that in the **Exercise 4.0.3**, we know that the number of all possible mappings from A to B is given by:

$$n_b^{n_a} \tag{5}$$

Next, we want to compute the number of all *not* surjections of A to B and we want to subtract that number form the number of all possible mappings shown above in (5).

This is where we win tactically because counting the *not* surjectons of A to B is, in some sense, *easy* or *straightforward*.

Let us agree to symbolize a set of mappings of A to B that do *not* have the element b_1 of B as their image with $A(b_1^-)$ - such mappings *do not* pair any elements a of A with the element b_1 of B and, by that virtue, do not constitute a surjection.

The number such mappings is also already known to us - it will be the same number shown in (5) with one element, b_1, thrown out of it:

$$(n_b - 1)^{n_a} \tag{6}$$

Likewise, let us symbolize a set of mappings of A to B that do *not* have the element b_2 of B as their image with $A(b_2^-)$ - such mappings *do not* pair any elements a of A with the element b_2 of B and, by that virtue, do not constitute a surjection.

The number of such mappings is also already known to us - it will be the same number shown in (5) with one element, b_2, thrown out of it:

$$(n_b - 1)^{n_a} \tag{7}$$

and so on.

In other words, the number of *non*-surjective mappings $A(b_k^-)$ of A to B that ignore an element b_k of B is the same for all the elements b_k of B.

Thus, in order to solve this problem and answer the question posed, we want to compute the number of elements in *the union* (or the sum) of all the sets of such non-surjective mappings that, one by one, exclude each element b from their images:

$$A(b_1^-) \cup A(b_2^-) \cup \ldots \cup A(b_k^-) \ldots \cup A(b_{n_b}^-)$$

The number of elements in such a union is given by the Inclusion Exclusion Principle and computing the size of each required *intersection* of the above sets of non-surjective mappings is straightforward. That is. We already know that for a fixed element b, say the element b_1, there will be:

$$(n_b - 1)^{n_a}$$

non-surjective mappings. But how many such mappings will there be across *all* the elements b?

Well, in order to answer that question, all we have to do is multiply the number of one-intersections of non-surjective mappings by the number of ways in which 1 item can be picked out of the set of n_b items in an order-insensitive way, which is the good old binomial coefficient n_b *choose one*:

$$\binom{n_b}{1} \cdot (n_b - 1)^{n_a} \tag{8}$$

Thus far, then, for the number of possible surjections sought-after we shall have:

$$n_b^{n_a} - \binom{n_b}{1} \cdot (n_b - 1)^{n_a}$$

But by subtracting the above number from the grand total, we threw out all the even, two-intersections of the non surjective mappings that ignore two elements of B.

For two fixed elements of B, say the elements b_1 and b_2, there will be:

$$(n_b - 2)^{n_a} = |A(b_1^-) \cap A(b_2^-)| \tag{9}$$

possible non-surjective mappings that ignore these *two* elements.

And how many such mappings will there be across *all* the elements of B?
The number of such mappings across all the elements of B will be the number shown in (9) multiplied by the number of ways in which 2 items can be picked out of the set of n_b items in an order-insensitive way, which is the good old binomial coefficient n_b *choose two*:

$$\binom{n_b}{2} \cdot (n_b - 2)^{n_a} \tag{10}$$

Thus far, then, for the number of possible surjections sought-after we shall have:

$$n_b^{n_a} - \binom{n_b}{1} \cdot (n_b - 1)^{n_a} + \binom{n_b}{2} \cdot (n_b - 2)^{n_a}$$

and so on - by adding the number from (10) to the grand total we *over*counted the number of the odd, three-intersections of the non-surjective mappings and the number of such mappings will, clearly, be given by:

$$\binom{n_b}{3} \cdot (n_b - 3)^{n_a}$$

Reasoning further in the same way, we see that if the sizes of the relevant sets agree, that is if $n_a \geqslant n_b$, then the number of possible surjections from A to B will be given by:

$$\sum_{k=0}^{n_b} (-1)^k \binom{n_b}{k} (n_b - k)^{n_a} \tag{11}$$

and because there exist efficient algorithms for the computation of the binomial coefficients shown in (11), such an number can be computed programmatically efficiently as well. □