SETS, RELATIONS, FUNCTIONS, TOPOLOGIES and DEFINITIONS

You're probably familiar with the notion of "function" by now. For many it is a difficult concept to understand. Maybe it's actually an undefined term for you, so far! For, what is a function? "A rule that assigns to each element of a set called the domain an element of another set, called the range." But what, mathematically, is a "rule?" What does "assign" mean, mathematically?

In this chapter we will give a definition of function, using sets, that answers these questions, and does a good job of making the concept of function precise, and, hopefully, clearer, but possibly disappointing. The definition will be made in terms of ordered pairs — we will define a "function" in terms of its graph, a set of points of the form $(x, f(x))$!

Perhaps you will meet topologies here for the first time! A topology is a collection of sets that satisfies three set-theoretic conditions, the "axioms" for a topology. All you've learned so far about "continuity" can be described in an abstract, *short*, set-theoretic way, that can be applied in many, many ways!

We will also develop, as we go, standard ways of working with definitions. So far, new terms have been defined by putting them in bold type, followed by words used to define the new term. We will continue this practice for a few paragraphs.

We start with sets A and B. We construct the set of *all* **ordered pairs** (a, b) of elements a in A, b in B. The set of *all* these ordered pairs is called the **Cartesian product** of A and B, and is denoted $\mathbf{A{\times}B}$. If one of A and B is empty, then $A{\times}B = \varnothing$. The name is in honor of Rene Descartes, the inventor of analytic geometry. There is some terminology to describe ordered pairs. If (a, b) is an ordered pair, then $\mathbf{a}$ is called the **first member of (a, b)**, and $\mathbf{b}$ is called the **second member of (a, b)**. Sometimes **coordinate**, **component**, **element**, or **term** are used in place of "member." By definition, two ordered pairs (a, b) and (c, d) are **equal** if $a = c$ *and* $b = d$. That is, their first members are equal *and* also their second members are equal. All this was done, in terms of sets, in an Exercise in Chapter 2. Hence my choice of "first and second members" of (a, b) to emphasize the connection with sets. But we can also take "ordered pair" to be an undefined term, and regard "ordered pair" as a "primitive" mathematical object.

(4.001) Exercise: Write out $A{\times}B$, given that $A = \{1, 2, 3\}$ and $B = \{2, 3, 4\}$.

Here is a sneaky definition: a **relation** is a set of ordered pairs, that is, a subset R of a Cartesian product! This is an extremely general definition. Any subset R whatever of $A \times B$ is a relation. In particular, $A \times B$, as well as the empty subset of $A \times B$, are relations. So the new term "relation" is not very useful in itself — but it *is* very useful in conjunction with further specifications on the nature of the relation! If the relation's name is R, and (a, b) belongs to R, we also write **aRb** to indicate that a and b are "related." We say that **R is a relation between (the elements of) A and (of) B**. We usually leave out the words in parentheses — but leaving them out can be misleading! The sets A and B need not have any elements in common, nor do they have to belong to the same universal set. However, if A and B *are* the same, so $R \subseteq A \times A$, we say that **R is a relation *on* A**. Here are some examples of relations. Most of these examples are relations usually expressed in the aRb format.

Example: The relation of **equality** is expressed as the subset of $A \times A$, also known as the **diagonal**, of $A \times A$, denoted Δ, and defined, using set selector notation, by the assignment equation $\Delta := \{ (a, a) : a \in A \}$. The name of this relation is usually given by the symbol "=." Thus, we say "a = b" to mean that $(a, b) \in \Delta$, and this means that there is an element c in A such that (a, b) = (c, c). By the way equality of two ordered pairs was defined, this means that a = c and b = c. That is, a and b are the same *element* of A, just bearing different names. Note that this relation involves elements of the same set, so equality is a relation *on* A. Please observe that this discussion does not define the relation "equality!" We already defined equality as "same mathematical object." The discussion only expresses equality in the format of "a relation."

Example: The relation **not equal** is the subset NE of $A \times A$ that is the complement of the diagonal Δ defined in the last example. That is, an ordered pair (a, b) is in NE if and only if a and b are not the same element of A. This relation is usually expressed "a b" instead of "$(a, b) \in$ NE." Inequality is a relation *on* A.

(4.002) Example: Here is an important relation: SI, for set inclusion! Suppose a set X is given. We construct a relation on 2^X, namely a subset of $2^X \times 2^X$, as follows: suppose that A and B are subsets of X. Then we construct SI by requiring that $(A, B) \in$ SI if and only if the mathematical statement $A \subseteq B$ is true.

> **(4.003) Problem**: Formulate a definition of the relation BT, "belongs to," that is a relation between X and 2^X, that describes elements of X belonging to subsets of X.

(4.004) Problem: Formulate a definition of equality for relations. That is, how can we tell when relations R and S are equal?

(4.005) Example: The relation **less than** on the set $A = \{0, 1, 2, 3, 4, 5, 6\}$ is the subset of $A\times A$ given, in list form, by $LT := \{ (0, 1), (0, 2), (0, 3), (0, 4), (0, 5), (0, 6), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 3), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5), (3, 6), (4, 5), (4, 6), (5, 6) \}$.

(4.006) Exercise: Identify this relation on $\mathbf{N}$ in words:
$\{ (0, 0), (1, 1), (2, 4), (3, 9), (4, 16), (5, 25), (6, ?), \ldots \}$.
The relation is a subset of $\mathbf{N}\times\mathbf{N}$.

(4.007) If R is a relation, $R \subseteq X\times Y$, its **inverse relation**, $\mathbf{R^{-1}}$, is the subset of $Y\times X$ given by $R^{-1} := \{ (y, x) \in Y\times X : (x, y) \in R \}$.

Thus the inverse of a relation is made by just interchanging the members of each ordered pair in R!

(4.008) Exercise: Consider the relation LT given in (4.005). What is LT^{-1}?

☛ **kinds of relations**

The most important kinds of relations for us are orderings, functions, equivalence relations, “belongs to” and set inclusion. We begin with some simple “kinds” of relation. Then we’ll study the concept of “function” carefully. We will then digress to discuss a basic subject used with functions, called “topology,” and define “continuous function,” using topologies. Then we will study equivalence relations. We have already discussed “belongs to” and set inclusion. The discussion of orderings will begin, in earnest, in the next chapter. Along they way, there will be digressions about Definitions.

Here are some conditions we can put on a relation. If a condition “X” is true about a relation R, then we say relation R “has the X **property**,” or “is a “X” relation.”

A relation R is called
(4.009) reflexive if “$R \subseteq A\times A$ and (for all a in A)($(a, a) \in R$)” is true.

That is, for all a in A, a is related to a. Also, this says “$\Delta \subseteq R$.”

(4.010) symmetric if "$R \subseteq A \times A$ and (for all $(a, b) \in R$)($(a, b) \in R \Rightarrow (b, a) \in R$)" is true.

That is, "a R b implies b R a" is true for all a, b.

(4.011) transitive if the following statement is true about R: "$R \subseteq A \times A$ and,

$$(\forall a \in A)(\forall b \in A)(\forall c \in A)\ (\ ((a, b) \in R \text{ and } (b, c) \in R) \Rightarrow ((a, c) \in R)\ ).\text{"}$$

That is, a R b and b R c implies a R c. An example of this would be the relation "less than."

We'll use these three properties of a relation later, to define "equivalence relation."

(4.012) Exercise: Which of the properties **reflexivity**, **symmetry**, **transitivity** does the relation "equality" have?

(4.013) Exercise: Which of the properties reflexivity, symmetry, transitivity does the relation "less than" have?

(4.014) Exercise: Which of the properties reflexivity, symmetry, transitivity does the relation "less than or equal to" have?

(4.015) Exercise: Show that set inclusion (see (4.002)) is a transitive relation.

(4.016) Exercise: How would you define "**antisymmetry**," as a property of a relation?

(4.017) Exercise: Here is a mathematical statement that is logically equivalent to one of the definitions just given. Determine which one it is, and say why.

$$(\forall a \in A)(\forall b \in A)(\ ((a, b) \in R) \Rightarrow ((b, a) \in R)\ )$$

The point of the last exercise is that definitions can be given in different ways. You are requested to make up more understandable versions of definitions, and bring them to my attention.

(4.018) Exercise: Make up a mathematical statement that is logically equivalent to the definition of transitivity, but uses quantifiers differently.

☛ introduction to the overwhelming importance of definitions

We are soon going to make a definition (of function). We will start listing almost all definitions in a more formal manner than before; they will have a heading of their own: ***Definition:*** By the way, notation is important in mathematics! We will treat notation like definition. In fact, notations will be defined like terms; it's just that notation defines a symbol, rather than a term, so notation, from now on, will get a heading too, like that for definition: ***Notation:*** As usual, the term, or the notation being defined will appear in bold type the first time. Related terminology will also appear in bold type. Sometimes terminology will be introduced in the text, simply by placing it in bold type while it is being explained.

FUNCTIONS

We are going to define "function" as a kind of relation! That is, we are going to list some conditions that a relation might, or might not, satisfy, and declare that a relation is a function if it satisfies these conditions. You probably already have the ideas behind these conditions !
A common "definition" of "function" is: a **rule** that **assigns** to every element, x, of the domain, X, a **definite** element y of the range, Y. We will use this method to define functions! But there are problems with this "definition!" What does "rule" mean? What does "assign" mean? What does "definite" mean? Obviously, these terms do mean something. The problem is, what? I claim that the "official" definition — to be given shortly — makes these terms meaningful.
We are now ready to define "function" as a special kind of relation — a graph. The definition is a bit long, because it needs to include terms such as "domain" and "range." In fact the definition defines three terms, and some synonyms! The definition we will use may be controversial! This is because I will insist that non-empty sets be used in the definition! Allowing empty domain or range produces exactly one "graph," the empty set, and constant hassles in dealing with the empty case, whether directly, or implicitly, by careful phrasing designed to work "vacuously." We don't need that sophistication right now. You can play with it on your own, or deal with it later!

☛ at last, the definition of "function"

(4.019) ***Definition:*** Let X and Y be non-empty sets. A **function (mapping, transformation, operator, etc.)**, F, with **domain** X and **range** Y is a subset, F, of $X \times Y$ that satisfies the following conditions:

> (1) For all x in the domain, X, of F, there exists y in the range, Y, of F such that the ordered pair (x,y) belongs to F.
>
> (2) For all ordered pairs (x, y) in F, and for all ordered pairs (u, v) in F, if x = u then y = v.

The official definition is not very clear, so here are some explanatory remarks.

Let's see whether the official definition works like the "usual definition: a **rule** that **assigns** to *every* element, x, of the domain, X, a **definite** element y of the range, Y. " Here is what "assigns" means, in the "official" definition.

For every x in X, there is, by (1), an ordered pair (x, y) in F that has x as its first member. According to (2), (x, y) is the *only* ordered pair in F that has x as its first member. Let y be the element "assigned" to x. We nearly always refer to this assigned y as F(x).

> Notice, please, that the same y can appear as the *second* member of different ordered pairs in F. What (2) says is that we can't have two different ordered pairs in F that have the same *first* member. Condition (2) is sometimes called the "vertical-line test," because a set G contained in the plane (the plane viewed as the set of ordered pairs (x, y) of real numbers) is the graph of a function with domain S (a subset of **R**) if and only if every vertical line thru a point (on the real axis, that is) in S contains exactly one point (x, y) in G.

Here is what "definite" means:

Is this y definite? Well, according to (2), if (x, y) is in F, and also (x, v) is in F (this is what $(u, v) \in F$ and $u = x$ means), then $v = y$. That is, for each $x \in X$, there is exactly one y such that $(x, y) \in F$.

Here is what "rule" means:

The vagueness of what "rule" means is replaced by this unsatisfying answer: **the set F of ordered pairs IS the "rule!"** This is unsatisfying because we are used to defining functions with formulas. Then we find people defining functions using two or more formulas, depending on where the x in the function is located, and so on. The point is, there are too many "rules," too many kinds of "rules" to classify. Our expectation that "rule" can be understood simply is doomed to disappointment, so we give it up.

(4.020) ***Notation:*** There are several notations used for functions. The most common is **F(x)**, and it usually means, in Calculus, the numerical value of a formula, e.g. $F(x) = \dfrac{1}{1 + x^2}$. For us, from now on, F(x) will *officially* mean the y that is the second member of the ordered

pair (x, y) in F. Thus, the ordered pairs in F can all be expressed as $(x, F(x))$. The function itself is simply denoted **F**. But this "officialese" does not always work well, so we will often use $F(x)$ to stand for the function itself, rather than a single value of it. This is especially useful when the function is defined by a formula, such as x^2. Another common notation is $\mathbf{F_x}$, so that ordered pairs in F are written (x, F_x). This notation is almost always used when the domain of the function is **N**. That is, a **sequence** is a function with domain **N**, and a sequence F is often denoted $\{F_n\}$. There can be some confusion when subscripted variables denote partial derivatives, so F_x is seldom used when F is a function of one or several real variables. In case F is a **linear function**, $\mathbf{Fx}$ is used instead of $F(x)$ to connote a matrix multiplied times a vector, even when the linear function is not multiplication by a matrix. For example, $\frac{d}{dx} f$ is often used to denote the linear operation of differentiation. Here, the domain of the function $\frac{d}{dx}$ is "differentiable functions," and the range is "functions." Sometimes we want to emphasize that the function has domain X and range Y, and then we use the notation $\mathbf{f : X \to Y}$. This is read "f maps X to Y," or "f is a function from X to Y," or "f takes X to Y," and so on. There is notation similar to $f: X \to Y$ that is useful sometimes for drawing diagrams that illustrate the effects of combining functions:
$\mathbf{X \xrightarrow{f} Y}$. We will use this when we deal with composite functions, to be defined shortly.

> **(4.021) Exercise**: Formulate a definition of equality for functions. That is, how can we tell when functions F and G are equal?

(4.022) ***Definitions and terminology:*** Let F be a function with domain X and range Y. (Note that this implicitly says that X and Y are non-empty — another reason to memorize definitions!) Then in the equation $y = F(x)$, that identifies y as the second member of the ordered pair (x, y) in F, x is called an **argument** of F, and y a **value**. We say a function F ***maps*** **X into Y**, ***transforms*** **x (in X) to F(x) (in Y)**, ***sends*** **x (in X) to F(x) (in Y)**, and so on. When F has the same domain and range, that is, when $X = Y$, we say **F is a function *on* X**.

(4.023) ***Definition and notation:*** If F is a function, with domain X and range Y, then the **image of** (a subset) **E** (of X) **under F** is the set of all y in Y such that $y = F(x)$ for some x in E. The image of E under F is denoted **F(E)**. In set selector notation,

$$F(E) := \{\, y \in Y : (\exists\, x \in X)(\, x \in E \text{ and } y = F(x)\,)\,\}.$$

When $E = X$, $F(X)$ is called the **image of F**. It consists of all the y in Y that are values of x's in X. Sometimes when people use the term "range" they mean "image" in the sense of this definition.

Example: Let $F = \{ (x, y) \in \mathbf{R}^2 : y = x^2 \}$. For each x in $\mathbf{R}$, $F(x) = x^2$. The image of F is $[0, \ )$, the image of $[-1, 1]$ is $[0, 1]$. Note that we have not officially defined intervals, or $\mathbf{R}$; they're just being used for an example. In this example it is natural to confuse the name of the function with the formula defining it. We usually say $F(x) = x^2$, thinking of x as a variable quantity, not just a particular but unspecified number. So the name of the function is usually x^2, not "squaring function," though "squaring function" is a fine name for the function x^2.

Example: Let $F = \{ (m, n) \in \mathbf{N}^2 : n = m/2,$ if m is even, and $n = 2m$ if m is odd $\}$. The image of F is $\mathbf{N}$. The image of the set of odd integers is the set of even positive integers that have remainder 2 when divided by 4:
$F(\{ n \in \mathbf{N} :\ n$ is odd $\}) = \{ n \in \mathbf{N} :\ n$ is even and $n > 0$ and $n - 2$ is divisible by $4 \}$.
Let $E = \{1, 2, 3\}$. Then $F(E) = \{1, 2, 6\}$. Notice that the image need not be written so that the values of F match the arguments that the values were assigned to.

(4.024) Problem: Prove that each of the following relations is true, for an *arbitrary* function, and arbitrary sets A and B in the domain of F.
(a) $F(A \cap B) \subseteq F(A) \cap F(B)$,
(b) $F(A \cup B) = F(A) \cup F(B)$.

(4.025) Problem: Find *examples* of functions F and sets A and B in the domain of F, such that the following *inequations* are true. You may use different functions in different parts.
(a) $F(A \cap B) \quad F(A) \cap F(B)$,
(b) $F(A^c) \quad (F(A))^c$.

To find out whether some given y is, or is not, in the image of F, we have to be able to solve the equation $F(x) = y$, where x is the unknown. To find out whether some given y is, or is not, in $F(E)$, the image of E under F, we have to be able to solve the equation $F(x) = y$, where x is the unknown, and show that there is such an x in E. Thus, in the squaring example, 25 is in the image of F. But 25 is not in the image of $[-1, 1]$.

(4.026) ***Definition and notation:*** Let F be a function with domain X and range Y. Let S be a subset of Y. Then the **inverse image of S under F** is the set of all x in X such that $F(x) \in S$. The inverse image of S under F is denoted $\mathbf{F^{-1}(S)}$. In set selector notation,

$$F^{-1}(S) := \{ x \in X :\ F(x) \in S \}.$$

Example: Let F be the squaring function. It has domain **R** and range **R**. Let S = {25}. Then $F^{-1}(S) = \{-5, 5\}$. If $S = [-1, 1]$, $F^{-1}(S) = [-1, 1]$. If $S = [0, 1]$, we still have $F^{-1}(S) = [-1, 1]$. Notice that, in contrast to images, $F^{-1}(S)$ can be empty for some non-empty S. For example, if $S = \{-1\}$, then $F^{-1}(S) = \{ x \in \mathbf{R} : F(x) \in \{-1\} \} = \{ x \in \mathbf{R} : x^2 = -1 \} = \varnothing$.

(4.027) Exercise: Let F be a function with domain X and range Y. Show that, for all subsets A and B of Y,
(a) $F^{-1}(A \cap B) = F^{-1}(A) \cap F^{-1}(B)$,
(b) $F^{-1}(A \cup B) = F^{-1}(A) \cup F^{-1}(B)$,
(c) $F^{-1}(A^c) = (F^{-1}(A))^c$,
(d) $F^{-1}(A \Delta B) = F^{-1}(A) \Delta F^{-1}(B)$.
Please notice how this Exercise contrasts with Exercises (4.024) and (4.025)! Because of this exercise, it's "nicer" to work with inverse images than with images.

(4.028) Problem: Let F be a function with domain X and range Y. Show that the definition of F^{-1} constructs a function from 2^Y into 2^X.

Just as there are kinds of relations (and "function" is officially a kind of relation), there are

☛ kinds of functions

(4.029) ***Definition:*** A function F is **one-to-one (one-one, injective, 1-1)** if, for all u and v that are in the domain of F, $F(u) = F(v)$ implies $u = v$. That is, if the ordered pairs (u, y) and (v, y) are in F, and F is one-to-one, then u = v.

Please compare this to condition (2) in the definition of "function:" (2): For all ordered pairs (x, y) in F, and for all ordered pairs (u, v) in F, if x = u then y = v. This was called the "vertical-line test." In this new definition, of one-to-one function, the condition gets turned around; we might write a condition (3): For all ordered pairs (x, y) in F, and for all ordered pairs (u, v) in F, if y = v then x = u. This we can call the "horizontal-line test." But there is a subtle difference! For this test says that every horizontal line that passes thru the *image* of F contains exactly one point of F. It does NOT say that every horizontal line that passes thru the *range* of F contains exactly one point of F! It does say that every horizontal line that passes thru the range of F contains *at most* one point of F.

Example: The function F, with domain **N** and range **N**, given by $F(n) = n^2$, is one-one, because $F(n) = F(m)$ means $n^2 = m^2$, so $n^2 - m^2 = (n - m)(n + m) = 0$. One of the facts about integers we take as known is that when the product of two integers is 0, at least one of the two integers is 0. Thus, from this fact and the equation $(n - m)(n + m) = 0$ we deduce that $n - m = 0$ or $n + m = 0$. If $n - m = 0$, then we know that $n = m$, so we are done if that is true. What if $n + m = 0$? Then, since n and m are non-negative, the only way their sum could be 0 is for each of them to be 0. But then, they were equal to start with, so the second part of the proof turns out to be redundant — but it "had to be done," from the point of view of proof.

A concept sort of obverse to one-one is **many-to-one**. A function F is many-to-one if it is not one-to-one. This means that there must exist at least one y in Y (and one is enough) such that there are two *different* ordered pairs (u, y), (v, y) in F. That is, $F(u) = y = F(v)$, with u v. An example: $X := [-1, 1]$, $Y := \mathbf{R}$, $F(x) := x^2$. Both $(-1, 1)$ and $(1, 1)$ are in F for $y = 1$.

(4.030) ***Definition:*** A function F with domain X and range Y is **onto (surjective)** if, for all y in Y there exists x in X such that $y = F(x)$. That is, for all $y \in Y$ there exists in F an ordered pair (x, y) that has y as its second member.

Example: Let $F = \{ (m, n) \in \mathbf{N}^2 : n = m/2, \text{ if } m \text{ is even, and } n = 2m \text{ if } m \text{ is odd} \}$. We saw this example before. Notice that $F(\{n \in \mathbf{N} : n \text{ is even} \}) = \mathbf{N}$. Therefore F is onto.
In particular, for all functions $f : X \to Y$, f is onto if and only if the image of f is Y.

If we want to know, for a given y, for just what x it is true that $F(x) = y$, then we have to solve the equation $F(x) = y$ with x as the unknown. A fancy way of saying F is onto is: for all y in Y, $F^{-1}(\{y\})$ $\varnothing$.

(4.031) Exercise: Use the F in the preceding example, and find all solutions, x, of $F(x) = 14$. Find all solutions of $F(x) = 3$.

(4.032) ***Definition:*** A function F with domain X and range Y is **invertible (has an inverse, is bijective)** if F is both one-to-one and onto. In the special case that F has domain X and range X, an invertible function is sometimes called a **permutation**.

Now we can state and prove a theorem! It will be an easy proof to follow, but not so easy to think how to do, until you have seen some examples. Now, we have proved some theorems already, but without ballyhoo. We will be using the following format for theorems pretty much from now on. It makes them easy to find, and separates their proofs from other stuff. This proof will be a lot longer than it needs to be, because it contains a lot of careful explanation.

(4.033) Theorem: A function F with domain X and range Y is invertible if and only if there exists a function G with domain Y and range X such that

(1) For all x in X, G(F(x)) = x,

and

(2) For all y in Y, F(G(y)) = y.

What the theorem "says" is that the statement "F is invertible," (which really means "F is one-to-one and onto," which really means what is in the definitions of "one-to-one" and "onto," and makes use of the definition of "function") is logically equivalent to the compound statement made of the two statements labeled (1) and (2) in the theorem. After stating a theorem, we need to present an "argument" — reasoning and explanation — designed to convince the reader that the theorem is indeed true. This argument is called a "proof." We will signal the start of a proof with "*Proof:* ." The end of a proof will either be obvious, or will be indicated with phrases such as "as desired," "completes the proof," and so on. The following proof will be much longer than necessary, because it contains a lot of explanation of why what is being done is being done!

Proof: The theorem has the logical form $A \Leftrightarrow B$. We must show two statements are true: (I): $A \Rightarrow B$, and (II): $B \Rightarrow A$. Sometimes people write (II) as $A \Leftarrow B$. Here are (I) and (II).

(I): "F is invertible" $\Rightarrow$ "There exists a function G with domain Y and range X such that

(1) For all x in X, G(F(x)) = x,

and

(2) For all y in Y, F(G(y)) = y."

(II): "There exists a function G with domain Y and range X such that

(1) For all x in X, G(F(x)) = x,

and

(2) For all y in Y, F(G(y)) = y." $\Rightarrow$ "F is invertible."

Let us begin with (I). By definition, "F is invertible" means that F is one-to-one and F is onto. This means that the statements in the definitions of function, of one-to-one function, and of onto function are all given to be true. Here they are:

(a) For all x in the domain, X, of F, there exists y in the range, Y, of F such that the ordered pair (x,y) belongs to F (Every x in X is the first member of an ordered pair in F).

(b) For all ordered pairs (x, y) in F, and for all ordered pairs (u, v) in F, if x = u then y = v (F passes the vertical-line test).

(c) For all u and v in X, F(u) = F(v) implies u = v (F is one-to-one).

(d) For all y in Y there exists x in X such that y = F(x) (F is onto).

Let us *construct* a relation G, as a subset of Y×X, as follows:

$$G := \{ (y, x) \in Y \times X : (x, y) \in F \}.$$

recall that $G = F^{-1}$, the inverse relation, is not the inverse function, yet. The construction amounts to "reversing" all the ordered pairs in F. We have to *show* that G is a function.

I assert that G is a function, and that the statements

(1) For all x in X, G(F(x)) = x,

and

(2) For all y in Y, F(G(y)) = y

are true mathematical statements about G. Thus, there are three things to do: "prove" that G is a function, and "prove" that (1) and (2) are true.

To prove that G is a function, I need to show that these mathematical statements are true:

(a*) For all y in the domain, Y, of G, there exists x in the range, X, of G such that the ordered pair (y, x) belongs to G.

(b*) For all ordered pairs (y, x) in G, and for all ordered pairs (v, u) in G, if y = v then x = u.

Let us show that (a*) is true.

We are given that statement (d), above, is true; namely, for all y in Y there exists x in X such that $y = F(x)$.
In other words, since F is onto, for all y in Y, there exists x in X such that $(x, y) \in F$.
By definition of G, (y, x) is in G.
Therefore, (a*) *is* true, because we chose an *arbitrary* y in Y, and, for that y, found that an x in X exists such that the ordered pair (y, x) is in G.

> The process used here was: **deduction** from true mathematical statements. Let y_o be given in Y. Since (d) is true, (3.3) Rule 5 allows us to deduce that "there exists x in X such that $y_o = F(x)$" is true. By (3.3) Rule 6, we can select an x_o in X such that "$y_o = F(x_o)$" is true. Consider such an x_o as selected now. Thus, for the given y_o, we have "found" x_o in X such that $y_o = F(x_o)$. This means that (x_o, y_o) belongs to F, so that, by construction, (y_o, x_o) belongs to G. We have shown that, for an arbitrary y in Y, there exists x in X (*not* arbitrary, because it depends on y) such that (y, x) is in G. But then the "universal and" statement (a*) is true. This was the goal, so now we can go on to the next statement.

Let us show that (b*) is true. Suppose that (y, x) and (v, u) are in G. Suppose that $y = v$. This means that (x, y) and (u, v) are in F, by our construction of G. Our notation for this is also that $F(x) = y$ and $F(u) = v$. We are given that statement (3), above, is true; namely, for all u and v in X, $F(u) = F(v)$ implies $u = v$. Recall that we have to substitute the letters we are working with for the letters used in the statements in definitions. So we choose $u = u$ and $v = x$. Then we are given that $F(x) = y$ and $F(u) = v$, and that $y = v$. This means that $F(x) = F(u)$, so from (3) we deduce that "$x = u$" is true. Thus, (b*) is true.

> **(4.034) Exercise**: Justify the steps in the argument proving that (b*) is true, using the Rules for deduction in (3.3).

Now we know that G is a function, because the two defining conditions, listed here as (a*) and (b*), that something be called a function, are true for G.

We next need to show that (1) and (2) are true.

> If you don't remember what (1) and (2) say, you need to look at them again, now. This sort of referencing happens a lot in proofs. We have a tendency not to look back, that can mean trouble!

To show that (1) is true, let x be an arbitrary element of X. Then $G(F(x))$ is the second member of the ordered pair $(F(x), G(F(x)))$ in G that has $F(x)$ as its first member. Also, $u := G(F(x))$ is an element u in X and $(F(x), u) \in G$, so $(u, F(x)) \in F$, by the way we constructed G. Since $(u, F(x)) \in F$, we know that $F(u) = F(x)$. Since F is one-to-one, there is only one element u in X such that $F(u) = F(x)$, namely x. Therefore, the ordered pair $(F(x), G(F(x)))$ in G must be $(F(x), x)$. Thus, $G(F(x)) = x$, by the way equality of ordered pairs was defined.

> **(4.035) Exercise**: Justify the steps in the argument proving that (1) is true, using the Rules for deduction in (3.3).

To show that (2) is true, we start with an arbitrary y in Y.
Since F is onto, we know there is an x in X such that $F(x) = y$.
That is, $(x, y) = (x, F(x))$ belongs to F.
Therefore, $(F(x), x) = (y, x) \in G$, so $G(y) = x$.
Then $F(G(y)) = F(x) = y$.
This is what had to be shown. Thus, (2) is true, and (I) is, at last, proved!

> **(4.036) Problem**: Prove that (II) is true. You might be able to start at the end of the argument for (I) and work back!

(4.037) ***Definition:*** If F is an invertible function, the function G, whose existence is proved in Theorem (4.033), is called the **inverse function of F**, or the **function inverse to F**.

(4.038) ***Notation:*** If F is an invertible function, the inverse function of F is denoted $\mathbf{F^{-1}}$.

> **(4.039) Problem**: Show that, if F is an invertible function, so is F^{-1}, and $(F^{-1})^{-1} = F$.

> **(4.040) Problem**: Show that F is an invertible function if and only if the inverse of the relation F, (also denoted F^{-1}), is a function.

ONE-TO-ONE CORRESPONDENCES

(4.041) ***Definition:*** Two sets X and Y **can be put into one-to-one correspondence**, or **are in one-to-one correspondence**, if there exists a function $f : X \to Y$ that is one-to-one and onto. The function f is then called a **one-to-one correspondence**.

There is nothing new here but the name, a nice name because it gives us the sense of matching each thing in one of the sets with exactly one thing in the other set. A one-to-one correspondence, f, as in the definition, has an inverse function, so if X and Y can be put into one-to-one correspondence, so can Y and X. Note that “are in one-to-one correspondence” is not a relation on “sets,” because of the technical fact that **there is no such thing as the set of all sets**! However, “are in one-to-one correspondence” IS an equivalence relation on 2^X, if X is not empty. Further discussion of this would be part of a course in set theory.

OPERATIONS INVOLVING FUNCTIONS

We have just seen an operation involving functions F that are invertible, namely, the construction of the inverse function, F^{-1}. In the case of functions $F : \mathbf{R} \to \mathbf{R}$, this has the geometric interpretation of “reflection in the diagonal.” You’ll see it in the next Exercise.

(4.042) Exercise: Draw the graph of $F(x) = x^2$, for x in the domain $0 \quad x \quad 1$, on a sheet of paper, using crayons, a marker, etc., that gives you bold lines. Label the x and y axes. Also draw in, not so heavily, the line $y = x$. Now, hold the paper with one pair of fingers at each end of the $y = x$ line, and rotate the page about the line so you are looking at the back side of the sheet. Hold it up to the light, so you are looking thru the paper. What you see is the graph of F^{-1}. What is the common name for this inverse function?

(4.043) Exercise: Draw the graph of $F(x) = x^2$, for x in the domain $-1 \quad x \quad 1$, on a sheet of paper, using crayons, a marker, etc., that gives you bold lines. Label the x and y axes. Also draw in, not so heavily, the line $y = x$. Now, hold the paper with one pair of fingers at each end of the $y = x$ line, and rotate the page about the line so you are looking at the back side of the sheet. Hold it up to the light, so you are looking thru the paper. What you see is the inverse relation F^{-1}, and you see why this relation is not a function — it violates the vertical-line test..

There is a basic operation involving two functions: constructing the **composite function**.

(4.044) ***Definition and notation:*** Let f be a function with domain X and range Y. Let g be a function with domain Y and range Z. The **composition of f and g** is the **composite function $g \circ f$**, with domain X and range Z, whose ordered pairs are $(x, g(f(x)))$. This is given in set selector notation by $g \circ f := \{ (x, z) \in X \times Z : z = g(f(x)) \}$.

We can draw some diagrams that illustrate functional composition:

$$\mathbf{X} \xrightarrow{\mathbf{f}} \mathbf{Y} \xrightarrow{\mathbf{g}} \mathbf{Z} \text{ combined give } \mathbf{X} \xrightarrow{\mathbf{g}\circ\mathbf{f}} \mathbf{Z}.$$

Note the peculiar thing that happens! The order of f and g get reversed!

Example: Let f be the squaring function, $f(x) = x^2$, and g the "add 1" function given by $g(y) = y + 1$. Then $g\circ f$ is the function whose value, $g\circ f(x)$, at x is found by first applying f to x, getting $y = x^2$, then by applying g to y, getting $y + 1 = x^2 + 1$. This example has been set up to give a double example, because the range and domain are the same. So we can look at $f\circ g$. The value of this function at x is found by forming $y = g(x)$, namely $x+1$, then by forming $f(y) = y^2$, namely $y^2 = (x + 1)^2 = x^2 + 2x + 1 = f\circ g(x) \quad g\circ f(x) = x^2 + 1$, unless $x = 0$.

This example shows that composition of functions is NOT an always-commutative operation, even when it *can* be carried out, for $f\circ g \quad g\circ f$ in the example.

(4.045) Theorem (Associativity of function composition): Let $f : X \to Y$, $g : Y \to Z$, and $h : Z \to A$. Then $h\circ(g\circ f) = (h\circ g)\circ f$.

Proof: For all x in X, $(h\circ(g\circ f))(x) = h((g\circ f)(x)) = h(g(f(x)))$, and $((h\circ g)\circ f)(x) = (h\circ g)(f(x)) = h(g(f(x)))$. By transitivity of equality, $(h\circ(g\circ f))(x) = ((h\circ g)\circ f)(x)$. Thus the two functions have the same ordered pairs, so they are equal. The theorem is proved.

The operations of function composition and function inversion are examples of operations that cannot always be done. But composition is always associative, when do-able.

(4.046) Exercise: Show that the composition of two one-to-one mappings is a one-to-one mapping.

(4.047) Exercise: Show that the composition of two onto mappings is an onto mapping.

(4.048) Exercise: Show that the composition of two one-to-one correspondences is a one-to-one correspondence.

(4.049) Problem: Show that the inverse of the composition of two one-to-one correspondences is the composition of their inverses in the reverse order, and express the result in an equation.

We now can return briefly to a topic in set theory that uses the idea of a function, in passing.

De Morgan's laws for set operations

We will sometimes have to work with sets of sets. When we do, we often use synonyms for "set" to avoid confusion. A set of sets might be called a "collection of sets," or a "family of sets," and the words "collection" and "family" do not have any special connotation, such as, that the sets somehow belong together, or are somehow related. They might be, they might not be. When we work with a family of sets, we may want to identify the sets in a uniform way. This is often done with subscripts, sometimes with superscripts. If there are, say, 5 sets in the family, we might denote the sets S_1, S_2, S_3, S_4, S_5. The subscripts are called **indices** collectively; each subscript is called an **index**. The indices comprise a set called the **index set**. When *we* work with families of sets, each set in the family will be a subset of some universal set for the family. Usually, but not always, different indices correspond to different sets in the family. Suppose we have a family of subsets of a set X, indexed by subscripts that belong to an index set A. Let's use S_α to stand for the set in the family that corresponds to the index α. **This is an instance of a function that maps A into 2^X.**

We don't use "class" to refer to a set or to a set of sets, because "class" is a technical term in Set Theory. But we could if you want.

(4.050) ***Definition and notation:*** Let Φ be a family of subsets S_α of a set X, indexed by an index set A. The **union over A of the sets S_α** is the set of all x in X that belong to at least one of the sets in Φ. In set selector notation, this new set, denoted $\bigcup_{\alpha\in A} S_\alpha$, or $\bigcup_{S\in\Phi} S$, is given by $\bigcup_{\alpha\in A} S_\alpha = \bigcup_{S\in\Phi} S = \{ x \in X : (\exists\, \alpha \in A)(x \in S_\alpha) \}$.

(4.051) Exercise: Show that, in the context of (4.050), $\{ x \in X : (\exists\, \alpha \in A)(x \in S_\alpha) \} = \{ x \in X : (\exists\, S \in \Phi)(x \in S) \}$.

(4.052) ***Definition and notation:*** Let Φ be a family of subsets S_α of a set X, indexed by an index set A. The **intersection over A of the sets S_α** is the set of all x in X that belong to every one of the sets in Φ. In set selector notation, this new set, denoted $\bigcap_{\alpha\in A} S_\alpha$, or $\bigcap_{S\in\Phi} S$, is given by $\bigcap_{\alpha\in A} S_\alpha = \bigcap_{S\in\Phi} S = \{ x \in X : (\forall\, \alpha \in A)(x \in S_\alpha) \}$.

Now we can state De Morgan's Laws, as a theorem, a theorem-with-a-name. The name makes the theorem easier to remember and to talk about. The theorem is going to "say" this: "the complement of a union is the intersection of the complements," and "the complement of an intersection is the union of the complements." In the theorem, we'll use the notation A^c for the complement of a set A that is a subset of a set X, so that $A^c = X \sim A$.

This will be one of our first FORMAL **Theorem**s. There will be many more. A theorem is a mathematical statement that is supposed to be a "big deal," so it gets a special heading. Each Theorem requires a proof. This will start with *Proof:* in this text. Each Theorem needs to be remembered, not memorized. There are as many ways as there are people to remember theorems. Your way needs to include understanding what the Theorem "says," and understanding how the proof proves the Theorem. You may also want to remember the ideas that lie behind the *Proof:*, and you may have to ask what the idea is. The idea is very often not apparent!

(4.053) Theorem (De Morgan's Laws): Let Φ be a family of subsets S_α of a set X, indexed by an index set A. Then

$$\left(\bigcup_{\alpha\in A} S_\alpha\right)^c = \bigcap_{\alpha\in A} S_\alpha^c, \text{ and } \left(\bigcap_{\alpha\in A} S_\alpha\right)^c = \bigcup_{\alpha\in A} S_\alpha^c .$$

Proof: We use the standard way of showing two sets are equal: show that each is contained in the other. Suppose $x \in \left(\bigcup_{\alpha\in A} S_\alpha\right)^c$. Then x does not belong to any of the sets S_α. That is, for all $\alpha \in A$, $x \notin S_\alpha$. Since $x \in X$, this means $x \in S_\alpha^c$, for all $\alpha \in A$. This is true for all $\alpha \in A$, so $x \in \bigcap_{\alpha\in A} S_\alpha^c$.

We have shown that $\left(\bigcup_{\alpha\in A} S_\alpha\right)^c \subseteq \bigcap_{\alpha\in A} S_\alpha^c$, and it remains to show that the reverse inclusion is true. Suppose $x \in \bigcap_{\alpha\in A} S_\alpha^c$. That is, x belongs to every one of the sets S_α^c. Thus, x belongs to none of the sets S_α. Therefore, x does not belong to any of the sets S_α, so x does not belong to $\bigcup_{\alpha\in A} S_\alpha$. Since x does belong to X, x belongs to the complement of $\bigcup_{\alpha\in A} S_\alpha$, namely to $\left(\bigcup_{\alpha\in A} S_\alpha\right)^c$. This completes the proof of the first statement, and the proof of the second one is similar.

(4.054) Exercise: Make the changes necessary in the preceding proof of the first of De Morgan's Laws that make it into a proof of the second of De Morgan's Laws.

We can now make the second version of the fundamental assumption that allows us to construct universal sets at will.

A fundamental assumption, second version: ***Let Φ be a family of sets S that are not contained in the same universal set. Then a set X exists that contains all the elements in the sets S, and no others. This set is called the union of Φ, and is denoted $\bigcup_{S \in \Phi} S$.***

☛ an extremely important general definition: a topology

A very important example of a family of sets is a "topology," a family of subsets of a set X that satisfies three special conditions invented to "abstract" the idea of "continuity.".

One application of De Morgan's Laws occurs after the definition of "closed set" in terms of "open set." You will be asked to prove that closed sets satisfy some conditions "dual" (via complementation) to those satisfied by open sets.

The word "topology" has a very different definition in a dictionary than the one we'll use here! Please look up the definition in your favorite dictionary, and see whether you agree! My dictionary says "topos" is Greek for "place." In keeping with the idea of sets being "places," we often call a set a **space**, and call its elements **points**.

(4.055) ***Definition:*** Let X be an arbitrary set. We say that a family $\boldsymbol{T}$ of subsets of X is **a family of open sets for X**, or **a topology for X**, or **a topology on X**, if

(i) The empty subset of X, and X, both belong to $\boldsymbol{T}$;
(ii) The intersection of two sets in T is also in $\boldsymbol{T}$;
(iii) The union of an arbitrary collection of sets in $\boldsymbol{T}$ is also in $\boldsymbol{T}$.

When $\boldsymbol{T}$ is a topology for a set X, X is said to **have** the topology $\boldsymbol{T}$. The set X, with a topology $\boldsymbol{T}$, is called **a topological space**.

Note the use of "*a* family of open sets," rather than "*the* family of open sets!" There can be many, many different topologies on a set. Also, notice that $\boldsymbol{T} \subseteq 2^{X}$. What set does a topology on X *belong* to?

☛ examples of topologies

The **indiscrete topology**, on any set X, is $\{\emptyset, X\}$.

The **discrete topology** on any set X is 2^X. That is, every subset of X is an open set in the discrete topology.

The indiscrete topology is the smallest topology that a set X can have (why?).
The discrete topology is the largest topology that a set X can have (why?).

> **(4.056) Exercise**: Show that the “indiscrete topology” and the “discrete topology” are indeed topologies. In other words, show that those collections of sets satisfy all three conditions in Definition (4.055). Just because we *call* a family of sets a topology doesn’t make it one! The thing that *makes* a set of subsets of X a topology on X is *our* verification that all three conditions in Definition (4.055) are true!

The **usual topology** for the Plane P of Euclidean Geometry consists of all subsets of P that contain all the points in a sufficiently small circular disc about each of their points. In other words, a subset U of P is **open** if for all $p \in U$, it is true that all the points inside some small circle with center at p are also points of U. In still other words, U is open if, for all p in U, there exists a point r p such that all points q, that are closer to p than p is to r, are also in U.

An open set in the Euclidean Plane
In the figure, the gray area represents an open set U. The dot at the center of the circle is p. The dot ON the circle is r. The dot inside the circle, but not at p, is a q. The dot on the edge of U is not IN U, because every circle with center at U and positive radius has to contain points that are not in U, as well as points that *are* in U.

We will use the usual topology on P as our main example of a topology for a while. The plane P is “unofficial;” we will not use the axioms of Euclidean Geometry in this course! But we *will* use Euclidean Geometry “in the back room” where we get and keep ideas!

✒ definition of “closed set”

(4.057) ***Definition:*** Let S be an arbitrary subset of a topological space X. We say that **S is an closed set** if S is the complement of an open set, in other words, S is closed if S^c is in the topology on X.

(4.058) Exercise: Use De Morgan’s Laws, and (4.055), to
(a) show that the empty set is a closed set, and that X is a closed set;
(b) show that the union of two closed sets is a closed set;
(c) show that the intersection of an arbitrary collection of closed sets is a closed set.

(4.050) Exercise: Show that the union of a finite family of closed sets is a closed set.

A family of sets is “a finite family” if the family of sets contains finitely many sets; the sets *in* the finite family don’t have to be finite sets!

✒ the topological definition of “continuous”

(4.059) ***Definition:*** Let X be a non-empty set, with a topology $\boldsymbol{U}$. Let Y be a non-empty set, with a topology $\boldsymbol{V}$. Suppose that $f : X \to Y$. Then **f is continuous on X** if the inverse image of every open subset of Y is an open subset of X.

Note that *we* must be careful to keep Y and *its* topology in mind! This terminology ignores Y’s topology!

If more than one topology is involved, we can say that f is continuous **with respect to** the topology $\boldsymbol{U}$ on X, and the topology $\boldsymbol{V}$ on Y.

In other words (and symbols), “f is continuous on X” means that, for every open subset V of Y, $f^{-1}(V)$ is an open subset of X. Even more abstruse (obscure?): f^{-1}, *viewed as a mapping of sets to sets*, maps open subsets of Y to open subsets of X. I.e., $f^{-1}: 2^Y \to 2^X$, for every function $f : X \to Y$, but continuous functions have the property that $f^{-1}(\boldsymbol{V}) \subseteq \boldsymbol{U}$.

Please learn to read the last inclusion so it tells you “the inverse image of a set in $\boldsymbol{V}$ is in $\boldsymbol{U}$.”

Example: An easy but important example of a continuous function on a topological space is the **identity function**, $id : X \to X$, constructed by the rule $id(x) := x$, for all x in X. Since, for *all* subsets S of X, $id^{-1}(S) = S$ (why?), the inverse image of an open set is itself, hence open, so the identity function (or **identity map**) is continuous on X.

Example: Any transformation of the Euclidean Plane P that we use to move triangles and circles around on the plane, including rotations, translations, reflections and similarities, is continuous, because these transformations all preserve circles and their centers and insides. Details will be skipped, since we're "in the back room" when we use the Euclidean Plane!

☛ an important theorem about *composing* continuous functions

(4.060) Theorem: Let X, Y and Z be non-empty topological spaces. For all functions $f : X \to Y$, and for all functions $g : Y \to Z$, if f and g are continuous, then the composite function $g \circ f : X \to Z$ is continuous.

Proof: Suppose that f and g are continuous.
To show that $g \circ f$ is continuous, we need to show that, for all open subsets W of (the topology on) Z, the inverse image, $(g \circ f)^{-1}(W)$, of W is open in X.
We are given that:

for all open subsets V of Y, $f^{-1}(V)$ is open in X;
for all open subsets W of Z, $g^{-1}(W)$ is open in Y.

By the definitions of inverse image and composite function,

$$(g \circ f)^{-1}(W) = \{ x \in X : (g \circ f)(x) \in W \} = \{ x \in X : g(f(x)) \in W \}$$
$$= \{ x \in X : f(x) \in g^{-1}(W) \} = \{ x \in X : x \in f^{-1}(g^{-1}(W)) \}.$$

The definition of inverse image was used three times. Suppose W is an open subset of Z.
Since g is continuous, and W is an open subset of Z, $V := g^{-1}(W)$ is an open subset of Y.
Since f is continuous, and $V = g^{-1}(W)$ is an open subset of Y, $U := f^{-1}(V) = f^{-1}(g^{-1}(W))$ is an open subset of X. This completes the proof!

When you learned about continuity in Calculus, the definition you worked with may have involved a limit, or perhaps "epsilons and deltas." If your definition of continuity involved a limit, then your definition of limit probably involved "epsilons and deltas." After we build the Real Number System, we'll return to this topic, and see how the "abstract, set-theoretic" definition of continuity is connected to the one you learned in Calculus. Next, we'll study the abstract way to talk about "continuity at a point."

There are some special terms that relate points of X and certain subsets of X.

(4.061) ***Definition:*** For all topological spaces X, a point **x is an interior point of a set S** contained in X if there exists an open set, Ω, — that is, a member of the topology, $\boldsymbol{T}$ — such that $x \in \Omega \subseteq S$ (that is, if x belongs to an open set that is contained in S).
We say that **a set S** (contained in X) **is a neighborhood of** (a point) **x** (in X) if x is an interior point of S.

Notice that x is a point of S, if x is an interior point of S. The points in the open set Ω are thought of as being "close to" x. The *meaning* of "close to" depends on the topology!

(4.062) Problem: Show that an open set is a neighborhood of each of its points.

(4.063) ***Definition:*** For all topological spaces X, a point **x is an exterior point of a set S** contained in X if x is an interior point of S^c, the complement of S.

Notice that x is NOT a point of S, if x is an exterior point of S, despite the "of!"

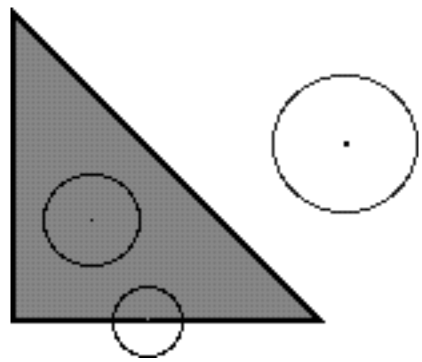
A point of X might be neither an interior point of a set S nor an exterior point of S. Such points will be called "boundary points."

(4.064) ***Definition:*** For all topological spaces X, a point **x is a boundary point of a set S** contained in X if every open set that contains x contains points of S and also contains points of S^c.

Note that "contains points of S" means "contains at least one point of S, not necessarily more than one! Also note that x can always be used for one such point.

Notice that x may, or may not be a point of S, if x is a boundary point "of" S!

Example: Consider a triangle in the Euclidean Plane, and the points inside it, to be our set S. Recall that a (usual) neighborhood of a point p must contain all the points inside some small circle centered at p. The points inside the triangle, but not the points on its sides, are interior points, because we can imagine a small enough circle, centered at any point in the shaded part, that has all shaded points inside it. And points outside the triangle, but not on the sides, are all exterior points, because we can imagine a small enough circle, centered at any point in the unshaded part, that has all its points in the unshaded part. Finally, any point that is on a side of the triangle is a boundary point, because every circle centered there has points inside it that are in the shaded part and points inside it that are in the unshaded part.



Interior, exterior and boundary points of a set S

(4.065) Exercise: Show that a point x is a boundary point of S if and only if *every* neighborhood of x contains a point of S and a point of S^c.

(4.066) Problem: Show that a subset S of a topological space is closed if and only if S contains all of the points that are boundary points of S. This is a "more descriptive" way to "define" "closed set" than just to say that a set is closed if its complement is open.

(4.067) Problem: Show that, if S is an arbitrary subset of a topological space X, then for every point x of X, exactly one of the following mathematical statements is true:

(i) x is an interior point of S;

(ii) x is a boundary points of S;

(iii) x is an exterior point S.

When you learned about continuity in Calculus, you learned about "continuity at a point." Here is the set-theoretic version. In this version, we will talk about functions defined "near" a point, namely, in a neighborhood of the point. So when we talk about continuity at a point, we're taking a "local" point of view, rather than the "global" point of view we used when we discussed continuity of functions defined on the whole space X.

☛ the *topological* definition of continuity at a point

Here are some concepts, topological and otherwise, that we will use right away:
neighborhood (of a point x), meaning the point x is an

interior point, of the neighborhood, meaning that there is an
open set containing the point that is, in turn, contained in the set called a neighborhood;
inverse image (of a set V) with respect to a function (or mapping), denoted $f^{-1}(V)$;
image (of a set U) with respect to a function (or mapping), denoted $f(U)$;
function as graph — $f : X \to Y$ is a set f of ordered pairs contained in $X \times Y$.

First, we will see the *topological* definition of a function being continuous at a point, then soon thereafter the *topological* definition of a function having a limit at a point.

(4.068) ***Definition:*** Let X and Y be non-empty topological spaces. Suppose that x_o is a point of X, and that f is a function, defined on a neighborhood of x_o, that takes values in Y. Then **f is continuous at $\mathbf{x_o}$** if the inverse image of every neighborhood of $f(x_o)$ is a neighborhood of x_o.

In other words, f is continuous at x_o if $f : S \to Y$, where S is a neighborhood of x_o, and for every neighborhood V of $f(x_o)$, $f^{-1}(V)$ is a neighborhood of x_o.

(4.069) Problem: Show that f is continuous at x_o if and only if for every neighborhood V of $f(x_o)$, there exists a neighborhood U of x_o such that $f(U) \subseteq V$. In other words, f is continuous at x_o if and only if for every (given) neighborhood V of $f(x_o)$, there exists a neighborhood U (depending on V) of x_o such that, for all x in U, $f(x) \in V$.

(4.070) Problem: Show that f is continuous at x_o if and only if for every neighborhood V of $f(x_o)$, there exists a neighborhood U of x_o such that $f \cap (X \times V) \subseteq U \times V$. Here, we regard f as a subset of $X \times Y$ — a graph!

☛ when is a function NOT continuous at a point?

To answer this question, we have to look at the denial of the definition of continuity at a point. Saying that f is NOT continuous at x_o means that there exists a particular neighborhood M of $f(x_o)$ such that $f^{-1}(M)$ is not a neighborhood of x_o. What does it mean for $f^{-1}(M)$ *not* to be a neighborhood of x_o? It means that x_o is not an interior point of $f^{-1}(M)$. Thus, x_o is a boundary point of $f^{-1}(M)$ (why?). This means that *every* neighborhood of x_o contains a point that is not in $f^{-1}(M)$. That is, every neighborhood of x_o contains a point $\hat{x}$ such that $f(\hat{x}) \notin M$.

(4.071) Exercise: Why is x_o a boundary point of $f^{-1}(M)$?

This is about all we can say in general: f is not continuous at x_o if there is a neighborhood of $f(x_o)$ that is avoided by the images of points arbitrarily close to x_o.

There is one thing that *might* be true: it could happen that f *would* be continuous at x_o if we would just assign the right value to f at x_o. When this is true, x_o is called a **removable discontinuity** of f, and **f has a removable discontinuity at x_o**. Thus, f has a removable discontinuity at x_o if there exists y_o in Y such that the inverse image of every neighborhood of y_o is a neighborhood of x_o with x_o (possibly) removed. For then we could remove the ordered pair $(x_o, f(x_o))$ from f and replace it by the ordered pair (x_o, y_o) to construct a new function that is continuous at x_o.

There is another name given to this situation, namely "limit." The idea of limit applies even when f is continuous at x_o. The idea of limit IGNORES the value of f at x_o, and simply asks **what happens to the values f(x) when x is arbitrarily "close" to x_o, without ever being equal to x_o**. To define limit, we need a minor definition, "deleted neighborhood."

(4.072) ***Definition:*** Let X be a non-empty topological space. Suppose that x_o is a point of X. Then a subset *D* of X is **a deleted neighborhood of x_o** if *D* does not contain x_o, and $D \cup \{x_o\}$ is a neighborhood of x_o.

Deleted neighborhoods are simply neighborhoods of x_o that have had x_o removed from them.

☛ the *topological* definition of limit

(4.073) ***Definition and notations:*** Let X and Y be non-empty topological spaces. Suppose that x_o is a point of X, and that f is a function, defined on a deleted neighborhood of x_o, that takes values in Y. Suppose y_o is in Y. Then **f has limit y_o at x_o** if the inverse image of every neighborhood of $f(x_o)$ contains a deleted neighborhood of x_o. This is denoted in many ways. One common notation is $\mathbf{lim_{x\to x_o} f(x) = y_o}$.
Sometimes this is written $\displaystyle\lim_{x\to x_o} f(x) = y_o$. We say that **f has a limit at x_o** if there exists y_o such that f has limit y_o at x_o. Often, we write **$f(x) \to y_o$ as $x \to x_o$** when f has limit y_o at x_o.

> **(4.074) Exercise:** Prove that, in the context of (4.073), $\lim_{x\to x_o} f(x) = y_o$ if and only if, for every neighborhood *M* of y_o there exists a deleted neighborhood *D* of x_o such that, for all $x \in D$, $f(x) \in M$ (or, in other symbols, $f(D) \subseteq M$).

(4.075) Problem: Prove that in the context of (4.073), f has a limit at x_o if and only if there exists a function g defined in a neighborhood N of x_o such that g is continuous at x_o and, for all x in $N \sim \{x_o\}$, g(x) = f(x). Prove also that, if such a g exists, then $\lim_{x \to x_o} f(x) = g(x_o)$.

This Problem can (and will) be used to convert Theorems about functions that are continuous at a point to Theorems about functions that have a limit at a point, and vice versa!

☛ when does a function NOT have a limit at a point?

To answer this question, we have to look at the denial of the definition of having a limit at a point. Saying that f does NOT have a limit at x_o means that, for ALL y_o in Y, there exists a particular neighborhood V of y_o such that $f^{-1}(V)$ is NOT a deleted neighborhood of x_o. What does it mean for $f^{-1}(V)$ not to be a deleted neighborhood of x_o? It means that x_o is NOT an interior point of $f^{-1}(V)$. This means that *every* neighborhood of x_o contains a point that is not in $f^{-1}(V)$. That is, every neighborhood of x_o contains a point $\hat{x}$ such that $f(\hat{x}) \notin V$. This has to be true for EVERY y_o in Y.

This means that f maps the points in deleted neighborhoods of x_o to at least two different, far-apart places, and probably more, in Y: f can't seem to make up its mind where to put the points that are near x_o!

☛ some more definitions, in the general topological context

We can gather interior points, exterior points, and boundary points into sets.

☛ interior, exterior, boundary and closure of a set

(4.076) ***Definitions and notations:*** For an arbitrary topological space X, for all subsets S of X,

the interior of S, denoted **int(S)**, is defined to be the set of all interior points of S;

the boundary of S, denoted **bdry(S)**, is defined to be the set of all boundary points of S, and

the closure of S, denoted **clos(S)**, is defined to be the union of S and bdry(S). The **exterior of S** is defined to be the interior of the complement of S. We can denote the exterior **ext(S)**.

Alternate notations are: $\mathring{\mathbf{S}}$, for int(S), S for bdry(S), $\overline{\mathbf{S}}$ for clos(S).

(4.077) Exercise: Show that, for all subsets of an arbitrary topological space X, the closure of S is the union of the interior of S and the boundary of S.

(4.078) Problem: Show that, for all subsets of an arbitrary topological space X, the boundary of S is equal to the boundary of S^c.

(4.079) Exercise: Show that, for all subsets of an arbitrary topological space X, the interior of S is the complement of the closure of S^c.

(4.080) Problem: Suppose that X is a non-empty set. Let S be a non-empty subset of X. With respect to the discrete topology on X, find S.

☛ limit point, or accumulation point, or point of accumulation

Here is another important topological concept, “limit point,” or “accumulation point.” At first, a limit point might seem to be a special kind of boundary point, BUT an accumulation point NEED NOT BE A BOUNDARY POINT!

(4.081) ***Definitions:*** Suppose that X is a topological space. Then, for all subsets S of X, and for all points p in X, **p is a limit point of S** if every neighborhood of p contains at least one point of S that is different than p. **A limit point of S need not be *in* S, nor need it be a boundary point of S.** A point p in S is **an isolated point of S** if there exists a neighborhood of p that contains no points of S other than p. **An isolated point of S *is* in S.** Limit points are also known as **accumulation points**, or **points of accumulation**.

Notice that an isolated point of a set S CANNOT be a limit point of S, and vice versa!

(4.082) Problem: Show that a set S in a topological space is closed if and only if S contains all the limit points of S.

☛ making new topologies out of old ones

We can perform set-theoretic operations on topologies, to construct other topologies.

☛ using a topology to construct topologies on subsets: relative topology

The next concept is “relative topology.” This will make it easier later for us to talk about such things as “functions that are continuous on a closed and bounded interval.”

(4.083) ***Definition:*** Suppose that X is a topological space. For all subsets Y of X, **the relative topology on Y** is the family of all subsets of Y constructed by forming the intersection of Y and an open set in the topology on X. We say that **a subset V of Y is relatively open** if there exists an open subset U of X such that $V = U \cap Y$. Some other terms used for “relative topology” are: **the topology on Y inherited from X**, **the topology on Y induced by X**.

I have never heard this called “**the topology of X restricted to Y**.” But it makes better sense to me, to think of “restricting” the topology on X to Y!

There is a minor point here that is a nuisance! We “officially” think of sets of the form $U \cap Y$ as subsets of X! They *are* subsets of Y, but they *officially* are members of 2^X. We can get around this problem by constructing a mapping ρ (for restriction) from 2^X to 2^Y by the rule $\rho(S) := \{ y \in Y : y \in S \}$. If you didn't think of this minor point, that's OK; but if you did, this is what we can do to solve the problem.

(4.084) **Problem**: Show that “the relative topology on Y” *is* a topology on Y.

☛ **combining two or more topologies on a set: the intersection of topologies**

(4.085) **Theorem:** Suppose that X is a non-empty set, and that $\boldsymbol{U_1}$ and $\boldsymbol{U_2}$ are two topologies on X. Then $\boldsymbol{U} := \boldsymbol{U_1} \cap \boldsymbol{U_2}$ is a topology on X. **Caution:** $\boldsymbol{U_1} \cap \boldsymbol{U_2}$ consists of subsets of X that are open subsets of X with respect to both of the topologies $\boldsymbol{U_1}$ and $\boldsymbol{U_2}$. The new collection, $\boldsymbol{U}$, is *not* to be construed as the collection of intersections of pairs of sets in $\boldsymbol{U_1}$ and $\boldsymbol{U_2}$! That would be a bigger collection!

Proof: We have to show that the three defining conditions for a topology are satisfied by $\boldsymbol{U}$.
Since X and $\varnothing$ are in both topologies, X and $\varnothing$ are in $\boldsymbol{U}$.
Thus, condition (i) is satisfied (see (4.055)).
If W_1 and W_2 are in $\boldsymbol{U_1} \cap \boldsymbol{U_2}$, then W_1 and W_2 are in $\boldsymbol{U_1}$ and in $\boldsymbol{U_2}$.
Therefore, $W_1 \cap W_2$ is in $\boldsymbol{U_1}$ and in $\boldsymbol{U_2}$, by (ii).
Hence $W_1 \cap W_2$ is in $\boldsymbol{U_1} \cap \boldsymbol{U_2}$. Thus, condition (ii) is satisfied for $\boldsymbol{U}$.

Now suppose $\mathcal{W}$ is an arbitrary collection of sets that are in $\boldsymbol{U}_1 \cap \boldsymbol{U}_2$.
Then $\mathcal{W}$ is contained in each of $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$,
so the union, W, of all the sets in $\mathcal{W}$ is also contained in each of $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$,
and this means that W, the union of all the sets in $\mathcal{W}$, is also contained in $\boldsymbol{U}_1 \cap \boldsymbol{U}_2$.
Thus, condition (iii) is satisfied, so $\boldsymbol{U}$ is a topology on X.

What worked for two topologies works for an arbitrary collection of them! You will now be asked to prove it; the proof can be made by adapting the preceding one.

(4.086) Theorem and Problem: Suppose X is a non-empty set, and that $\mathcal{U}$ is an arbitrary collection of topologies on X. Then the intersection of all the topologies in $\mathcal{U}$ (namely, the collection of all subsets of X that belong to all the topologies in $\mathcal{U}$) is a topology on X.

Prove the Theorem.

It may be useful to you, someday, to know that a definition that consists of statements about a (variable) family of sets (each statement of which having to be true of arbitrary members of the family, or all subfamilies that meet stated conditions), always has the property that the intersection of an arbitrary collection of such families also satisfies the definition's statements. Some people call this a "meta-theorem;" others call it "general nonsense."

(4.087) Theorem and *Definition:* Suppose that X is a non-empty set, and that $\mathcal{A}$ is an arbitrary family of subsets of X, namely a subset of 2^X. There exists a smallest topology that contains $\mathcal{A}$, known as **the topology generated by** $\mathcal{A}$.

Proof: The collection $\mathcal{U}$ of all topologies on X that contain $\mathcal{A}$ is non-empty, because $\mathcal{U}$ contains the discrete topology, namely 2^X.
Then the intersection, $\boldsymbol{U}$, of all the topologies in $\mathcal{U}$ is a topology on X, by (4.086).
Also, $\boldsymbol{U}$ contains $\mathcal{A}$ because each topology in $\mathcal{U}$ contains $\mathcal{A}$.
Now $\boldsymbol{U}$ is the *smallest* topology that contains $\mathcal{A}$, because every topology that contains $\mathcal{A}$ is in $\mathcal{U}$, and therefore contains $\boldsymbol{U}$, the intersection of $\mathcal{U}$.
This is what "smallest" means, in terms of set inclusion.
The proof is done.

(4.088) Problem: Suppose that X is a non-empty set, and that $\mathcal{A}$ is an arbitrary family of subsets of X. Construct the family $\boldsymbol{W}$ of arbitrary unions of finite intersections of sets that are in $\mathcal{A}$, and include X and $\varnothing$ in $\boldsymbol{W}$. Prove that $\boldsymbol{W}$ is the smallest topology on X that contains $\mathcal{A}$.

☛ combining two topological spaces: the product topology

Suppose that X and Y are topological spaces. We can construct a topology on the Cartesian product, $X \times Y$, by using the idea that the Cartesian product of two open sets is the *analog* of an open rectangle, so we can say that a subset S of $X \times Y$ is open if every point of S belongs to the Cartesian product of an open subset of X and an open subset of Y.

(4.089) Theorem and *Definition:* Suppose that X and Y are topological spaces. The **product topology on $X \times Y$** is the collection of all subsets W of $X \times Y$ that satisfy the condition:

for all (x, y) in W, there exist open sets U of X and V of Y such that
$(x, y) \in U \times V \subseteq W$.

The product topology is a topology on $X \times Y$, and it is the smallest topology that contains all the Cartesian products of open sets U of X and V of Y. Each of X and Y is called a **coordinate space** of the product.

Proof: We have two statements to prove: that the product topology *is* a topology, and that it is the smallest topology that contains all the Cartesian products of open sets U of X and V of Y.
Since the empty subsets of X and Y, respectively, are open, the empty subset of $X \times Y$ can be expressed as $\varnothing \times \varnothing$, hence is in the product topology.
Since X and Y, respectively, are open, $X \times Y$ can be expressed as the Cartesian product of two open sets, hence $X \times Y$ is in the product topology.
Now suppose that W_1 and W_2 are in the product topology.
Then for all (x_o, y_o) in $W_1 \cap W_2$, we know that there exist open sets U_i and V_i, $i = 1$ and 2, such that $U_i \times V_i \subseteq W_i$, $i = 1$ and 2, so $(x_o, y_o) \in (U_1 \times V_1) \cap (U_2 \times V_2)$.
Now we need

☛ a useful property of Cartesian products

$$A \times B \cap C \times D = (A \cap C) \times (B \cap D).$$

Let's use this, and prove it after.
Thus, $(x_o, y_o) \in (U_1 \times V_1) \cap (U_2 \times V_2) = (U_1 \cap U_2) \times (V_1 \cap V_2) \subseteq W_1 \cap W_2$, so the product topology is closed under finite intersections.

(4.090) Exercise: Explain the last sentence in more detail.

Now suppose $\mathcal{W}$ is an arbitrary collection of sets that are in the product topology. If (x_o, y_o) is in the union of all the sets in $\mathcal{W}$, then (x_o, y_o) is in a set W of the union, and so there are open sets U and V such that $(x_o, y_o) \in U \times V \subseteq W$, and W is contained in the union of all the sets in $\mathcal{W}$. Thus the union is in the product topology, so the product topology is a topology on X.

(4.091) Exercise: Explain the last two sentences in more detail.

Since the Cartesian product $U \times V$ of sets U and V that are open in X and Y, respectively, is in the product topology (why?), and sets in the product topology are unions of Cartesian products of open sets (why?), and because of the fact that $A \times B \cap C \times D = (A \cap C) \times (B \cap D)$, we can deduce, by (4.088), that the product topology is the smallest topology that contains all the Cartesian products of open sets U of X and V of Y.

(4.133d) Exercise: Explain the last sentence in more detail.

The proof is complete, except for showing that $A \times B \cap C \times D = (A \cap C) \times (B \cap D)$.

☛ **proof of $A \times B \cap C \times D = (A \cap C) \times (B \cap D)$**

If $(x, y) \in A \times B \cap C \times D$, then $(x, y) \in A \times B$ and $(x, y) \in C \times D$, so $x \in A \cap C$ and $y \in B \cap D$. This means that $(x, y) \in (A \cap C) \times (B \cap D)$.
The steps in this last argument can be reversed, to show that $(A \cap C) \times (B \cap D) \subseteq A \times B \cap C \times D$.

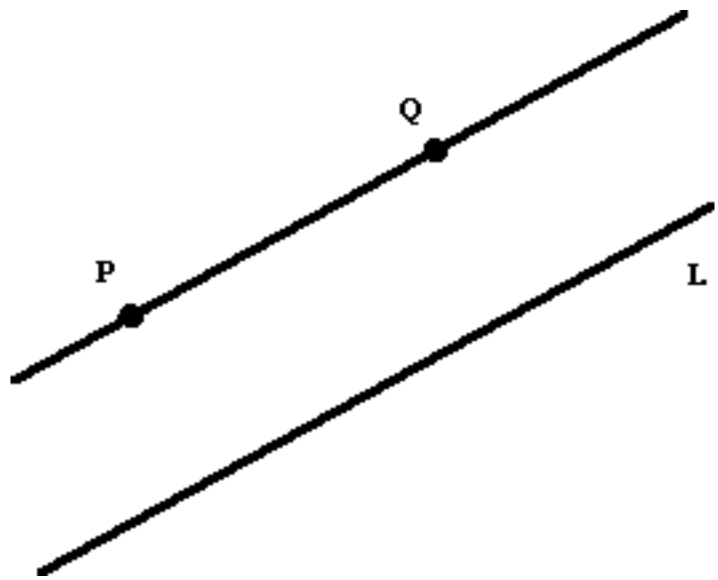
The proof of the Theorem is complete.

Now we are done with topologies for a while. It is time to turn to another major topic, that of "equivalence relations." You will probably have to look up the definitions of reflexive, symmetric and transitive, as applied to relations. See (4.009), (4.010) and (4.011).

EQUIVALENCE RELATIONS

(4.092) ***Definition:*** A relation E on a set A is an **equivalence relation** if E is reflexive, symmetric, and transitive. If $(x, y) \in E$, we say that **x and y are equivalent**. If x and y are each equivalent to z, then they are, by transitivity and symmetry, equivalent to each other, so we say that **x, y, and z are mutually equivalent**.

(4.093) Example: This example will be based on Euclidean Geometry. Let L be a line in the plane, $\mathcal{P}$. Then two points P and Q in the plane **are on a line parallel to L** if there is a line thru P and Q that is parallel to L. A relation E is now constructed: the set of all ordered pairs (P, Q) of points in the plane such that P and Q are on a line parallel to L. The relation E is a subset of $\mathcal{P} \times \mathcal{P}$. In order to tell whether or not E is an equivalence relation, we have to check whether or not *all* the defining statements in the definition of equivalence relation are true mathematical statements when interpreted in terms of this newly constructed E. Let us give to E the (cumbersome!) name "are on a line parallel to L."



First question: **Is E reflexive?** That is, is "R $\subseteq$ A$\times$A and (for all a in A)((a, a) $\in$ R)" true when we replace the variable A by $\mathcal{P}$, and replace the variable R by E? That is, is "E $\subseteq \mathcal{P}\times\mathcal{P}$ and (for all P in $\mathcal{P}$)((P, P) $\in$ E)" true? Note that the variable a was replaced by P. That was not necessary, because a was a bound variable. But we changed it because we are used to using capital letters to name points in the Plane $\mathcal{P}$.
Answer: Yes! Here is why. First, E is indeed a set of ordered pairs of points of $\mathcal{P}$, so "E $\subseteq$ $\mathcal{P}\times\mathcal{P}$" is true. Second, "(for all P in $\mathcal{P}$)((P, P) $\in$ E)" is true because there is exactly one line thru P that is parallel to L. Do you remember how to construct it, using straightedge and compass?

Second question: **Is E symmetric?** That is, is the following mathematical statement true? "E $\subseteq \mathcal{P}\times\mathcal{P}$ and (for all (P, Q) $\in \mathcal{P}\times\mathcal{P}$)((P, Q) $\in$ E $\Rightarrow$ (Q, P) $\in$ E)."

(4.094) Exercise: Verify that the answer to the Second question is Yes. Why is the statement in the Second question a *mathematical* statement, instead of just a statement?

Third question: **Is E transitive?** That is, is the following statement true?
"$E \subseteq \mathcal{P}\times\mathcal{P}$ and
$(\forall P \in \mathcal{P})(\forall Q \in \mathcal{P})(\forall R \in \mathcal{P})$ (((P, Q) $\in$ E and (Q, R) $\in$ E) $\Rightarrow$ ((P, R) $\in$ E))."
Answer: Yes! Why?
First, E is contained in $\mathcal{P}\times\mathcal{P}$.
Second, if P, Q, and R are *arbitrary* points in $\mathcal{P}$, the statement

"((P, Q) $\in$ E and (Q, R) $\in$ E) $\Rightarrow$ ((P, R) $\in$ E)"

is true because,
either "((P, Q) $\in$ E and (Q, R) $\in$ E)" is false, in which case the whole statement "((P, Q) $\in$ E and (Q, R) $\in$ E) $\Rightarrow$ ((P, R) $\in$ E)" is true, because it form is "A $\Rightarrow$ B," with "A" false,
or "((P, Q) $\in$ E and (Q, R) $\in$ E)" is true, in which case there is a line thru P and Q that is parallel to L, and there is a line thru Q and R that is parallel to L.
Since there is only *one* line thru Q that is parallel to L, (this is where we use Plane Geometry) and P and R both belong to that line, there is a line thru P and R that is parallel to L, namely, the one thru Q.
Since P, Q and R were arbitrary, the mathematical statement
"((P, Q) $\in$ E and (Q, R) $\in$ E) $\Rightarrow$ ((P, R) $\in$ E)" is true for all P, Q, and R in $\mathcal{P}$.

> **(4.095) Exercise**: Draw a figure in $\mathcal{P}$ that includes a line L, and several pairs of points P and Q, and several triplets of points P, Q, R, that provide examples of equivalent points and inequivalent points, and transitive and intransitive triplets of points.

☛ some more remarks about Definitions,
using the definition of equivalence relation as an example

A definition can involve several statements that have both bound and unbound, or free, variables. For example, in the definition of equivalence relation, the term "transitive" appears. This really means that the statement
"$R \subseteq A\times A$ and, $(\forall a \in A)(\forall b \in A)(\forall c \in A)$ (((a, b) $\in$ R and (b, c) $\in$ R) $\Rightarrow$ ((a, c) $\in$ R))."
is part of the definition, but with R replaced by E. We might have had to replace A by X, had the definition of equivalence relation begun "A relation E on a set X is an **equivalence relation** if…"
The statement
"$E \subseteq A\times A$ and, $(\forall a \in A)(\forall b \in A)(\forall c \in A)$ (((a, b) $\in$ E and (b, c) $\in$ E) $\Rightarrow$ ((a, c) $\in$ E))."
has a free variable: E. It has bound variables: a, b and c. Thus, it is not a mathematical state-

ment until a definite E is chosen. Once E is specified, the statement has a truth value. The bound variables then become part of the selected E in our example.

A definition has "blanks" to be filled in — namely the free variables.
Once values have been chosen for the free variables in the statements in a definition — and the same value has to be used for the same variable in all the statements in the definition — the statements are either all true, or at least one of them is false. If all are true, the free variables are said to **satisfy** the (conditions of) the definition. So we say, for example, that the relation "are on a line parallel to L," defined in (4.093), "is" an equivalence relation, because all the statements in the definition of equivalence relation are true when the set of ordered pairs of points in P that "are on a line parallel to L" is substituted for the free variable E.

The definition of "equivalence relation" means this: that from now on, in the context of this document, the two-word term "equivalence relation" has the meaning "set of ordered pairs in some $A \times A$ that is reflexive, symmetric and transitive." This is how every definition works. The *meaning* of the term being defined *is the set of statements in the definition.*

There is something slightly "wrong" with the way definitions are phrased in mathematics! The definition of equivalence relation began "A relation E is an **equivalence relation** *if…*" We "should" say "A relation E is an **equivalence relation** *if and only if*, " — but we don't.

A definition gives precise meaning to a term; after it is defined, the term or symbol has that meaning every time it is used.

Unfortunately, a term is sometimes used in different ways. That happens for reasons of history or analogy. Context usually helps make things clear.

There are two ways Definitions come up. Let us see how this works, using the definition of equivalence relation as an example. You can be given that a relation E is an equivalence relation, or you can have the task of telling whether or not a given relation is an equivalence relation.

One: suppose you are *given* that a relation E *is* an equivalence relation. What do you know? You know that there is a set A such that $E \subseteq A \times A$, and you know that each of the three following mathematical statements is true:

$(\forall a \in A)((a, a) \in E)$ (reflexive)

$(\forall\ (a, b) \in E)(\ (a, b) \in E \Rightarrow (b, a) \in E\ )$ (symmetric)
$(\forall a \in A)(\forall b \in A)(\forall c \in A)\ (\ ((a, b) \in E$ and $(b, c) \in E) \Rightarrow ((a, c) \in E)\ )$ (transitive).

One synonymous version of "(the mathematical statement) A is true" is "A holds." Another is "A is satisfied." These variations are used for variety and convenience.

Notice that we replace the "R" in the definitions of reflexive, symmetric, transitive with "E," because E is the given name of the relation, about which the mathematical statements in the respective definitions of reflexive, symmetric, transitive are true.

What can we deduce from the truth of all these quantified mathematical statements? Since all the quantifiers in them are universal, we can deduce that each of the statements
$(a, a) \in E$, $(a, b) \in E \Rightarrow (b, a) \in E$, $((a, b) \in E$ and $(b, c) \in E) \Rightarrow ((a, c) \in E)$ becomes a true mathematical statement whenever *any* specific values of the variables a, b, and c are chosen. That's about it. Then we'd presumably be able to use the given information, usually to deduce that some other mathematical statement is true.

(4.096) Example: Let C denote the set of all finite, ordered lists of positive integers taken from the set {1, 2, 5, 10, 20}. Repetitions are allowed in the lists. For example, {1, 1, 1, 5} and {2, 5, 1} are elements of C, as is the list {1, 2, 5, 10, 20}, which is different than {10, 20, 5, 1, 2}, also in C! We will say that two lists in C are *equivalent* if the sums of the numbers in each list are the same. Thus, {1, 1, 1, 5} and {2, 5, 1} are equivalent. Let E denote the relation, the set of all ordered pairs of *equivalent* lists from {1, 2, 5, 10, 20}. Suppose you are given that E *is* an equivalence relation. What, then, do you know is true? You know that, for every list L in C, (L, L) is in E. You know that, for every pair of lists (L, M), that if (L, M) is in E, then (M, L) is in E. You know that, if (L, M) and (M, N) are in E, then (L, N) is in E.

> **(4.097) Exercise**: How many lists are there in C that are *equivalent* to {1, 2, 5, 10}? Show that the number of lists that are *equivalent* to an arbitrary list in C is finite.

Two: You might be given a relation E — that is, detailed instructions about how to tell whether an ordered pair (a, b) is in E, and then be asked to answer the question *whether* E is an equivalence relation. Now, *you* have to find out whether or not each of the statements

$(\forall a \in A)((a, a) \in E)$ (reflexive)
$(\forall (a, b) \in E)((a, b) \in E \Rightarrow (b, a) \in E)$ (symmetric)
$(\forall a \in A)(\forall b \in A)(\forall c \in A)(((a, b) \in E \text{ and } (b, c) \in E) \Rightarrow ((a, c) \in E))$ (transitive)

is true.

Example: Let $A = \mathbf{N}$, the set of non-negative integers. We are given that an ordered pair (a, b) of non-negative integers is in E if and only if a and b have the same remainder when divided by 2. Is E an equivalence relation?

Solution: We have to check the truth-values of three mathematical statements. Let's try to take them in order.
Is $(\forall a \in \mathbf{N})((a, a) \in E)$ true? We are given that $(a, a) \in E$ if and only if a and a have the same remainder when divided by 2. They do, because $a = a$. Since this reasoning applies to an arbitrary element a of $\mathbf{N}$, $(\forall a \in \mathbf{N})((a, a) \in E)$ is true.

Notice that we replaced the "b" in the definition of "(a, b) ∈ E" by "a" to match the definition of reflexivity.

Is $(\forall (a, b) \in E)((a, b) \in E \Rightarrow (b, a) \in E)$ true? *Suppose* that $(a, b) \in E$. Then a and b have the same remainder when divided by 2. Thus, b and a have the same remainder when divided by 2. By the definition of E, $(b, a) \in E$. Therefore, $(a, b) \in E \Rightarrow (b, a) \in E$ is true.

Recall that, if $(a, b) \in E$ were false, $(a, b) \in E \Rightarrow (b, a) \in E$ would be true, vacuously. Therefore, we usually don't ever *consider* the possibility that the antecedent is false when we set out to prove that an implication is true, because it would be vacuously (automatically) true.

Since this reasoning applies to an arbitrary element (a, b) of E, the mathematical statement $(\forall (a, b) \in E)((a, b) \in E \Rightarrow (b, a) \in E)$ is true.

Is $(\forall a \in \mathbf{N})(\forall b \in \mathbf{N})(\forall c \in \mathbf{N})(((a, b) \in E \text{ and } (b, c) \in E) \Rightarrow ((a, c) \in E))$ true? Well, let a, b, c be given in $\mathbf{N}$, and suppose that $(a, b) \in E$ and $(b, c) \in E$ is true about a, b, and c. We have to deduce, using this information, that "$(a, c) \in E$" is true. Well, we are given that a and b have the same remainder when divided by 2, and that b and c have the same remainder when divided by 2. Thus a and c, whose remainders when divided by 2 are each equal to b's remainder when divided by 2, have the same remainder when divided by 2. By the definition of

E, (a, c) $\in$ E. Since this reasoning applies to arbitrary elements a, b, and c in **N**, the mathematical statement

$$(\forall a \in \mathbf{N})(\forall b \in \mathbf{N})(\forall c \in \mathbf{N})\ (\ ((a, b) \in E \text{ and } (b, c) \in E) \Rightarrow ((a, c) \in E)\ )$$

is true.

Having found that all three mathematical statements comprising the definition of "equivalence relation" are true for our given relation E, we conclude that the relation E is an equivalence relation.

The solution just given is an example of a "proof."

(4.098) Exercise: Show that *equivalent*, defined in (4.096), is an equivalence relation on C, also defined in (4.096).

(4.099) Exercise: Which of the following pairs are in the equivalence relation E of the last example? Why?

(1, 5), (7, 12), (28, 19), (0, 10000010), (1, 2), (6, 3), (9, 7), (33, 0), (33, 3).

(4.101) Exercise: Let f be a function with domain X and range Y. Define a relation E on X this way: E := { (u, v) $\in$ X$\times$X : f(u) = f(v) }. Show that E is an equivalence relation, and that F is one-to-one if and only if E = Δ, the diagonal in X$\times$X.

(4.100) Problem: Discuss, in the manner of the last Example, how to show that a relation is NOT an equivalence relation. Illustrate your discussion with an example of a relation that is not an equivalence relation.

You need to memorize Definitions!

☛ why equivalence relations are important

If E is an equivalence relation, we can treat elements that are equivalent to each other as though they were the same. This allows us to concentrate on only those properties of the elements that we are interested in. We can ignore, even if only temporarily, the differences between them.

Example: Consider the set C and the relation *equivalent* defined in (4.096). We might think of a list in C as a list of the denominations of bills we take out of a wallet, one at a time, listed in that order. Two *equivalent* lists then represent the *same amount of money*.

Equivalence relations are used by people every day, though unconsciously! If you're paying a restaurant check, two \$5 bills, are equivalent to one \$10 bill, which is, in turn, equivalent to ten \$1 bills. You can think of many examples, and I hope you will observe, to see more. Equivalence relations are used in mathematics to express "essential" equality — relative to some property of interest at the moment. We will use equivalence relations to build our number systems on the foundation of the natural numbers.

(4.102) Exercise: Verify that, if E is an equivalence relation, then $E^{-1} = E$. See (4.007) for the definition of E^{-1} in terms of E.

☛ **more on equivalence relations: *equivalence classes* and the *quotient space***

(4.103) ***Definition and notation:*** If E is an equivalence relation on a set X, and $x \in X$, we define **the equivalence class containing x** to be the set denoted by **[x]** and given by

$$[x] := \{ y \in X : (x, y) \in E \}.$$

Some synonyms for $[x]$ are **the equivalence class generated by x**, or **the equivalence class of x**.

Each equivalence class $[x]$ is called an **equivalence class determined by E**. This makes sense, because different equivalence relations on X usually have different equivalence classes containing an element x of X. Sometimes, confusingly, equivalence classes are called equivalence classes of E, and this gives the impression that an equivalence class of E is a member of E, but that is not so! It is just another of the many mathematical ambiguities that exist!

(4.104) Exercise: Show that, if E is an equivalence relation on a set X, and x_1 and x_2 are elements of X, then $[x_1] = [x_2]$ if and only if x_1 and x_2 are equivalent, that is, if and only if (x_1, x_2) belongs to E.

(4.105) Exercise: Show that, if E is an equivalence relation on a set X, and x_1 and x_2 are elements of X, then $x_1 \in [x_2]$ if and only if $x_2 \in [x_1]$.

Let us call an arbitrary element of an equivalence class a **representative** of the equivalence class; a representative element of an equivalence class **represents** the equivalence class.

Example: Recall the equivalence relation "are on a line parallel to L," introduced in (4.093). Let P be a point in $\mathcal{P}$. Then the equivalence class $[P]$ consists of all points Q in the plane such that there is a line thru P and Q that is parallel to the given line L. This is the line thru P that is parallel to L. Each point Q on the line thru P that is parallel to L thus represents the whole equivalence class.

Example: Recall the equivalence relation *equivalent*, on the set C of all finite, ordered lists of positive integers taken from the set {1, 2, 5, 10, 20}, with repetitions allowed in the lists. This was introduced in (4.096). The equivalence class of the list {1, 1, 1, 5} is denoted [{1, 1, 1, 5}], and [{1, 1, 1, 5}] consists of the lists {5, 2, 1}, {5, 1, 1, 1}, {2, 2, 2, 1, 1}, {2, 2, 1, 1, 1, 1}, {2, 1, 1, 1, 1, 1, 1}, and {1, 1, 1, 1, 1, 1, 1, 1}, as well as 5 other rearrangements of {5, 2, 1}, 3 other rearrangements of {5, 1, 1, 1}, 9 other rearrangements of {2, 2, 2, 1, 1}, 14 other rearrangements of {2, 2, 1, 1, 1, 1}, and 6 other rearrangements of {2, 1, 1, 1, 1, 1, 1}.

(4.106) Problem: Show that, if E is an equivalence relation on a set X, then any two equivalence classes $[x_1]$ and $[x_2]$ are either equal (as sets) or are disjoint.

(4.107) Theorem: If E is an equivalence relation on a set X, and S is a subset of X that consists of elements that are mutually equivalent, then for all x in S, $S \subseteq [x]$. If, in addition, S contains all the elements of X that are equivalent to any one of the elements of S, then for all x in S, S = [x].

"Mutually equivalent" here means that any two elements x, y in S are equivalent.

Proof: Let $x \in S$. Since [x] contains all elements of X that are equivalent to x, and every element of S is equivalent to x, $S \subseteq [x]$. If there is some a in S such that $S \supseteq [a]$, then for any x in S, $[a] \subseteq S \subseteq [x]$. But then [a] = [x], so [a] = S = [x] for all x in S, as desired.

(4.108) Exercise: Why is the phrase "But then [a] = [x]" in the proof of (4.107) true?

(4.109) Exercise: Prove: X is the union of all the equivalence classes determined by E.

Example: Recall the equivalence relation "are on a line parallel to L," introduced in (4.093). Then any line M that is parallel to L consists of mutually equivalent points, because there is a line, namely M, parallel to L that passes thru each pair of points on M. And, if P is an arbitrary point of M, then M contains all the points Q in $\mathcal{P}$ such that the line thru Q that is parallel to L also contains P. Therefore, each equivalence class [P] determined by "are on a line parallel to L" is a line parallel to L, and conversely, every line parallel to L is an equivalence class [R] for some R in the $\mathcal{P}$.

The next Theorem includes a Definition, and some notation. The term being defined is "quotient space," and it is a tough definition, because it introduces a "new" concept. The newness is in the realization that we create "quotient spaces" all the time in everyday life. The reason for introducing quotient spaces is that, technically, each of the number systems "integers (positive and negative)," "rational numbers," and "real numbers" that are used so much in Calculus ***is*** a quotient space. We try to think of the equivalence classes determined by an equivalence relation as "points," that is, as single "indivisible" objects. Certainly, equivalence classes that contain more than one point *are* "divisible," because each equivalence class can be separated into its constituent points. But we use equivalence classes to mentally "lump together" all elements of X that are mutually equivalent. Otherwise, we could become bogged down in our thinking if we had to think of each of equivalent elements separately, when the same thoughts apply to each of equivalent elements of X.

(4.110) Theorem, *Definition and notation*: Let E be an equivalence relation on a set X. Then we will denote by **X/E** the subset of 2^X that consists of all the equivalence classes determined by E. We will call X/E the **quotient space of X by E**. There exists a function : $X \to X/E$, called **the projection of X onto X/E**, that is onto, and such that (x) = (y) if and only if $(x, y) \in E$. Moreover, X is the union of all the sets in X/E, and two elements of X/E, as subsets of X, either coincide or are disjoint. Sometimes we denote X/E by $\dfrac{\mathbf{X}}{\mathbf{E}}$.

The only things to prove are the statements about and the statements in the last sentence of the Theorem. The proof of this theorem has actually already been done, bit by bit, since "equivalence class was introduced in (4.103), but the theorem will be proved "from scratch."

Proof: **We construct the function** this way: for each x in X,
(x) := { $y \in X : (x, y) \in E$ } = [x]; namely (x) is the equivalence class that contains x.
The *elements* of X/E will be these equivalence classes, which are elements of 2^X.
To show that is onto, suppose that S is an equivalence class determined by E. Then there exists x in X such that S = [x]. Thus S = (x), so is onto.
Now **suppose that (u) = (v)**. Then there is an element x_o in both equivalence classes. This means that (u, x_o) and (v, x_o) belong to E. **Then** by symmetry and transitivity, **(u, v) belongs to E**. **Suppose that (u, v) $\in$ E**. Then (u) = { $y \in X : (u, y) \in E$ }, and
(v) = { $y \in X : (v, y) \in E$ }. Now suppose that x is *any* element of (u). This means that $(u, x) \in E$. Thus, $(x, u) \in E$. But since $(u, v) \in E$, and E is transitive, $(x, v) \in E$. Hence

$(v, x) \in E$, by symmetry. Thus, $x \in$ (v). This shows that (u) $\subseteq$ (v). The same argument works in the other direction. We have shown that, if u and v are equivalent (meaning: $(u, v) \in E$), **then (u) = (v)**. Therefore have shown that **(x) = (y) if and only if (x, y) $\in$ E**.

To show that **X is the union of all the sets in X/E**, let $x \in X$.

Then $x \in [x]$, so $X \subseteq \bigcup_{x \in X} [x]$. But since $[x] \subseteq X$ for each x in X, $\bigcup_{x \in X} [x] \subseteq X$. Thus

$X \_ \bigcup_{x \in X} [x]$.

Finally, let us show that **two elements of X/E, as subsets of X, either coincide or are disjoint**. Let $[x_1]$ and $[x_2]$ be elements of X/E. If $[x_1]$ and $[x_2]$ are disjoint, there is nothing to prove. If $[x_1]$ and $[x_2]$ both contain an element x of X, then by definition of $[x_1]$ and $[x_2]$, (x_1, x) and (x_2, x) belong to E. Thus, by symmetry, (x, x_2) belongs to E, so by transitivity (x_1, x_2) belongs to E. Again, by symmetry, (x_2, x_1) belongs to E. Now let w be any element of $[x_1]$. This means that (x_1, w) belongs to E. Thus, since (x_2, x_1) and (x_1, w) belong to E, and E is transitive, (x_2, w) belongs to E so, by definition, w belongs to $[x_2]$. This shows that $[x_1] \subseteq [x_2]$. The same argument works in the other direction. Thus, **if $[x_1]$ and $[x_2]$ both contain an element x of X, then $[x_1] = [x_2]$**. This completes the proof of the Theorem.

We could have used the notation (x) instead of [x] in the definition of the projection. Both notations are used. I chose this one because (x) "looks more like a function" than [x]. Did you notice that the last part of the proof is a repetition of an earlier part?

Example: Recall the equivalence relation "are on a line parallel to L," introduced in (4.093). The set of equivalence classes is precisely the set of lines in P that are parallel to L. Thus, the quotient space $\dfrac{P}{\text{"are on a line parallel to L"}}$ consists of the lines parallel to L. But we want to think of these lines as points! We can draw a line K perpendicular to L, and let the point that is on $K \cap [P]$ be the point to think about. Thus, the quotient space can be thought of as the set of points on K.

(4.111) Exercise: Draw a diagram that illustrates some equivalence classes of "are on a line parallel to L," and the quotient space $\dfrac{P}{\text{"are on a line parallel to L"}}$.

(4.112) Exercise: Recall the equivalence relation *equivalent*, on the set C of all finite, ordered lists of positive integers taken from the set {1, 2, 5, 10, 20}, with repetitions allowed in the lists. This was introduced in (4.096). What "is" the quotient space C/*equivalent* ? Your answer should single out some pertinent feature of each equivalence class that "describes" it efficiently.

We can "lift" a function defined on a quotient space back to the original space by constructing its values to be constant on equivalence classes. We can "condense" *some* functions defined on the original space to be functions on the quotient space, but only if they are constant on equivalence classes. This is what the next Theorem is about.

(4.113) Theorem: Suppose E is an equivalence relation on a set X, and that $f : X/E \to Y$ is a function defined on the quotient space, X/E. Then there exists a unique function $F : X \to Y$ such that, for all x in X, $F(x) = f([x])$. Moreover, if $F : X \to Y$ is a function, then there exists a function $f : X/E \to Y$ such that, for all x in X, $F(x) = f([x])$, if and only if, for all x in X and for all y in X, xEy implies $F(x) = F(y)$.

Proof: This is really two Theorems in one. First, suppose that $f : X/E \to Y$. We construct F by the rule $F(x) := f([x])$. By construction, the function F is unique. This proves the first part! Now suppose $F : X \to Y$ is a function, and for all x in X and for all y in X, xEy implies $F(x) = F(y)$. Therefore, F is constant on each equivalence class. Note that F *may* take the same value for inequivalent arguments! For example, F could be a constant function. Now we can construct f. Construct f by the rule that $f([x]) := F(x)$. Does this really define a function on X/E? To check that the answer is YES, suppose that y x, but xEy. By hypothesis, $F(x) = F(y)$. Thus, $f([y]) = F(y) = F(x) = f([x])$, so equal equivalence classes are indeed assigned the same value, no matter which element of the equivalence class is used to construct the value of f. Finally, suppose that $F : X \to Y$ is a function, and that there exists a function $f : X/E \to Y$ such that, for all x in X, $F(x) = f([x])$. We have to show that, for all x in X and for all y in X, xEy implies $F(x) = F(y)$. We know that, if xEy, then $[x] = [y]$. Therefore, $F(x) = f([x]) = f([y]) = F(y)$ by the Substitution Rule. This completes the proof.

(4.114) Exercise: Show that, under the conditions in Theorem (4.113), $F = f \circ$.

(4.115) Problem: Recall the equivalence relation "are on a line parallel to L," constructed in (4.093). Imagine a line K that is not parallel to L. Then K meets any line parallel to L in exactly one point. Construct two functions, F and G that map $\mathcal{P}$ into $\mathcal{P}$, as follows: For all P in $\mathcal{P}$, $G(P) := P$, and $F(P) :=$ the unique point where the line thru P that is parallel to L meets K. Show that G does NOT satisfy the conditions of Theorem (4.113), and that F does satisfy them. Find the image of the function f corresponding to F that maps the quotient space to P.

☛ how we will use quotient spaces

In the next Chapter, we'll officially introduce the Natural Numbers, **N**. No equivalence relations will be used there, and no quotient spaces. The *order* relations, less than, less than or equal, and so on, *will* be developed. So relations, functions, sets, lots of definitions, and lots of logic will be used.

In the Chapter after that, we'll do some "magic tricks" with **N**×**N**: we'll define an equivalence relation on **N**×**N**, and construct the quotient space, and thus produce the Integers, **Z**, both positive and negative, and we will have done it in a constructive way, designed to make the operations from **N** "transferable" to **Z** "naturally."

In the Chapter after *that*, we'll do some "magic tricks" with **Z**×**Z**: we'll define an equivalence relation on **Z**×**Z** (involving fractions) and construct the quotient space, and thus produce the Rational Numbers, **Q**.

In the Chapter after that, we *could* do some "magic tricks" with certain sequences of rational numbers, to construct the Real Numbers as equivalence classes. But we won't. Instead, well construct the Real Numbers using a method developed in the nineteenth century — Dedekind Cuts. Even there, we *could* use an equivalence relation, but we won't because it would be silly: the equivalence classes would either be singletons or doubletons!

So our use for quotient spaces is quite serious: we will use them to build the numbers on which Calculus is based!