Constructing Mathematics

Tyler Wang

Modified March 23, 2025

Preface

As technology propels our world forward, the days of calculating complex lunar trajectories by hand are behind us. In their place has risen a new era: The era of the computer. Today, we can entrust the burdens of computation to machines, freeing humans to pursue deeper, more meaningful endeavors. To reflect our changing world, a revolution in education must revitalize our schools to suit modern times.

Many math courses (especially in America) were primarily designed for an era when computers were insufficient for the computational demands of science. However, despite significant technological advancements, it seems as if the classroom has largely remained the same, if not stepped in a direction contrary to the needs of the modern world. In the 1950s to the 1970s, the *New Math* movement began as an effort to dramatically change the way we taught math in schools to reflect the changing times. However, this revolution was immediately rejected, and in its wake, it seems as if math education went backward in terms of the needs of the modern world with the introduction of Advanced Placement and, eventually, Common Core. While these were attempts to solve certain problems of education, in turn, it caused education to trend in the wrong way.

Now, more than ever, people feel that the things they learn in school (especially math) provide very little utility in their everyday lives. But (at least for math), it really doesn't have to be this way. Math is a subject that goes way beyond memorizing formulas and painstakingly crunching numbers; it's the study of the way humans think and rationalize concepts whose applications go far beyond physics, engineering, etc. The following text serves as my response to this issue: A revolution to change the way people perceive math and use it in a way that helps students beyond school.

What the subject of math really captures is a formalization of thought. Throughout history, math has had a profound impact on the development of human civilization. From the ancient Greeks to the many modern mathematicians, the goal of math has always been the same: To comprehend the uncomprehendable, to break down the complexity of the universe to concepts better suited for human understanding. This eventually led to the development of numerical computation, but this kind of math represents only a very small portion of the subject. There are plenty of things one can rationalize that can't be strictly represented numerically. Without proofs, it's very difficult to capture the true essence of math. A proof-based approach shines in its ability to demonstrate math's abilities to solve problems that are not strictly numerical.

Math goes far beyond the painstaking computation that we are all taught in schools to the point that when someone brings up mathematics, computation is all that comes to mind. But this perception of math couldn't be further from the truth. While most students might find the idea of proofs daunting, I believe this is largely due to their inexperience with them. My purpose with this book is to bridge that gap since I find that the many skills and lessons (as opposed to the content) of research math can readily be applied to our everyday lives and careers far outside of any immediate application of math. Whilst students may not go their entire lives remembering the definition of set or function, the mastery of critical thinking will remain for a lifetime...

Contents

Chapter 1

Mathematical Logic

1.1 Formal Logic

To begin our mathematical journey, let's start by formalizing the notion of logic. If one ponders it deeply, math, at its core, is just a game of manipulating a finite set of symbols based on a given set of rules and procedures. So, let's begin by defining the following:

An **alphebet** is a set of symbols that can consist of letters, numbers, arrows, mathematical symbols, etc.

A **formula** is a string of symbols along with some rule of determining whether a given string is a valid formula. We called these valid formulae *well-formed*. For example, if our alphabet consisted of the symbols $\{A, B, +\}$, we can say every valid formula in our system is a finite string where A and B are not adjacent. Hence, the following formulae are well-formed:

$$A + B+, \quad A + +B,$$

and the following are not well-formed:

$$ABA + AB, \quad A + +BA.$$

With these definitions out of the way, let's define the idea of a proof. Because we want our formal system to be as general as possible, our definition of a proof might not seem at first like a proof. But as we use these ideas to develop first propositional logic and then first-ordered logic, the motivation for such a definition might start to become more apparent.

With that, we start by defining a **proof** to be an ordered list of formulae where each element is either given by an **axiom** or by a **rule of inference** from any previous statements.

An **axiom** is a set of well-formed formulae. To help define an axiom, it is common to specify an **axiom schema**, which is a mechanical rule for determining whether a statement is an axiom. Continuing with our system as defined before, we can define the set

$$\mathbb{A} = \{A + A, A + B, A + B, B + B\}$$

to be the set of axioms. Alternatively, we define this set using an axiom schema by saying

A formula is an axiom if it's of the form

$$a + b$$

$$\textit{for } a, b = A \textit{ or } a, b = B.$$

Then, a **rule of inference** is a rule for assigning a certain set of formulas to the new formula. As an example, let's define the following rule in our system:

$$\textit{From } A + A \text{ and } B + B, \textit{ infer } +.$$

Taking all of this, we could construct the following proof:

(1) B+B

(2) A+B

(3) A+A

(4) +

where (1), (2), and (3) are from our axioms, and (4) is inferred by our rule of inference.

Finally, the formula listed at the end of a proof is a **theorem**, where the entire proof can be denoted as the proof of said theorem. Our previous example can be thought of as the proof for the theorem +.

1.2 Propositional Logic

Now that we have formalized the notion of logic, let's start to actually do something a bit more useful. Due to the generality of a formal system of logic, without additional structure, it's really hard to make sense of just some arbitrary system. For example, in our preceding section, we made no mention of the notion of truth, which should be a core concept in any logic system that does anything mathematical.

1.2.1 Seeking Motivation

In this section, we will start by using our intuition to understand how we want to build our formal system in a way that fits our perception of reality. After all, a system that resembles very little what we might think may be of little use in our journey to formalize *human* thought.

Let's start by examining the properties of manipulating simple true/false expressions. When examining true/false statements, our variables only evaluate to two outcomes; it's often convenient to prove theorems or define operations in what's called a **truth table**. Let's show this by defining the operators we seek to use in this way.

First, we define the **conjuction** operator, denoted as $p \wedge q$ where p, q are variables that either evaluate as T or F. As seen in figure (1.2.1), the possible evaluations of p are listed in the leftmost column. Similarly, the middle column lists the possible evaluations of q. The rightmost column lists the evaluations of $p \wedge q$, given the evaluations of p and q in the same row.

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

Figure 1.2.1

As such, examining the truth table, one finds that the conjunction operator functions eerily similarly to *and* in the English language. This is no mistake as the word *and* can be thought of as the conjunction operator in the English language.

Next, we can define the disjunction operator, denoted as $p \vee q$, which has the truth table given in figure (1.2.2). As one might expect, disjunction is essentially the word *or* in the English language.

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

Figure 1.2.2

Finally, we define the negation operation, denoted as $\neg p$, which contrary to disjunction and conjunction, is a unary operator[1] The truth table of negation is shown in figure (1.2.3) and, as one expects, is essentially the not operator.

p	$\neg p$
T	F
F	T

Figure 1.2.3

We will, however, not be referring to these operators by their technical names and instead be referring to $p \vee q$ as **or**, $p \wedge q$ as **and**, and $\neg p$ as **not**.

Now, let's examine the properties of these operations. First, we note that the binary operations are both **commutative** and **associative**. The truth tables for either are similar. Therefore, I only show **or** in figure (1.2.4).

For the binary operations, we notice that T and F act in a way like identity elements,

$$p \vee \mathrm{F} = p \quad \text{and} \quad p \wedge \mathrm{T} = p$$

[1]Acting on one one input.

p	q	$p \vee q$	$q \vee p$
T	T	T	T
T	F	T	T
F	T	T	T
F	F	F	F

p	q	r	$(p \vee q) \vee r$	$p \vee (q \vee r)$
T	T	T	T	T
T	T	F	T	T
T	F	T	T	T
T	F	F	T	T
			⋮	

Figure 1.2.4

like the elements 0 and 1 do for addition and multiplication. Now, to this point, it seems more and more that the **or** and the **and** operators are operating more and more like addition and multiplication.

Applying our observations that these operations operated quite similarly to addition and multiplication, we should expect

$$p \wedge (q \vee r) = (p \wedge q) \vee (q \wedge r).$$

If one were to construct these statements' truth tables, we would find they indeed match up.

Now, whilst these parallels are nice, we have gotten to the point where this analogy starts to diverge. By constructing the truth table for

$$p \vee (q \wedge r),$$

one would find this is equivalent to the truth table for

$$(p \vee q) \wedge (p \vee r)$$

suggesting that $\vee$ *also* distributes over $\wedge$.

Up to this point, it seems that by interchanging every $\vee$ and $\wedge$, then every T and F, the truthfulness of an identity remains unchanged. If this statement were true, it would help us substantually, as if we establish some theorem, by simply interchaning the symbols, we would get another true statement. Well, with a stroke of luck, this is precisely the case for $\vee$ and *wedge* in. Math, we called these relations **duality principles**, and when they exist, they become a very powerful tool. Without getting too specific with the definitions here, concepts are dual roughly when they are defined anti-symmetrically to each other. Thus, due to the symmetry, with the inverted parts reversed, the same proof may be repeated.

In our case, $\wedge$ is *dual* to $\vee$. The proof of this duality principle should be immediately obvious by observing our definitions for $\vee$ and $\wedge$: If we interchange T and F in the definition for $\vee$, we get the definition for $\wedge$ and vise versa.

Interchanging T and F, we find $\vee$ can by defined in terms of $\wedge$ as follows:

$$p \vee q = \neg(\neg p \wedge \neg q)$$[2]

Then by observing

$$\neg\neg r = r$$

for any r, we have the identity

$$\neg(p \vee q) = \neg\neg(\neg p \wedge \neg q) = \neg p \wedge \neg q.$$

This is what's known as **De Morgan's law**, and by duality, we also have

$$\neg(p \wedge q) = \neg p \vee \neg q.$$

After these, we have the complement law, which also comes in pairs, namely

$$p \vee \neg p = \mathrm{T}, \quad p \wedge \neg p = \mathrm{F}.$$

In some sense, we can describe $\neg$ as the *inverse assigning* operator, similar to how $-$ assigns inverses with addition (ie $a + (-a) = 0$), except the inverse assigning operator is the same for both operations by duality.

Two remaining identities that are of importance are fairly trivial to see, and we won't go too in-depth into them. These results are the

[2]I should note, to simplify notation, we will be imposing an order of operations: Namely we evaluate $\neg$ before $\vee$ and $\wedge$ when parenthesis are not present.

$$\text{Idempotent Laws:} \quad p \vee p = p, \quad p \wedge p = p,$$
$$\text{Absorption Laws:} \quad p \vee (p \wedge q) = p \wedge (p \vee q) = p.$$

1.2.2 Formalization

Now, let's take our observations and define a formal system that reproduces these properties. It might be good to stop at this time and first discuss why it might be a good idea to formalize logic. After all, isn't what we have above almost good enough for most purposes?[3]

Well, if we were to just stick with higher level structures, like those seen in algebra or geometry, this is as far as we probably need to go in terms of *formalizing* logic. However, defining propositional logic as a formal system also has its advantages. For one, by giving it a rigorous definition, one could treat propositional logic as a *mathematical object*. Throughout this book, we will develop different methods to interpret and manipulate different mathematical objects to reveal hidden properties the said objects might have. As it turns out, understanding our underlying logic systems has very important applications in not just math but also computer science. Many studies in theoretical computer science emerge from the study of logical systems; as such, formalizing logic has become of great importance in the past century as computers have slowly started taking over the world.

Now, with this side tangent out of the way, let's get back to describing what we did in the previous section, by instead in a formal setting.

Whilst we could stick with our **or**, **and**, and **not** operators, as with convention, we usually just reduce this down to just two operators: Namely **not** and **implies** ($\rightarrow$), or more formaly, the **conditional** operator.

As the name suggests, we should define this new operation in a way that closely resembles the word *implies* in the English language. Moving back to our informal system of true/false statements, let's examine how we should define this operator.

Let's take the statement

Grey skies implies it rains tomorrow.

[3] I say almost here since to actually describe math, we are still missing a few key pieces like quantifiers. We will discuss this soon.

Examining when this statement is true, let us first suppose the first part of the statement is never true, hence

The skies are never grey.

In this situation, it doesn't matter whether it rains tomorrow or not; the statement is **vacuously true** since the first part of the statement is never satisfied.

Now, suppose we observe a grey sky; then, for the statement to be true, we would require it to rain the next day, or the statement becomes a false implication.

With this observation in mind, we come up with the truth table as shown in figure (1.2.5). Comparing this with the truth table of $\neg p \vee q$, we can conclude:

$$p \to q = \neg p \vee q.$$

p	q	$p \to q$
T	T	T
T	F	F
F	T	T
F	F	T

Figure 1.2.5

Now, suppose we had

$$p \to q \quad \text{and} \quad q \to p,$$

We could replace this with a single new operator, namely

$$p \leftrightarrow q.$$

When we have a conditional operator going in both directions, we call this in logic a **biconditional**. In plain English, we might call this relation **if and only if**; hence, we can restate the previous statement as

$$p \textit{ if and only if } q.$$

You can also interpret the biconditional operator as a sort of *logical equiv-*

alence. We are quite ready to discuss the notion of equivalence yet, but I introduce this terminology since in math, often you might hear the phrase *the following are equivalent*, and I wanted to make sure I define it here. The biconditional has the truth table depicted in figure (1.2.6)

p	q	$p \leftrightarrow q$
T	T	T
T	F	F
F	T	F
F	F	T

Figure 1.2.6

Now, we are ready to define propositional logic. First, we have our alphabet, which will consist of our alphanumeric characters, which each on their own will be a well-formed formula, called an **atomic formula**; the symbols $\rightarrow$ and $\neg$, which will make up our **connectives**; and brackets, as to enforce an order of operations.[4]

We will define every atomic formula as well-formed. Then we can generate new formulae inductively from the following set of rules:

1) If p and q are both well-formed, then $(p \vee q)$ is also well-formed,

2) If p and q are both well-formed, then $(p \wedge q)$ is also well-formed,

3) If p is well-formed, then $\neg p$ is well-formed.

Hence, the following formulae are well-formed:

$$(p \rightarrow \neg q) \rightarrow \neg q, \quad p, \quad \neg p, \quad \ldots$$

and the following are *not* well-formed:

$$(p \rightarrow\rightarrow q), \quad \neg \rightarrow p, \quad \ldots$$

We will also impose that all well-formed formulae are finite.

As for axioms, we will have three schemes, namely

(A1) $\phi \rightarrow (\psi \rightarrow \phi)$

[4]We will stick with our previous convention of evaluation $\neg$ before $\rightarrow$ when parenthesis are not present.

(A2) $[\phi \to (\psi \to \theta)] \to [(\phi \to \psi) \to (\phi \to \theta)]$

(A3) $[(\neg\psi \to \neg\phi) \to (\phi \to \psi)]$

where ϕ, ψ, θ is any valid formula.

The idea of defining these axioms is that we can provide proofs for true statements in our formal system when we use these. Now, we haven't determined what it means for a statement to be true or how notions of true and false fit into this model, but once we do, we would need to show that statements with proofs in this system are indeed true. Also, if we expect proved statements to be true, naturally, the axioms listed here must be true since a one-element proof with any of these axioms is a valid proof of the axiom.

Then, for our rules of inference, we will have one, known by *modus ponens*, which states

$$\textit{Given } p \textit{ and } p \to q, \ \textit{infer } q.$$

Using our formal system, we can prove the following:

Theorem 1.2.1. *For any valid formula p, $p \to p$ is a theorem in propositional logic.*

Proof. The formal proof for this formula goes as follows:

1) $p \to [(p \to p) \to p]$, by (A1)

2) $(p \to [(p \to p) \to p]) \to ([p \to (p \to p)] \to (p \to p))$, by (A2)

3) $[p \to (p \to p)] \to (p \to p)$, by modus ponens

4) $p \to (p \to p)$, by (A1)

5) $p \to p$, by modus ponens.

□

1.2.3 Interpretation of Truth

Now that we've defined how to manipulate and construct proofs in our formal system, let's provide our system with some meaning and discuss the notion of truth. In propositional logic, we generally deal with two types of statements, **tautologies** and **contradictions**. To define these terms, we need to discuss the **valuation** of propositional statements.

A valuation is a rule for assigning truth values to formulae in propositional logic. We recall that each formula in our system can be defined inductively starting from our atomic formulae. A connective provides the means to construct new formulae from previous ones. Each connective has an associated truth table, which can be used to define the truth value of our new statement. Therefore, to consider the truth value of a statement, we require that a truth value be assigned to each atomic formula.

We will define a valuation as the truth value of a statement under our system of true/false statements as defined in our previous section. By the above observation, we conclude it suffices to consider the valuation of a statement given the valuation of each of its atomic formulae. We will denote the valuation of a statement ϕ with $v_S(\phi)$, where S is a rule for assigning valuations to each atomic formulae.[5]

Example 1.2.1.

Let's consider the valuation of each of our axioms in propositional logic. Let's draw our attention first to our first axiom. If (p) and (q) are atomic formulae, we can define our rule S by letting

♠

If for every choice of S, the valuation of our statement under each S is true, then we claim this statement is a tautology. Therefore, by example (1.2.1), we deduce that each of our axioms are tautologies.

We define a contradiction as the dual concept to a tautology; if a statement has the valuation of false under any choice of S, we deduce this statement is a contradiction.

In the remainder of this section, we will restrict our attention solely to the discussion of tautological formulae.

Now, the natural question to ask is whether formulae that can be derived as proof are therefore also tautological. This is a remarkably non-

[5]The notation that we've chosen here for valuation resembles the notation for a *function*. This is no mistake since the valuation operation can be understood as a function taking propositional statements to either a value of true or false.

trivial question since, in general, our rules of inference need not interact with valuation in any sensible way. In other words, formulae derived from our rules of inference need not be tautological for an arbitrary formal system. In addition, our axioms need not be tautologies which complicates matters even further. But, when statements derived from a proof are indeed tautologies, we say this formal system is **sound**.

Theorem 1.2.2. *Propositional logic is sound.*

Proof. We've already shown in example (1.2.1) that our axioms are tautologies. Thus, we check that modus ponens preserves tautologies.

Suppose p and $p \to q$ are both tautologies. Then the result is immediate after examining the truth table in figure (**??**).

With these two facts, we prove the soundness of propositional logic. Suppose we have a theorem p of length arbitrary length N.

□

Chapter 2

Set Theory

2.1 Set Theory

Often, in mathematics, we like to group abstract things, also known as a mathematical object, into sets. A mathematical object is, simply put, a *thing* we like to talk about in math, such as functions, numbers, matrices, other sets, etc. Although seemingly simple, these sets are incredibly powerful, so powerful in fact that most modern mathematics today is built off the backbones of set theory. For our purposes here, we only need a very elementary understanding of set theory, as a deeper dive, though fascinating, can be very confusing, difficult, and at times controversial.[1]

2.1.1 Notation

Before we move any further, we should get some basic notation out of the way, below is a list of all of the important symbols that we will be using in this class:

Symbol	Meaning

[1]This controversy mostly arises from the infamous axiom of choice. Simply put, given any set of sets that are not empty, one can always find a way to way to map a child set to an element in that child set. Though seemingly simple, due to its complications with infinite sets, it has deep consequences in all mathematics and how we deal with infinities in the universe.

$\in, \notin$	in and not in respectively. Typically to denote that an object belongs or doesn't belong to a set
$\subset, \subseteq$	proper subset, subset, we discuss the distinction later
$\not\subset, \not\subseteq$	not a proper subset, or subset
$\setminus$	Set difference
$=$	two sets are equivalent to one another
$\cup$	Union
$\cap$	Intersect
$\overline{A}$	Compliment of set A
$\{x : \phi(x)\}$	Defines a new set using variable x on the condition $\phi(x)$
$A = (a, b), [a, b]$	Defines a set on the interval where for any x $a < x < b$ or $a \leq x \leq b$ respectively, $x \in A$.
$\varnothing$	empty set, a set with no elements

Of the following, let us first draw our attention to the idea of a subset. A subset, in most simple terms, is just a set that is entirely contained within a larger parent, as shown in figure (2.1.1). As you may have noticed,

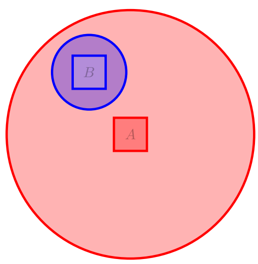


Figure 2.1.1: Here we claim $B \subset A$

I've decided to use $\subset$ as opposed to $\subseteq$ in the figure. Although both would have been equally valid, I specifically chose the former to shed light on the fact the two sets B and A are not equal. Similar to $>$ and $\geq$, the line underneath simply implies the possibility that the subset is equal to the original set.[2]

Now we have some basic intuition for this concept, let's shed some light on the rigorous definition of a subset, as such a definition will help us prove that a set is indeed a subset of another.

Definition 2.1.1. *$A \subseteq B$ if and only if [3]for every $a \in A$, $\in B$.*

With this definition, we can construct a definition of set equality. In particular, when we say two sets are equal, we should expect that every element in one set should be in another and vice versa. In other words, we expect

$$a \in A \text{ if and only if } a \in B.$$

Therefore, a definition that would naturally follow is

Definition 2.1.2. *We say $A = B$ is and only if $A \subseteq B$ and $B \subseteq A$.*[4]

The next symbol I'd like to focus on is the set difference. Very similar to subtraction, if we have two sets A and B, $A \setminus B$ is simply defined as the set of elements that are in A, but not in B as shown in figure (2.1.2).

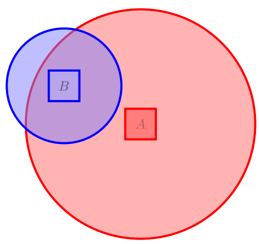


Figure 2.1.2: $A \setminus B$ defined in red

We can also formalize this idea in a more formal definition.

[2]It is important to note, however, that in certain texts, the distinction is not made between proper and non-proper subsets and the symbol $\subset$ is used for either case. In all my texts, I will be making this distinction, and $\subset$ will only refer to proper subsets.

[3]If you're unfamiliar with the term *if and only if*, when we say A if and only if B, we imply if A then B and if B then A, hence the implication goes both ways. Another way to say this is B is *necessary and sufficient* for A, or vice versa.

[4]As an exercise, try to prove this is equivalent to the characterization we gave before.

Definition 2.1.3. $x \in A \setminus B$ *if and only if* $x \in A$ *and* $x \notin B$.

2.1.2 Set Theory and Logic

The idea of the intercept or union can be shown most clearly with a Venn diagram like in figure (2.1.3). From here, it is clear that when taking the

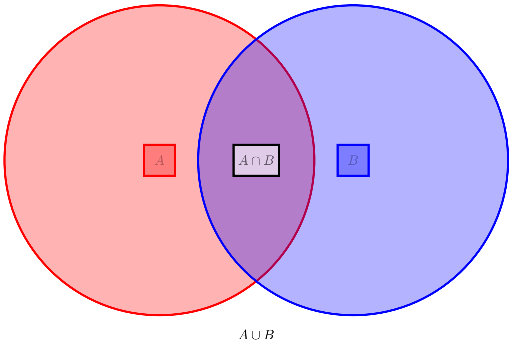


Figure 2.1.3

union of the two sets, we are essentially *adding* the two sets together, or other words, every element $x \in A \cup B$ is a member of either in A **or** B. An important note here is that a set doesn't contain duplicates of an object, therefore, when taking the union, an object is only included once in the output.

Then the intersect is the region where every element $x \in A \cap B$ is both in A **and** B.[5] I emphasize the "and" and "or" parts of the previous statements since I believe it's important to see the relationship between these very mathematical ideas and words we use in our everyday lives.[6]

Now, with this intuition of these operations, we can move on to rigorously define them as the following:

Definition 2.1.4. $x \in A \cup B$ *is and only if* $x \in A$ *or* $x \in B$.

[5] the intersect also bears stark resemblance with an elementary operation, which, as you may have assumed, is multiplication. We won't get into the details here, but just understand that it follows many of the same properties as multiplication

[6] It is important to note, however, that the or used in this case is slightly different from the or we use in everyday life. This is because we tend to deal more with exclusive or, where the or we deal with in math is non-exclusive. If you have no idea what I'm talking about, then forget that you even read this section.

Definition 2.1.5. *$x \in A \cap B$ is and only if $x \in A$ and $x \in B$.*

The complement of a set can be thought of exactly as the word **not**. We can generalize figure (2.1.4) by saying that every $x \notin A$ must be in

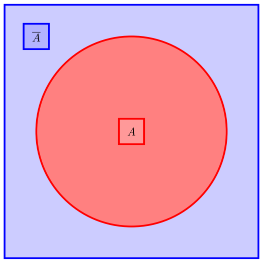


Figure 2.1.4: If the entire square is B, then $\overline{A}, A \subseteq B$

$\overline{A}$. It is important to note here that generally, when talking about a compliment, you need to know what you are "complementing" over, in essence, you must know what your set is a subset of. Unless you go into more advanced mathematics, generally, you can assume that this will be the real numbers, or in more isolated cases, the set of complex numbers, but nothing more than that. Now, formalizing our intuition gives the following definition:

Definition 2.1.6. *Suppose $A \subseteq U$. Then $x \in \bar{A}$ if and only if $x \in U$ and $x \notin A$.*[7]

2.1.3 Defining Sets

The last two notations in the chart are simply ways in which we can define new sets. Let's say we are trying to define a set that consists of only points on the unit circle. So, in this case, we would be defining a set of ordered pairs. We first note that any point that lies on a circle satisfies

$$x^2 + y^2 = 1.$$

Then, we can simply define S as

$$S = \{(x, y) : x^2 + y^2 = 1\}.$$

[7]Notice, we could have also defined $\bar{A}$ as $U \setminus A$, which has the advantage of specifying a global set which we are taking the compliment over.

This might still seem confusing, but typically, colons in math can be read as *such that* or *given* hence the statement above is equivalent to saying: *the set S is equivalent to the set of all real ordered pairs given they satisfy $x^2 + y^2 = 1$.* This notation is known as set builder notation since, in essence, we are building a set based on a given set of conditions. Sometimes, we like to be a bit more specific and explicitly declare that (x, y) is an ordered pair and write

$$S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

where $\mathbb{R}^2$ is the set of all real ordered paris.

We then draw our attention to an alternative way to define sets in real numbers, which is interval notation. Generally, in precalc, you will see this more often, as it is a quick and easy way to define a set on a given interval. Let's say we are trying to define a set S, where S is the set of all reals between 2 and 5, excluding the values 2 and 5 (the endpoints). Then we write

$$S = (2, 5).$$

Alternatively, we can change the parentheses to square brackets, where then we would have

$$S = [2, 5] = \{x : 2 \leq x \leq 5\}.$$

We see here that where parentheses correspond to strict inequality, square brackets correspond to a non-strict inequality. We can also mix and match and define sets like

$$[1, 5) = \{x : 1 \leq x < 5\},$$

or we can use it in combination with the union and intersect to combine multiple intervals such that

$$[1, 5) \cup (6, 10] = \{x : 1 \leq x < 5 \text{ or } 6 < x \leq 10\},$$

or

$$[1, 5) \cap (3, \infty) = \{x : 3 < x < 5\}.[8]$$

[8] When we deal with infinity, always use () since infinity is not in the real numbers, and hence we cannot include ∞ in our set.

2.1.4 Very Important Sets

Certain sets in mathematics are so important that we give them a special symbol

Symbol	Definition
$\mathbb{Z}$	Set of all integers
$\mathbb{N}$	Set of all natural numbers
$\mathbb{W}$	Set of all whole numbers
$\mathbb{R}$	Set of all real numbers
$\mathbb{C}$	Set of all complex numbers
$\mathbb{Q}$	Set of all rational numbers

Some of these, I hope, are very familiar to all of you, but I'd like to shed some light on $\mathbb{N}$ and $\mathbb{W}$, as these might not be so familiar to many of you. The whole numbers are simply defined as positive real numbers, including 0 or

$$\mathbb{W} = \{x \in \mathbb{Z} : x \geq 0\}.$$

Then, the natural numbers are simply

$$\mathbb{N} = \mathbb{W} \setminus \{0\},$$

or in other words, positive integers without 0.[9]

2.2 Proof by Induction

2.2.1 Well-Ordering Principle

An important idea that I'd like to draw our attention to is the idea of a well-ordered set. This might seem at first like a bunch of mathematical mumble jumble, but it has serious consequences when it comes to

[9]In different countries, this might be defined differently, like in France, $\mathbb{N}$ is defined in the same way as we defined the $\mathbb{W}$.

mathematical reasoning. But before we get to the implications, let's first examine what exactly this well-ordering principle is.

Let's start by looking at the natural numbers. Our first and most obvious observation is that the natural numbers are *totally ordered*, meaning you can take any two elements of the set and declare one is larger than the other, given they are not equal. Then, looking further, we realize by taking any subset of the natural numbers, finite or infinite, as long as this set isn't empty, there must be a least element since no matter what is in any of these subsets, 1 is always less than everything, and if 1 isn't in the set, we go to 2 or 3, and the same logic applies. This is a well-ordered set; we take this as a defining property of natural numbers.[10] We can formalize our observation here and propose a definition for a well-ordered set:

Definition 2.2.1. *Let S be a set where for every $a, b \in S$, exactly one of*

$$a < b \quad or \quad a > b \quad or \quad a = b$$

is true. Then S is ***well-ordered*** *if and only if for every $T \subseteq S$, there exists x_l such that for every $x \in T$, $x_l \leq x$.*

It is important to note here that this property occurs in other types of sets other than just the natural numbers, as we will soon find, we can extend this idea to a set of ordered statements.

Example 2.2.1.

We will demonstrate the well-ordered principle in this example. Let's use the set

$$S := \{2n : 2n > 7, n \in \mathbb{N}\}.$$

It should be immediately clear that this is the set of positive even numbers greater than 7. Therefore, we conclude $S \subseteq \mathbb{N}$. We can also check that 8 is indeed the least element of S, i.e., for every $k \in S$, $8 \leq k$.

Then, let's see why a set like $\mathbb{Z}$ is not well-ordered. Suppose

$$T := \{n : 2n, n \in \mathbb{Z}\}.$$

[10] Rigoursly proving this fact $\mathbb{N}$, using the formal definition of $\mathbb{N}$ is very difficult and confusing.

It should be clear that T is the set of all even numbers, positive or negative. Since for every element $k \in T$, we can find a smaller even number, we conclude that T cannot have the least element. Since the well-ordering principle requires that any arbitrary subset has a least element, we have $\mathbb{Z}$ is not well-ordered. ♠

2.2.2 Example of Proof by Induction

This may or may not be a new idea for some, but first, I will attempt to demonstrate the power of proof by induction by proving a very simple statement.

Example 2.2.2.

In the following example, I'd like to show for a sequence $1, 2, 3, ..., n$, where $n \in \mathbb{N}$, the sum is exactly

$$\frac{n(n+1)}{2}.$$

First, let's check that our formula works by checking against a sequence with $n = 1$. We find the sum is 1 and

$$\frac{1(1+1)}{2} = 1$$

hence verifying our formula works for $n = 1$. Let's call this case, where $n = 1$ is the base case. We can keep checking for $n = 2, 3, 4...$, but since we want to know that this formula works for all possible n, this is a never-ending process, thereby making it impossible to construct a rigorous proof. This means we must look for another technique to prove our desired statement. Let's suppose, for a minute, that for some $m \in \mathbb{N}$, our formula produces the right result. Then, using our assumption, we claim that

$$1 + 2 + ... + m + m + 1 = (1 + 2 + ... + m) + m + 1 = \frac{m(m+1)}{2} + m + 1$$

Then, by manipulating the expression, we get

$$\frac{m(m+1)}{2} + m + 1 = \frac{m(m+1) + 2(m+1)}{2} = \frac{m^2 + m + 2m + 1}{2}$$

$$= \frac{m^2 + 3m + 1}{2} = \frac{(m+1)(m+2)}{2} = \frac{(m+1)([m+1]+1)}{2}$$

which is exactly the expression we would get if we substituted $m + 1$ into our formula. Therefore, we have proven if the formula is true for m, then it must be true for $m + 1$. Let's call this part the inductive step. Here, I claim, without proof, that the base case, in combination with the inductive step, proves by induction that our formula works for all n. But a question remains: how exactly am I able to make such a bold conclusion?[11] ♠

2.2.3 Proof of Induction by Deduction

This brings us back to the well-ordered principle. First, we need to organize all of our statements into a set. This idea might seem quite unnatural, but if you think about our statements as functions that take in a natural number and output a true or false, this idea comes very naturally since we know we can organize functions into a set. As in our previous example, take for instance the statement $1 + 2 + ... + n$ **is equal to** $\frac{n(n+1)}{2}$ as a function of n, called $P(n)$, where n is any natural number, we now know that $P(n)$ is true for every n, hence the statements $P(1), P(2), P(3), ...$ are all true. We can make another example of this idea: suppose we define a statement $Q(n)$ as n **is odd**. Then we can define a set $S = \{Q(n) : n \in \mathbb{N}\}$. Then we know $Q(1), Q(3), Q(5), ...$ is true and $Q(2), Q(4), Q(6), ...$ is obviously false. Now equipt with this, we will attempt to prove our first lemma:

Lemma 2.2.1. *Let $P(n)$ be an arbitrary statement with argument $n \in \mathbb{N}$. If*

$$S := \{P(n) : n \in \mathbb{N}\},$$

then S is well-ordered.

Proof. We will attempt to prove this proposition with the help of figure (2.2.1). As shown in the diagram, we can map every natural number to a unique statement and vice versa. Then we can define an ordering on S such that for any $P(m_1), P(m_2) \in S$, $P(m_1) < P(m_2)$ if and only if

[11] Philosophically speaking, in general everyday life, inductive reasoning is inherently flawed since your conclusions rely on possibly flawed assumptions about the past. In math, we only accept statements as true if we can deduce them from other true statements, which contradicts what we've done here since we did exactly what I said was inherently flawed. But somehow, math has a way of deductively proving that what I did here, in fact, is okay.

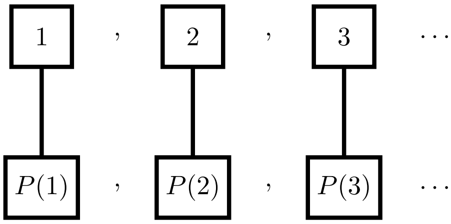


Figure 2.2.1

$m_1 < m_2$. Therefore, for any $T \subseteq S$. Then we can construct

$$N := \{n \in \mathbb{N} : P(n) \in T\}$$

Then, we have $N \subseteq \mathbb{N}$ implying N has a least element. Let n_l be the least element of N. Then, since $P(n_l) \in T$ by definition, we claim this is the least element of T. To show this is indeed the least element, consider any $P(m) \in T$. By definition of N, we have $m \in N$. Therefore $n_l \leq m$, implying $P(n_l) \leq P(m)$ thereby completing the proof. □

Theorem 2.2.1 (Induction)**.** *Suppose $P(n)$ is some arbitrary statement that depends on n. Then if*

1) $P(1)$ is true (Base case)

2) If $P(n)$ is true, then $P(n+1)$ is true. (Inductive step)

Then $P(n)$ is true for all n.

Proof. This proof will be by contradiction. First, assume there exists $P(n)$ that satisfies the above conditions, where for some $k \in \mathbb{N}$, $P(k)$ is false. Then define

$$S = \{P(n) : n \in \mathbb{N}\}.$$

Then let

$$F := \{P(n) \in S : P(n) \equiv \text{false}\}.$$

Then we have $F \subseteq S$. Then by our assumption, $P(k) \in F$ hence $F \neq \varnothing$. Since S is well-ordered by lemma (2.2.1), F must have a least element $P(n_l)$. $n_l \neq 1$, since by assumption $P(1)$ is true. Therefore, $P(n_l - 1) \in S$. But $P(n_l - 1) \notin F$, since $P(n_l - 1) < P(n_l)$, implies $P(n_l - 1)$ is true. This means $P(n_l)$ must also be true by the inductive step assumption. This is a contradiction of our original assumption, hence $F \equiv \varnothing$. This concludes our proof of induction. $\square$

To summarize, if we can simply prove that some base case $P(b)$ is true, by then proving that $P(n)$ being true implies $P(n+1)$ is true, we can confidently claim that for every $n \geq b$, $P(n)$ is true.

2.2.4 More Examples

Example 2.2.3.

In this example, we'd like to prove $9^n - 1$ is divisible by 8 for all $n \in \mathbb{W}$. To begin, we'd first like to prove that the base case is true, and in this case, the base case is $n = 0$.[12] Since

$$9^0 - 1 = 0$$

is divisible by 8, we can establish that our base case is true. Then, assume the statement is true for some arbitrary k. Then there exists, by the definition of divisibility, some $m \in \mathbb{Z}$ such that

$$9^k - 1 = 8m$$

which implies

$$9^k = 8m + 1$$

Then for $k + 1$,

$$9^{k+1} - 1 = 9(9^k) - 1 = 9(8m + 1) - 1$$

$$= 72m - 8 = 8(9m - 1)$$

Since $8(9m - 1) \in \mathbb{Z}$, this implies $9^{k+1} - 1$ is divisible by 8. Therefore, by induction for all $n \in \mathbb{W}$, $9^n - 1$ is divisible by 8. ♠

[12]For the base case, n does not necessarily have to be 0, but the index of the base case is the least that is guaranteed to be true with our current approach at induction (try to prove this fact by modifying the above proof).

Example 2.2.4.

In this example, we'd like to prove $n^2 < 2^{n+1}$ for any $n \in \mathbb{N}$. This will be a little less straightforward than in our previous examples, but what is the same is proving the base case is true, which is trivial by substitution of 1 for n. For our inductive step, we assume for some k, $k^2 < 2^{k+1}$. Then

$$2^{k+2} = 2 \cdot 2^{k+1} > k^2 + 2^{k+1}.$$

If we notice, what we want on the right side is

$$(k+1)^2 = k^2 + 2k + 1.$$

Somehow, we must convert the 2^{k+1} to $2k+1$. We could do this by proving

$$2^{k+1} > 2k + 1$$

for all $k \in \mathbb{N}$, which

$$2^{k+2} > (k+1)^2$$

would immediately follow. To prove the previous statement, we leverage induction again. By substituting $k = 1$, we find the base case is trivial. Then, for our inductive step, we let be $m \in \mathbb{N}$ such that $2^{m+1} > 2m + 1$. Then

$$2^{m+2} = 2 \cdot 2^{m+1} > 2(2m+1) = 2m + 2m + 2.$$

Since $m \geq 1$,

$$2m + 2m + 2 > 2m + 4 > 2m + 2 + 1 = 2(m+1) + 1.$$

Therefore, by induction, we conclude

$$2^{k+1} > 2k + 1$$

hence

$$2^{k+2} > k^2 + 2^{k+1} > k^2 + 2k + 1 = (k+1)^2.$$

Therefore, by induction, we claim $n^2 < 2^{n+1}$. ♠

2.3 Exercises

1. Let $A = \{1, \beta, 20\}$, $B = \{\text{two}, \beta, 20\}, C = \{20, \text{one}, \alpha\}$. Give the sets resulting from

a) $A \cup B$

b) $A \cap B$

c) $C \setminus A$

d) $A \cup (B \cap C)$

2. Define each of the following sets in set builder notation.

a) The interval $-5 < x < 3$ or $-10 < x < -3$.

b) The set of all even numbers.

c) The set of all positive perfect squares (excluding 0).

3. Let q be any real number such that $q \neq 1$. Show, for any $n \in \mathbb{N}$,

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

(Hint: If you've never seen the summation notation, $\sum_{k=1}^{n} q^k = q + q^2 + ... + q^n$.)

4. Prove $5^n + 5 < 5^{n+1}$ for all $n \in \mathbb{N}$.

5. **(Challenge)**. For any statement, $P(n)$ for $n \in \mathbb{N}$, suppose

1) $P(1)$ is true

2) For $n \in \mathbb{N}$, $P(n)$ implies $P(n + 2)$

Prove $P(n)$ is true for every positive odd integer.

Chapter 3

Functions and Relations

3.1 Binary Relations

Before we dive into functions, I'd like to spend some time talking about binary relations, as we will come to learn, functions are just a very special kind of relation.

> **Definition 3.1.1.** *Let* $\Box$ *be a binary relation. Then* $\Box \subseteq X \times Y$.[1] *Then we say* $x \Box y$ [2]*if and only if* $(x, y) \in \Box$.

This definition of a binary relation might seem a bit confusing at first, but after pondering this for a second, this is, in fact, a very natural way to view a binary relation. Though the term binary relation is quite new, you've probably been working with binary relations in math your entire life.

A good example would be the relation $\leq$. For our purposes, I will represent this relation with the symbol $\triangle$, where $a \triangle b$ if and only if $a \leq b$. Let us examine how this works.

First, we need to establish for what sets we are relating. For this example, I will use the natural numbers less than or equal to 3, so defining

$$S := \{n \in \mathbb{N} : n \leq 3\},$$

[2] x is related to y by $\Box$

[3] The $\times$ here is a cartesian product; it means the set of all ordered pairs (x,y) where $x \in X$ and $y \in Y$.

we have $\triangle \subseteq S^2$.[4] Let's fix, for now, $a = 1$. Then we find since

$$1 \triangle 1, \quad 1 \triangle 2, \quad 1 \triangle 3,$$

hence we find *1 is related to 1, 2, and 3*;

$$\{(1,1),(1,2),(1,3)\} \subseteq \triangle.$$

Repeating this exercise, we also find

$$2 \triangle 2, \quad 2 \triangle 3, \quad 3 \triangle 3$$

hence

$$\{(2,2),(2,3),(3,3)\}\triangle$$

where we can conclude by stating

$$\triangle = \{(1,1),(1,2),(1,3),(2,2),(2,3),(3,3).\}$$

Interestingly enough, however, what we have done here is the exact opposite of how modern mathematicians define the greater than and less than operators to create ordered sets in more generality. A discussion of this is too complex for our purposes now, but if anyone is interested, I've left some links in the footnotes, for if anyone wants to learn more.[5]

3.2 Functions

3.2.1 Definition

As I mentioned before, a function is just a type of relation, so in this section, we will explore what a function actually is from a more fundamental perspective.

Definition 3.2.1. *Define relation* $\square \subseteq X \times Y$. *If*

1) for every $x \in X$, *there exists* $y \in Y$ *such that* $(x, y) \in \square$,

[4] $S^2 = S \times S$
[5] https://proofwiki.org/wiki/Definition:Ordering

2) $(x, y) \in \square$ and $(x, z) \in \square$ implies $y = z$,

then we say $\square$ is the graph[6] *that defines the function $f : X \to Y$, where $(x, y) \in \square$ implies $f(x) = y$.*

This definition might be a bit tough to swallow, but before diving into this, we should quickly review some vocabulary for functions. I assume many of you have never defined a function using the notation I've introduced above, so let's break that down first.

The statement $f : X \to Y$ can be read simply as a *function f from X* ***to*** *Y* where X is the **domain** and Y is the **codomain**. You may have also seen a function be referred to as a *mapping*[7] since all a function does is *map* things from one set to another.

Now, with that out of the way, let's look at our conditions. Looking at the first condition requires that every element in X must have a corresponding element in Y. This matches what we should know about functions: that the function should be defined on its entire domain. The next condition says that if $f(a) = b$, and $f(a) = c$, then $b = c$, or in other words, every input can have only one output. You may have formulated this idea with a vertical line test in your previous math classes. But that definition requires the function to have a graphical representation. Since we know that the domain and codomain need not be sets of numbers, there exist functions that have no meaningful graphical representations. What this definition does is that it generalizes that fact to a more general class of functions, even when we can't plot them.

But notice, nowhere in this definition of a function did we specify anything about the codomain of a function. With this, we can conclude that even if an element is in a function's codomain, it need not be related to anything in the domain. This is important since we can define some functions $f(x) = g(x) = x^2$ for $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}^+$[8], where both of these functions are well-defined under our definition of a function. It's easy to tell the graph of the functions are equivalent, and using our prior knowledge, we conclude that the functions are equivalent[9]. But since

[6] Notice here that I'm not referring to the *plot* of a function as you've seen in your previous classes. Instead, I'm talking about the underlying structure that connects inputs to outputs, sort of like a web. From now on, I will refer to the thing we draw as a plot to save some confusion.

[7] or a functional, depending on context

[8] Positive reals

[9] If you are still unsure, just observe that the same inputs relate to the same outputs.

if two functions are equal, we should expect the functions to be defined equivalently, what we get is that $f \not\equiv g$, since the domains of f and g are different.

Definition 3.2.2. *We call $f(A)$ the **image of f under** $A \subseteq X$ if and only if*

$$f(A) := \{f(a) : a \in A\}.$$

This definition might seem very similar to the definition of the range that you might already be familiar with. We will denote the range as $f(X)$ or the image of f under the entire domain. Now, contrast this definition of range with our definition of a codomain. While not every element in a function's codomain needs to be related to an element in the domain, every element in a function's range must be related to an element in its domain. This idea is highlighted in the figure (3.2.1), where the arrows between elements show the structure of the function's graph. Since the element f is unrelated to any element in X, we say f is not in the range.[10]

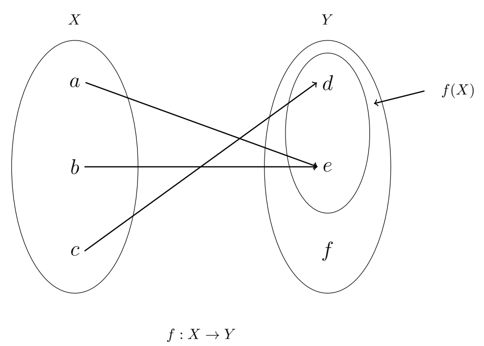


Figure 3.2.1

Definition 3.2.3. *We call $f^{-1}(B)$ the **preimage of f under** $B \subseteq Y$, where*

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

We will not be using this definition that much in precalculus, but it is still a good one to know and will help us with the theorem in a later section. A way we can understand this definition is if that $f^{-1}(B)$ is the set that contains all elements of X that gets mapped to B.[11] We will note

[10]We note that $f(X) \subseteq Y$ for any well-defined function.

[11]Note: this operation *is not* equivalent to the inverse of a function, we will discuss inverses in the next sections.

that $f^{-1}(Y) = X$. I will leave this as an exercise for you to check your understanding of this definition.

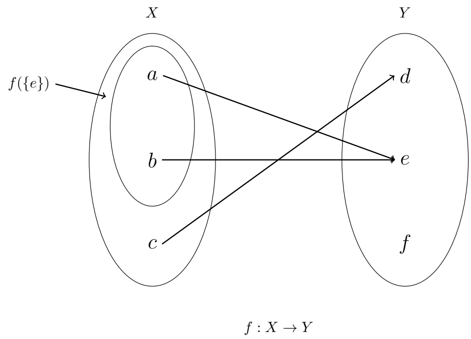


Figure 3.2.2

Example 3.2.1.

In this example, I'd like to demonstrate how to find the maximum possible domain and image for the function

$$f(x) := \sqrt{x},$$

such that

$$X \subseteq \mathbb{R} \quad \text{and} \quad f(X) \subseteq \mathbb{R},$$

where X is the domain of f. The maximum possible domain is just simply the domain for the graph of a function for which we exclude every possible input that causes a function to be ill-defined. Since we know $\sqrt{x}$ doesn't have real outputs for $x < 0$, we let $X = [0, \infty)$. Then, to find the image, with how $\sqrt{x}$ is defined, we only take the positive value of the two theoretically possible outputs hence, we find

$$f(X) = [0, \infty).$$

♠

Example 3.2.2.

In this example, let's find the maximum possible domain for

$$f(x) := \frac{x-1}{x^2+x-2}$$

where X is the domain and

$$X \subseteq \mathbb{R}.$$

Since this is a rational function, the only potential issue here is that f will run into a divide by 0. We can see this is the case when $x = -2$. But this is only one of the possible issues we might run into with f. Therefore, to find the rest, we want to find x such that

$$x^2 + x - 2 = 0$$

hence

$$x^2 + x - 2 = (x+2)(x-1) = 0$$

where we find when $x = 1, -2$, the denominator is zero. Thus we define

$$X = \{x \in \mathbb{R} : x \neq 1, -2\}.$$ [12]

The image is slightly more difficult to find, so we will leave that for a future discussion. ♠

3.2.2 Odd or Even

Definition 3.2.4. *Let $f : D \to \mathbb{R}$[13]. We say f is odd if and only if $f(-x) = -f(x)$. Furthermore, we say f is even if and only if $f(-x) = f(x)$.*

Now, this may seem like a rather arbitrary definition, but rather than worry about the particular wording of this concept, just interpret this as any other mathematical notation or definition as if the wording bears no resemblance to its counterpart when used to refer to numbers.

[12] Notice since $f(x) = \frac{x-1}{(x+2)(x-1)}$, so we might be inclined to say $f(x) = \frac{1}{x+2}$ and conclude $X = \{x \in \mathbb{R} : x \neq -2\}$. But this is indeed incorrect since when we divided out the $x - 1$, we inadvertently divided by zero since we never excluded $x = 1$ from our domain.

[13] From now on, we will use D to signify a domain in $\mathbb{R}$.

Theorem 3.2.1. *Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ be even functions. Then the following are true:*

1) $f + g$ *is even,*

2) $f \cdot g$ *is even,*

3) $f - g$ *is even,*

4) $\frac{f}{g}$ *is even, for* $g(x) \neq 0$ *for all* $x \in D$*.*

Proof. Since

$$(f + g)(-x) = f(-x) + g(-x) = f(x) + g(x) = (f + g)(x)$$

hence, the function $f + g$ is even, therefore proving (1).

Then, using the same technique,

$$(f \cdot g)(-x) = f(-x) \cdot g(-x) = f(x) \cdot g(x) = (f \cdot g)(x)$$

hence proving (2).

Let $h(x) = -1$. Then observe $h(-x) = h(x) = -1$, hence h is even. Then

$$(f - g)(x) = f(x) + h(x)g(x)$$

Since $h \cdot g$ is even, we conclude that $f - g$ is also even, hence proving (3).

Then since

$$\left(\frac{f}{g}\right)(x) = \frac{f(-x)}{g(-x)} = \frac{f(x)}{g(x)},$$

we can conclude $\frac{f}{g}$ is even. □

Theorem 3.2.2. *Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ be even and odd functions, respectively. Then*

1) $f \cdot g$ *is odd*

2) $\frac{f}{g}$ *is odd, for* $g(x) \neq 0$ *for all* $x \in D$*.*

Proof. The proof works similarly to how we proved theorem (3.2.1). Since

$$(f \cdot g)(-x) = f(-x)g(-x) = f(x) \cdot (-g(x)) = -(f \cdot g)(x)$$

hence proving $f \cdot g$ is odd. Then since

$$\left(\frac{f}{g}\right)(-x) = \frac{f(-x)}{g(-x)} = \frac{f(x)}{-g(x)} = -\left(\frac{f}{g}\right)(x)$$

Thereby proving the theorem. □

We can actually reverse the functions in theorem (3.2.2) and get the same conclusion (the proof follows exactly as above). Notice, in theorem, (3.2.2), we didn't mention anything about when these functions were added. This is because the odd or evenness of a function cannot be determined if nothing else is known about the function. Nevertheless, something interesting will happen, as we will explore soon.

Theorem 3.2.3. *Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ be odd functions. Then the following are true:*

1) $f + g$ is odd,

2) $f \cdot g$ is even,

3) $f - g$ is odd,

4) $\frac{f}{g}$ is even, for $g(x) \neq 0$ for all $x \in D$.

Proof. Exercise. □

Theorem 3.2.4. *Let $f : D \to \mathbb{R}$ be any arbitrary function. Then there exists $g : D \to \mathbb{R}$ and $h : D \to \mathbb{R}$, even and odd respectively such that*

$$f = g + h$$

Proof. Define

$$g(x) = \frac{f(x) + f(-x)}{2}$$

$$h(x) = \frac{f(x) - f(-x)}{2}$$

Then the sum $g + h$ is trivially f. To prove g is even:

$$g(-x) = \frac{f(-x) + f(-(-x))}{2} = \frac{f(-x) + f(x))}{2} = g(x).$$

And h is odd:

$$h(-x) = \frac{f(-x) - f(-(-x))}{2} = \frac{f(-x) - f(x))}{2} = -\frac{f(x) - f(-x)}{2} = -h(x).$$

Since our definitions of g and h are even and odd, respectively, this concludes our proof of the theorem. □

We will conclude our discussion of odd and even functions in this lesson and pick up on their applications in future lessons. For now, just remember the definition of this idea and its basic properties.

3.2.3 Compositions and Inverses

Definition 3.2.5. *Let $f : X \to Y$ be a function. f is said to be **one-to-one (injective)** if for every $x_1, x_2 \in X$ such that $f(x_1) = f(x_2)$ implies $x_1 = x_2$ or equivalently, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$.*

Definition 3.2.6. *Let $f : X \to Y$ be a function. f is said to be **onto** Y **(surjective)** if for every $y \in Y$, there exists $x \in X$ such that $f(x) = y$.*

Definition 3.2.7. *Let $f : X \to Y$ be one-to-one and onto Y. Then f is said to be a **bijection**.*

The mathematical jargon may be a bit tough to comprehend, but let's break each down individually.

A one-to-one function, simply put, is a function where there can be a maximum of **one** input that maps to each output. As in figure (3.2.3), each output has a maximum of one connection. A common way to visualize a real one-to-one function on a plot is that a one-to-one function passes a horizontal line test. Like the vertical line test, if for every horizontal line, the line only passes through the plot at most once, then the function is one-to-one. If the function wasn't defined on a number line, the horizontal

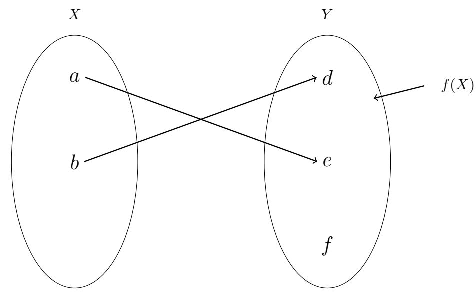


Figure 3.2.3: one-to-one function

line test wouldn't apply, but it's still a useful test for when we are getting still getting used to these ideas.

An onto function is a function where the image is equal to the codomain. As you can see in the figure (3.2.4), every element in Y has at least one corresponding element in X that maps to it.

Then, when you put these together, the resulting function passes the horizontal line test and maps to every element in Y, as in figure (3.2.5), we get a bijection.

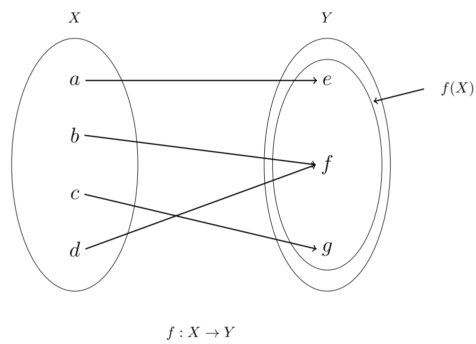


Figure 3.2.4

With this newfound vocabulary, we are one step closer to talking about inverses, but before we do, we must talk about functional composition. When we compose functions, what we do is take the output from one function and put it into another function. Because of some of the properties of this operation, the symbol we use for this looks very similar to multiplication.[14]

[14] If you ever take linear algebra, functional composition distributes on addition very similarly to how multiplication does. It also has an interesting connection to matrix

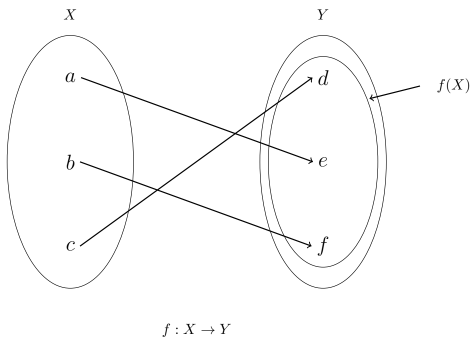


Figure 3.2.5

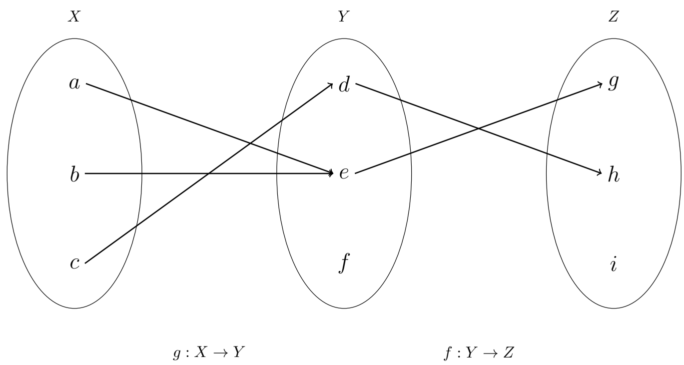


Figure 3.2.6

Definition 3.2.8. *Define* $f : Y \to Z$ *and* $g : X \to Y$*. The composition of* f *and* g*, notated as* $f \circ g$ *is a function* $f \circ g : X \to Z$ *such that* $(f \circ g)(x) = f(g(x))$*, or as seen in figure* (3.2.6)*.*

An important property that's useful later is that functional composition is associative, or $f \circ (g \circ h) = (f \circ g) \circ h$. The proof is fairly trivial, so we ignore it here.

Definition 3.2.9. *Define* $f : X \to Y$*. If there exists* $g : Y \to X$ *such that* $g \circ f = I_X$*, then* g *is said to be the left inverse (retraction) of* f*.*

I_X in definition (3.2.9) is simpily the identity function on domain X, where $I_X(x) = x$. Like how the symbol 1 works with multiplication, any composition with the identity function results in the same function (e.g.,

multiplication, which is where I assume we get the symbol from (I'm not completely sure).

$f \circ I = I \circ f = f$).

Definition 3.2.10. *Define $f : X \to Y$. If there exists $g : Y \to X$ such that $f \circ g = I_Y$, then g is said to be the right inverse (section) of f.*

The left inverse is illustrated by figure (3.2.7). As you can see, a left inverse, essentially *retracts* the action of f and sends everything back to where it came from.

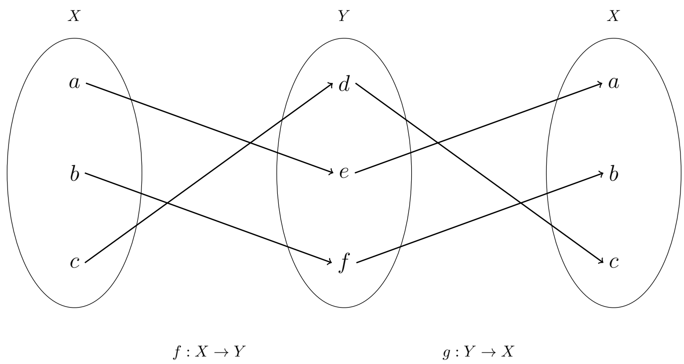


Figure 3.2.7

The right inverse, as seen in figure (3.2.8), can be thought of as a function that takes a *section* of the domain of f and reverses it.

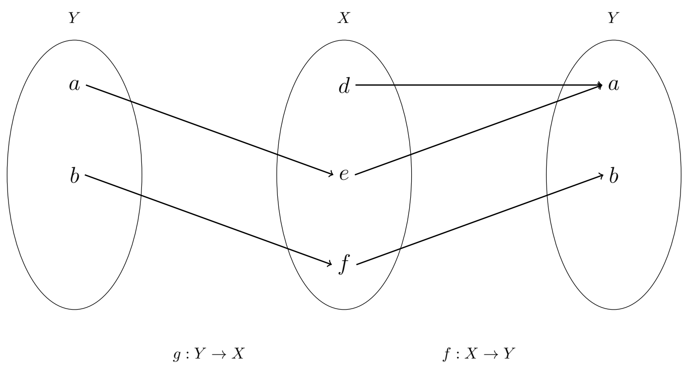


Figure 3.2.8

Now that we've characterized these basic types of inverses let's start understanding these functions more with the following theorems.

Theorem 3.2.5. *Define $f : X \to Y$. If f has a left inverse, then f is one-to-one.*

Proof. For any $x_1 x_2 \in X$ such that $f(x_1) = f(x_2)$, let $g : Y \to X$ be the left inverse of f. Therefore, by the definition of the function, we have

$$g(f(x_1)) = g(f(x_2)).$$

But since $g \circ f = I_x$, we deduce

$$x_1 = x_2$$

hence f is one-to-one. □

Theorem 3.2.6. *Define $f : X \to Y$. If f has a right inverse, then f is a surjection.*[15]

Proof. Exercise. □

Definition 3.2.11. *Define $f : X \to Y$. If f has left inverse $g : Y \to X$ and right inverse $h : y \to X$, then f is said to have a two-sided inverse.*

$$f^{-1} = g = h$$

From here on out, I will use the term inverse for a two-sided inverse and specify when I only mean a left or right inverse.

Theorem 3.2.7. *Suppose $f : X \to Y$ has an inverse, then f is a bijection.*[16]

Proof. By theorems (3.2.5) and (3.2.6), f is both one-to-one and onto, concluding that f is a bijection. □

[15] The converse of theorems (3.2.5) (3.2.6) is also true, but proving them requires explicitly constructing a left or right inverse without making explicit assumptions about the underlying function, hence a bit tricky.

[16] Using the converse of theorem (3.2.5) (3.2.6), we can also get the converse of this theorem.l

Now, from theorem (3.2.7), the function being a bijection is a necessary condition for the existence of an inverse. This means, that inverses only exist for functions if they are bijections, which might seem to contradict what you know about function inverses.[17] While this might be true by the strict definition of an inverse, there are ways to build meaningful inverses by changing the function definition just a bit. If this doesn't make sense to you, I will show in just a moment how this all makes sense.

Definition 3.2.12. *Let $f : X \to Y$. Then we define $f|_A : A \to Y$, where $f|_A(a) := f(a)$, to be the* ***restriction of** f **on** A*

Definition 3.2.13. *Let $f : X \to Y$. Then if $f|^B : X \to B$, where $f|^B(x) := f(x)$, is a well-defined function, then we call $f|^B$ the* ***corestriction of** f **on** B.*

We note that whilst the restriction of a function always exists, so long as A is non-empty, the corestriction might not always be a well-defined function. So, let's examine what guarantees the existence of $f|^B$.

Lemma 3.2.1. *Suppose $f : X \to Y$. $f|^B$ is well-defined if and only if $f(X) \subseteq B$.*

Proof. Suppose $f|^B$ is well-defined and let $y \in f(X)$. Then, by the definition of an image, there exists $x \in X$ such that $f(x) = y$. Then, by the definition of $f|^B$,

$$f(x) = f|^B = y.$$

Since $f|^B$ is well-defined, we have $y \in B$ implying $f(X) \subseteq B$.

Then suppose $f(X) \subseteq B$. Then, we prove $f|^B$ is a well-defined function by checking with definition (3.2.1). First, we check that $f|^B$ satisfies (1). This is done by noticing that for every $x \in X$, we have

$$f|^B(x) = f(x) \in f(X) \subseteq B,$$

hence $f|^B$ satisfies (1).

To show $f|^B$ satisfies (2), we notice if $x_1, x_1 \in X x_1 = x_2$, we have

$$f|^B(x_1) := f(x_1) = f|^B(x_2) := f(x_2).$$

[17]think of $\sin^{-1}(x)$ or $\cos^{-1}(x)$

Since f is a well-defined function, we have $x_1 = x_2$. □

Lemma 3.2.2. *For any* $f : X \to Y$*, there exists* $f|^B$*,* $B \supseteq f(X)$*, where* $f|^B$ *is a onto* B*.*

Proof. Let $B = f(X)$. By lemma (3.2.1), $f|^B$ is well-defined; hence it remains to show that $f|^B$ is onto B. By the definition of the image, for every $y \in B$, there exists $x \in X$ such that $f(x) = f|^B(x) = y$. Therefore, this suffices to prove $f|^B$ is onto B. □

Lemma 3.2.3. *For any* $f : X \to Y$*, there exists* $f|_A$*,* $A \subseteq X$ *where* $f|_A$ *is one-to-one.*

Proof. Define $C : P(X) \to X$, where $P(X)$ is the power set of X[18] and for every $D \in P(X)$, $C(D) \in D$. We don't need to worry about the details of how C is defined. We just care that such a function exists and is well-defined, which is guaranteed by the axiom of choice.[19] Let

$$S := \{f^{-1}(\{y\}) : y \in Y\},$$

where $S \subseteq P(X)$ since $f^{-1}(\{y\}) \subseteq X$ for each $y \in Y$ by definition. Then let $A := C(S)$. We wish to show this definition of A makes $f|_A$ one-to-one. Suppose $x_1, x_1 \in A$ such that

$$f(x_1) = f(x_2).$$

Since $x_1, x_2 \in A$ hence $x_1, x_2 \in C(S)$, there exists $y_1, y_2 \in Y$ such that

$$C(f^{-1}(\{y_1\}) = x_1 \quad \text{and} \quad C(f^{-1}(\{y_2\}) = x_2$$

Then, by definition of C, we have

$$x_1 \in f^{-1}(\{y_1\}) \quad \text{and} \quad x_2 \in f^{-1}(\{y_2\}).$$

[18] or the set of all subsets of X

[19] C here is what we call a choice function. You can think of it as a map from a subset of X to a corresponding element, or choice, within that subset. It's okay if you don't understand this and the axiom of choice. The only thing that's important here is that C is a well-defined function, and it exists if we choose to believe in the axiom's validity

Therefore

$$f(x_1) \in \{y_1\} \quad \text{and} \quad f(x_2) \in \{y_2\}$$

hence

$$f(x_1) = y_1 \quad \text{and} \quad f(x_2) = y_2.$$

By assumption of $f(x_1) = f(x_2)$, we have $y_1 = y_2$, hence

$$f^{-1}(\{y_1\}) = f^{-1}(\{y_2\}).$$

Then, since C is well-defined, we have

$$C(f^{-1}(\{y_1\}) = C(f^{-1}(\{y_2\})$$

hence $x_1 = x_2$ therefore, $f|_A$ is one-to-one. □

Theorem 3.2.8. *Suppose $f : X \to Y$. Then there exists $A \subseteq X$ and $B \supseteq f(X)$ such that $f|_A^B$ is a bijection.*

Proof. By lemmas (3.2.2) and (3.2.3), there exists A and B such that, $f|_A^B$ is one-to-one and onto B, hence $f|_A^B$ is a bijection. □

Now, with theorem (3.2.8), we now know we can transform any function into a bijection, which has an inverse. With this, let's try this in an example.

Example 3.2.3.

Let's try to construct a meaningful inverse for the function

$$f(x) = \sqrt{x},$$

where $f : [0, \infty) \to \mathbb{R}$. By imagining the function's plot in our heads, f is one-to-one but isn't onto. Since we know f doesn't output anything less than 0, we can construct a new function

$$g(x) = f(x)$$

where $g : [0, \infty) \to [0, \infty)$. Now checking, g is clearly onto since the range is exactly equal to the codomain, our new function g is a bijection;

therefore, there exists g^{-1} such that

$$g^{-1} \circ g = g \circ g^{-1} = I_{[0,\infty)}$$

Doing some simple algebra, we get

$$g^{-1}(x) = x^2.$$ [20]

♠

Example 3.2.4.

In this next example, let's try to determine the domain and codomain of $\sin^{-1}(x)$. By imagining the plot of $f(x) = \sin(x)$, we conclude that f is not one-to-one nor onto. We first solve the "onto" problem by setting our new function's codomain to $f(X)$, which in this case is $[-1, 1]$. Like in the proof of theorem (3.2.3), there is a choice involved here, so unlike in the previous example, answers can vary depending on what we'd like to achieve. For our purposes, we will cut the number line, leaving the interval $(-\frac{\pi}{2}, \frac{\pi}{2}]$. We can check this is the case by plotting f on this domain, and f clearly passes the horizontal line test. Hence we define $g : (-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$, where

$$g(x) = \sin(x)$$

and

$$g^{-1}(x) = \sin^{-1}(x)$$

where $g^{-1} : [-1, 1] \to (-\frac{\pi}{2}, \frac{\pi}{2}]$. ♠

Example 3.2.5.

In this example, we'd like to find a meaningful inverse for

$$f(x) = \frac{x}{1 + x}$$

on the largest possible domain X such that f is well-defined and a bijection. We conclude our largest possible domain for this function is

[20] Even though it's trivial, check this is indeed the case by computing $g^{-1} \circ g$ and $g \circ g^{-1}$.

$X = \{x \in \mathbb{R} : x \neq -1\}$. We do not yet have the tools to compute the image of f under this domain directly so we will let R represent this set hence

$$f : \mathbb{R} \setminus \{-1\} \to R.$$

Then, let us solve for the graph of f^{-1} algebriclly. Let $y = f(x)$ for any x. Then

$$y = \frac{x}{1+x}$$

which by eliminating the dominator, we get

$$y + xy = x$$

hence

$$y = x - xy$$

$$y = x(1-y)$$

$$x = \frac{y}{1-y}$$

Then letting $f^{-1}(y) = x$, we conjecture the graph of f^{-1} is

$$f^{-1}(x) = \frac{x}{1-x}.$$ [21]

Being the inverse of f, we have $f^{-1} : R \to \mathbb{R} \setminus \{-1\}$. Since the maximum possible domain of f^{-1} is $\mathbb{R} \setminus \{1\}$, we also conjecture

$$R = \mathbb{R} \setminus \{1\}.$$

First, we can check that f is well defined with respect to our definition of R by checking $f(x) \neq 1$ for any x in our domain. To do this by way of contradiction. Thus if

$$f(x) = \frac{x}{1+x} = 1,$$

then

$$x = 1 + x$$

implying $0 = 1$ which is a contradiction. Therefore, our definition of f is well-defined.

It remains to check that f^{-1} is well-defined with respect to our codomain,

[21] I just switch x and y, for if anyone is confused. It just looks better this way. If you'd like, just substitute in y to get $f^{-1}(y)$.

that this is indeed the inverse of f, and that f is a bijection. We check the first claim by checking that f^{-1} doesn't output -1. Again, we proceed by contradiction hence, if there exists x such that

$$f^{-1}(x) = \frac{x}{1-x} = -1,$$

then

$$x = x - 1$$

implying $0 = -1$ which is a contradiction. We check the remaining claims by computation. Since

$$f(f^{-1}(x)) = \frac{f^{-1}(x)}{1 + f^{-1}(x)} = \frac{\frac{x}{1-x}}{1 + \frac{x}{1-x}}$$

$$= \frac{x}{(1-x) + x} = x = I$$

and

$$f^{-1}(f(x)) = \frac{f(x)}{1 - f(x)} = \frac{\frac{x}{1+x}}{1 - \frac{x}{1+x}}$$

$$= \frac{x}{(1+x) - x} = x = I$$

we conclude that f^{-1} indeed is an inverse, hence as a necessary condition, f is a bijection. Therefore, the range of f is exactly R implying the range of f is exactly $\mathbb{R} \setminus \{1\}$. ♠

3.3 Exercies

1. Prove theorem (3.2.3).

2. Prove theorem (3.2.6).

3. Determine, with proof whether the following functions are odd, even, or neither

   a) $f(x) = \sqrt{x^4 + x^2} + 4$.

   b) $g(x) = x\sqrt{x^2 + 1}$.

4. Find two function f, g, even and odd respectively such that

$$f + g = x^2 + 2^x.$$

5. For the following functions, determine the maximum possible domain without a calculator. Then, when possible, construct an inverse without changing the domain of each function (you may need a graphing calculator for this second part).

a) $\frac{1}{\sqrt{1+x^2}}$.

b) $\frac{x}{x^3+1}$.

c) $\frac{2}{\sqrt{x-1}}$

Chapter 4

Polynomials

4.1 Quadratics

This section will be mostly review, for those who may feel less comfortable with factoring and solving quadratics. Those who feel comfortable with these skills should feel free to jump to section two, where we will expand on these ideas.

4.1.1 Factoring

In this first section, we are going to focus on factoring quadratics. When we factor, what we want is to break down the polynomial into binomials of the form $x + p$ or constants. As you know from previous math classes, if we work only on real numbers, then not every polynomial is completely factorable, so in general, we will try our best to find as many factors as we can. We will expand on the idea of factorability in a later section.

Now in the case of quadratic expressions or polynomials of order 2, if the polynomial is factorable, we should expect to be able to write our factored quadratic in the form

$$a(x + p)(x + q).$$

We should expect exactly two factors since, if there's any less than when expanding, we would not get the x^2 term, and if there's any more, then we would get a term of order higher than 2.[1] Expanding the previous expression gives:

$$a(x^2 + [p + q]x + pq)$$

Then in the case $a = 1$, to factor a given polynomial, we require $p + q$ to be the coefficient for x and pq to be the constant.

Example 4.1.1.

In this example, we'd like to factor the expression

$$x^2 + 3x - 4$$

Now since we require the two factors of -4 to add up to 3. Since pairs of factors of -4 must be of opposite sign, we additionally require the negative factor to be the smaller of the two. Now listing out factors of 4:

$$4 \cdot 1 = 4$$

$$2 \cdot 2 = 4$$

Therefore, since $4 - 1 = 3$, we let $p = 4$ and $q = -1$ hence

$$x^2 + 3x - 4 = (x + 4)(x - 1)$$

♠

Example 4.1.2.

In this example, we'd like to factor the expression

$$x^2 - 6x + 8$$

Since in this case, our constant is positive and our x coefficient is negative, we should expect both p and q to be negative. Then listing our factors of 8, we get

$$8 \cdot 1$$

$$4 \cdot 2$$

[1] Check this!

Therefore, since $(-4) + (-2) = -6$, we the quadratic in factored form is

$$(x - 4)(x - 2)$$

♠

Now suppose $a \neq 1$, instead of leaving the a outside the parenthesis, we will rewrite our factored form as

$$(ax + p)(bx + q).$$ [2]

Therefore expanding the previous expression gives

$$abx^2 + (aq + bp)x + pq.$$

It turns out this form is particularly useful for factoring since

$$(abx^2 + aqx) + (bpx + pq) = ax(bx + q) + p(bx + q)$$

$$= (ax + p)(x + q).$$

Thus, it might be smart to transform the expression to this form to begin factoring. We can sum up our observations with the following: to factor a polynomial with a non-zero leading coefficient, we should try to find two values aq, and bp, that add to our middle coefficient and who is the product of our leading coefficient and the constant term. Let's try this in the following example.

Example 4.1.3.

In this example, we'd like to factor the expression

$$2x^2 + 7x + 3$$

We notice since $2 \cdot 3 = 6$, we'd like to split the middle term into two terms with coefficient with the product 6. Running through all possible factors of 6, we get

$$2x^2 + 6x + x + 3$$

[2] This looks very different from our previous expression, but if we pull out a and b from our two factors, our a, b, p, q values are arbitrary hence, by making a few substitutions, we would arrive at the same expression only with different letters.

We factor and get

$$2x(x+3)+(x+3)=(2x+1)(x+3)$$

♠

Example 4.1.4.

In this second example, we'd like to factor the expression

$$6x^2+5x-6$$

We notice $6\cdot -6=-36$, so we like to find two numbers with the same product that sum to 5. Running through all the possibilities we notice since

$$9\cdot -4=-36 \quad \text{and} \quad 9-4=5$$

we get

$$6x^2+9x-4x+6$$

Then we factor and get

$$=3x(2x+3)-2(2x+3)=(3x-2)(2x+3)$$

♠

4.1.2 Solving quadartics

Typically, when we say to solve a quadratic, what we want is to find the roots of the expression, or in other words, what value of x makes the expression give a value of 0? Let's try this as an example.

Example 4.1.5.

Using a polynomial which we factored from above, let's find the roots of

$$x^2-6x+8$$

Factoring the expression gives

$$(x-4)(x-2)$$

Since we want to find when this expression equals zero, we require

$$(x-4)(x-2)=0$$

Notice here, that we cannot just divide a factor out, since we could be inadvertently dividing by zero. But what we can say is if the entire expression is 0, thus there must be a factor that is zero.[3] Hence either

$$x-4=0 \quad \text{or} \quad x-2=0$$

therefore $x = \{4, 2\}$. ♠

Example 4.1.6.

In this example, let's find the roots of the expression

$$2x^2+7x+3$$

As in example (4.1.3), this factors to

$$(2x+1)(x+3)$$

Removing the 2 from inside the parenthesis, we get

$$2(x+\frac{1}{2})(x+3)$$

Then since we are finding the roots, we require

$$2(x+\frac{1}{2})(x+3)=0$$

Here, we can divide out the 2 since $2 \neq 0$. Actually, in general, we can always divide that coefficient, since it cannot be zero, or our entire

[3] We should give a proof of this fact with some basic properties of the real numbers. This will be left to the reader.

expression is identically zero. Hence, we require

$$(x + \frac{1}{2})(x + 3) = 0$$

and using the same idea as before, we have

$$x = \{\frac{1}{2}, -3\}$$

♠

4.2 Polynomials of Higher Order

In this case, we wish to expand on the ideas established in the previous section for polynomials of order 2 or higher. Now, unfortunately, this is going to get a little messy, and we will soon find there aren't always established methods for finding these factors, and most of the time we will have to resort to guessing and checking. While things like the cubic formula and quartic formula (for orders 3 and 4, respectively) exist, it has been mathematically proven that a quartic formula (order 5) cannot exist.[4] Therefore, to find these roots, we must resort to numerical methods, or in other words, strategic guessing. Unfortunately, since our brains can't work as fast as a computer, we will have to resort to a more primitive method of strategic guessing, but rest assured, with some practice, you will get the hang of this.

4.2.1 Polynomial Long Division

Let's start this section with some remarks about notation. We will let $\mathbb{R}[x]$ denote polynomials with real coefficients and $\mathbb{C}[x]$ denote polynomials with complex coefficients. Then define the function $\deg : \mathbb{R}[x] \to \mathbb{N} \cup \{0\}$ such that for any $p \in \mathbb{R}[x]$, $\deg(p)$ gives the degree of p. Hence $\deg(x^2+x+1) = 2$ $\deg(2) = 0$ ...

With this, let's lead off the section with the following theorem.

Theorem 4.2.1. *Let $p, d \in \mathbb{R}[x]$ such that $d \neq 0$. Then there exists*

[4] This is done in what is known as Galois theory.

unique $q, r \in \mathbb{R}[x]$ *such that* $r = 0$ *or* $\deg(r) < \deg(d)$ *and*

$$p = dq + r$$ [5]

Proof. I will only prove the uniqueness part here.[6] Assuming the existence part of the theorem, suppose for some $p, d \in \mathbb{R}[x]$,

$$p = dq_1 + r_1 \quad \text{and} \quad p = dq_2 + r_2.$$

Then, taking the difference here gives

$$dq_1 + r_1 - (dq_2 + r_2) = d(q_1 - q_2) + (r_1 - r_2) = 0$$

therefore

$$r_2 - r_1 = d(q_1 - q_2).$$

If $q_1 - q_2 \neq 0$, then

$$\deg(d(q_1 - q_2)) = \deg(d) + \deg(q_1 - q_2) \geq \deg(d). \tag{4.2.1}$$

Since $\deg(r_1), \deg(r_2) < \deg(d)$ by definition, we have

$$\deg(r_2 - r_1) \leq \max\{\deg(r_1), \deg(r_2)\} < \deg(d).$$

Therefore, by equation (4.2.1), we also require

$$\deg(r_2 - r_1) = \deg(d(q_1 - q_2)) \geq \deg(d)$$

Since no non-negative integer satisfies the relation

$$\deg(d) \leq \deg(r_2 - r_1) < \deg(d),$$

we require $r_2 - r_1 = 0$. Since $d \neq 0$, we also require $q_1 - q_2 = 0$. Therefore,

$$r_1 = r_2 \quad \text{and} \quad q_1 = q_2$$

completing the proof.[7] □

[5] This theorem remains true if we replace all real polynomials with complex ones.

[6] The existence part requires a confusing and lengthy induction argument, hence I've left it out here.

[7] I used two identities involving the deg function namely $\deg(a - b) \leq \max\{\deg(b), \deg(b)\}$ and $\deg(ab) = \deg(a) + \deg(b)$ which I didn't give a proof for. These proofs aren't difficult hence, I've left them as an exercise for the reader.

This is actually a result that you've seen with integers. Therefore, like integers, we also have long division for polynomials. We will not prove that this method works in general for finding the q and r as prescribed by theorem (4.2.1). Now it might have been some time since the last time many of you have done long division, but let's try a few examples to jog your memory.

Example 4.2.1.

In this example, let's try to divide $x^2 + 2x - 7$ by $x - 2$.

$$x - 2 \,|\overline{x^2 + 2x - 7}$$

To begin, we wish to eliminate the leading term. To do so, we must multiply the $x - 2$ by x, hence

$$\begin{array}{rl} & x \\ x - 2 & |\overline{x^2 + 2x - 7} \\ & x^2 - 2x \end{array}$$

Then, like we did in elementary school, we subtract, hence

$$\begin{array}{rl} & x \\ x - 2 & |\overline{x^2 + 2x - 7} \\ & \underline{x^2 - 2x} \\ & 4x - 7 \end{array}$$

Then repeating the same process, we get

$$\begin{array}{rl} & x + 4 \\ x - 2 & |\overline{x^2 + 2x - 7} \\ & \underline{x^2 - 2x} \\ & 4x - 7 \\ & \underline{4x - 8} \\ & 1 \end{array}$$

1 here is a reminder, since we can't divide that any further by $x - 2$, so we can write

$$x^2 + 2x - 7 = (x - 2)(x + 4) + 1.$$

♠

Example 4.2.2.

Let's divide $6x^4 - 9x^2 + 3x + 6$ by $x^2 - 2$.

$$\begin{array}{r|l} & 6x^2 \\ \hline x^2 - 2 & 6x^4 + 0x^3 - 9x^2 + 3x + 6 \\ & \underline{6x^4 - 12x^2} \\ & 3x^2 + 3x + 6 \end{array}$$

$$\begin{array}{r|l} & 6x^2 + 3 \\ \hline x^2 - 2 & 6x^4 + 0x^3 - 9x^2 + 3x + 6 \\ & \underline{6x^4 - 12x^2} \\ & 3x^2 + 3x + 6 \\ & \underline{3x^2 - 6} \\ & 3x + 12 \end{array}$$

hence

$$6x^4 - 9x^2 + 3x + 6 = (x^2 - 2)(6x^2 + 3) + 3x + 12$$[8]

♠

4.2.2 Roots of Higher Order Polynomials

Now from before, you might have noticed that the roots of polynomials line up quite nicely with factors of a polynomial. We can summarize this observation with the following theorem:

[8]Whilst we don't run into an issue here when a term like x^3 is missing, we mustn't forget the polynomial is $6x^4 + 0x^3 - 9x^2 + 3x + 6$. It didn't matter here, but it could in the future.

Theorem 4.2.2. *Suppose f is a polynomial (real or complex), then $x - p$ is a factor if and only if $f(p) = 0$.*

Proof. For this proof, we most show that if $x-p$ is a factor, then $f(p) = 0$ and $f(p) = 0$ implies $x - p$ is a factor of f.

To prove the first statement, since $x - p$ is a factor of f, there exists $g(x)$ such that

$$f(x) = (x - p)g(x)$$

Therefore, $f(p) = 0$.

Then to prove the second statement, using theorem (4.2.1), we have some q, r such that

$$f(x) = q(x)(x - p) + r(x).$$

Since we assume p is a root of this polynomial, we have

$$f(p) = q(p)(p - p) + r(p) = r(p) = 0$$

Since $\deg(r) < \deg(q) = 1$ or $r = 0$, either r is a constant or zero. Thus by $r(p) = 0$, we conclude $r \equiv 0$ hence

$$f(x) = q(x)(x - p).$$

□

Armed with polynomial long division and theorem (4.2.2), if we can show $f(p) = 0$ for a polynomial f, then $x - p$ is a factor of f. Let's try this as an example.

Example 4.2.3.

Let's try to factor the expression

$$x^3 - 6x^2 - x + 30.$$

First, we need to try to find a factor by guessing and checking. We can do this by plugging in what we think will be a factor in our polynomial. To make our lives easier, we will assume our polynomial only has integer roots. This is not always the case in general, but making our assumption allows us not to waste too much time trying random fractions. Here I will skip all of the guessing and try -2. Since for $x = -2$[9]

$$(-2)^3 - 6(-2)^2 - (-2) + 30 = 0$$

hence, $x + 2$ must be a factor. I will skip showing the long division here since it's not easy to type out, but regardless, we should find

$$x^3 - 6x^2 - x + 30 = (x + 2)(x^2 - 8x + 15).$$

Then factoring the resulting quadratic we find

$$x^2 - 8x + 15 = (x - 5)(x - 3)$$

hence we can write our cubic as

$$(x + 2)(x - 5)(x - 3)$$

with roots

$$x = \{-2, 5, 3\}$$

♠

4.3 Complex Numbers

In this section, we will discuss introduce the notion of complex numbers. The most important symbol when talking about complex numbers is i which is defined as

$$i := \sqrt{-1}.$$

In this section, we will only introduce this idea and we will go into the detail at another time. Let's first give a brief definition of complex numbers.

Definition 4.3.1. *Let $\mathbb{C} := \{a + ib : a, b \in \mathbb{R}\}$. Then $z \in \mathbb{C}$ is said to be a complex number.*

Notice because of the way we define complex numbers, the set of all complex numbers isn't ordered, so there isn't a notion of bigness or small-

[9]Since all of our roots are assumed to be integers, it follows that every factor must divide the constant term. Try to see if you can show why this is the case.

ness of complex numbers. Even still, there is a notion of distance, but before we discuss that, let's discuss one of the most important operations of a complex number: the conjugate.

Definition 4.3.2. *Let $z \in \mathbb{C}$. If $a, b \in \mathbb{R}$ such that $z = a + ib$, then the conjegate $\bar{z} := a - ib$.*

Keeping this in mind, let's prove some basic properties of the conjugate.

Theorem 4.3.1. *Let $z_1, z_2 \in \mathbb{C}$. Then the following are true:*

1) $\overline{(z_1 + z_2)} = \bar{z_1} + \bar{z_2}$

2) $\overline{(z_1 \cdot z_2)} = \bar{z_1} \cdot \bar{z_2}$

Proof. Let

$$z_1 = a_1 + ib_1 \quad \text{and} \quad z_2 = a_2 + ib_2$$

Then

$$\bar{z_1} + \bar{z_2} = (a_1 - ib_1) + (a_2 - ib_2) = (a_1 + a_2) - i(b_1 + b_2)$$

$$= \overline{(z_1 + z_2)}$$

since $\overline{(z_1 + z_2)} = (a_1 + a_2) - i(b_1 + b_2)$, this proves (1).

Then to prove (2):

$$\bar{z_1} \cdot \bar{z_2} = (a_1 - ib_1)(a_2 - ib_2) = a_1a_2 - ia_2b_1 - ia_1b_2 - b_1b_2$$

$$= (a_1a_2 - b_1b_2) - i(a_1b_2 + a_2b_1).$$

Then since

$$\overline{(z_1 \cdot z_2)} = \overline{(a_1 + ib_1)(a_2 + ib_2)}$$

$$= \overline{[a_1a_2 + ia_2b_1 + ia_1b_2 - b_1b_2]} = \overline{[(a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1)]}$$

$$= (a_1a_2 - b_1b_2) - i(a_1b_2 + a_2b_1)$$

we complete the proof. □

Corollary 4.3.1. *Let $z \in \mathbb{C}$. Then $\overline{(z^n)} = (\bar{z})^n$, for any $n \in \mathbb{Z}$.*[10]

Proof. Since by definition of the exponential,

$$\overline{z^n} = \underbrace{\overline{z \cdot z \cdot \ldots \cdot z}}_{n}$$

Then by theorem (4.3.1),

$$= \overline{z} \cdot \overline{z} \cdot \ldots \cdot \overline{z}$$

$$= (\overline{z})^n$$

□

Theorem 4.3.2. *Let $z \in \mathbb{C}$. Then $z\bar{z} \in [0, \infty)$.*

Proof. Let $a, b \in \mathbb{R}$ such that $z = a + ib$. Then

$$z\bar{z} = (a + ib)(a - ib) = a^2 + abi - abi - b^2 = a^2 - b^2$$

Multiplication is closed under multiplication and subtraction, we conclude $a^2 + b^2 \in \mathbb{R}$. Then since $a^2 + b^2 \geq 0$, this implies $a^2 + b^2 \in [0, \infty)$ hence proving the theorm. □

Theorem 4.3.3. *Let $z \in \mathbb{C}$. Then $z + \bar{z} \in \mathbb{R}$*

Proof. Since $z = a + ib$, for $a, b \in \mathbb{R}$,

$$z + \bar{z} = a + ib + a - ib = 2a$$

Since $2a \in \mathbb{R}$, $z + \bar{z} \in \mathbb{R}$ hence proving the theorem. □

4.4 Algebric Completeness

4.4.1 Fundemental Theorem of Algebra

As I noted above, not every polynomial can be factored completely, but why isn't this the case? The answer to this question lies in the fact that the real numbers are not *algebraicly complete*. A big advantage of using

[10]Notice, we only prove this theorem for integer powers, when in fact, this corollary holds for any n, even complex. But for this, we would need a more generalized definition of the exponential, and even still, we would first need to prove it for integer powers before we can prove the general case.

complex numbers as opposed to real numbers is that complex numbers are algebraically complete, or in other words, every polynomial of degree n can be completely factored into a series of n binomials of degree 1. We illustrate this in the following theorem.

Theorem 4.4.1 (Fundemental Theorem of Algebra[11]). *Let $p_n(z) \in \mathbb{C}[z]$[12] be a n degree polynomial of order n. Then $p_n(z)$ has exactly n roots, counting multiplicity.*

Proof. Postponed indefinitely. □

One thing I think might be slightly confusing to some students is the notion of multiplicity. Let's illustrate this idea using an example.

Example 4.4.1.

Let's find all the roots of the following polynomial:

$$x^3 - 5x^2 + 8x - 4$$

First, to simplify the problem, we can assume that this polynomial only has integer roots. Then, since 4 is divisible by $1, 2, 4$, and this is a third-order polynomial, if it's completely factorable, we should have three factors, implying we must find three numbers that have a product of -4. We try the value 1 first. If $x = 1$, the polynomial simplifies to

$$(1)^3 - 5(1)^2 + 8(1) - 4 = 0.$$

Then by theorem (4.2.2), we know $x - 1$ must be a factor. Then, removing this factor using polynomial long division, we get the polynomial

$$x^2 - 4x + 4$$

which factors to $(x-2)^2$, hence we conclude the final factoring of our cubic is

$$(x-2)^2(x-1).$$

[11] Ironiclly, because this theorem requires complex analysis to prove (which is like complex calculus), this theorem is neither fundamental nor a theorem about algebra.

[12] Generally, we use the variable z for complex numbers.

Since the factor $(x-2)$ occurs twice in the factoring, we claim this polynomial has the roots $x = \{1, 2\}$, with the root $x = 2$ having a multiplicity of 2, which satisfies theorem (4.4.1). ♠

Example 4.4.2.

In this example, let's examine all the roots (real or complex) of the expression

$$x^3 - 1$$

Naïvely, we might be tempted to move the 1 to the other side and take the cube root but doing so gets us

$$x^3 = 1$$

$$x = \sqrt[3]{1} = 1$$

Since this is the only root we can find in this manner, using theorem (4.2.2) and theorem (4.4.1), we conclude that 1 must be a root of multiplicity 3 implying

$$x^3 - 1 = (x - 1)^3$$

which is false. Now at this point, we might question the validity of the theorem, since our algebra is seemingly flawless here.

Let's reexamine the situation here. Since we know 1 is a root, let's divide out the factor $x - 1$ from our polynomial. Performing polynomial long division, we get

$$\frac{x^3 - 1}{x - 1} = x^2 + x + 1$$

Then using the quadric formula,[13] we get the remaining roots to be

$$-\frac{1}{2} \pm \frac{i\sqrt{3}}{2}$$

hence we conclude there are 3 roots to the polynomial and using theorem (4.2.2), we conclude that

$$x^3 - 1 = (x - 1)\left(x + \frac{1}{2} + \frac{i\sqrt{3}}{2}\right)\left(x + \frac{1}{2} - \frac{i\sqrt{3}}{2}\right).$$

[13] If you don't remember, $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

So, what went wrong for the first time around? Well, as it turns out, like the square root, when we take the codomain of the cube root function to be complex, we find the cube root actually gives three values,[14] which exactly correspond to the roots of the cubic. We will learn how to find these values in a later lesson. ♠

4.4.2 Factoring of Real Polynomials

As established with theorem (4.4.1), every polynomial can be factored completely to a first-order binomial (assuming we allow complex roots), but exactly how much factoring is guaranteed for real polynomials? Before we can answer that question, we have to answer a few smaller questions.

Theorem 4.4.2. *Let $p \in \mathbb{R}[x]$. If $z \in \mathbb{C}$ is a root of p, then $\bar{z}$ is also a root of p.*

Proof. Since

$$p(z) = \sum_{k=0}^{n} a_k z^n$$

where $a_k \in \mathbb{R}$. If z is a root of P_n, then

$$\sum_{k=0}^{n} a_k z^n = 0$$

Then taking the conjugate of both sides

$$\overline{\sum_{k=0}^{n} a_k z^n} = \sum_{k=0}^{n} \overline{a_k z^n} = \sum_{k=0}^{n} \overline{a_k}\,\overline{z^n} = \sum_{k=0}^{n} a_k \bar{z}^n = 0$$

by theorem (4.3.1). Therefore $\bar{z}$ is a root of p_n, proving the theorem. □

Lemma 4.4.1. *Let $z \in \mathbb{C}$. Then $((x-z)(x-\bar{z}))^n$ is a real polynomial, for any $n \in \mathbb{W}$.*

Proof. Since

$$(x-z)(x-\bar{z}) = x^2 - \bar{z}x - zx + z\bar{z} = x^2 - (x^2 - (z+\bar{z}) + z\bar{z})$$

[14] So the complex cube root isn't even a function for this matter.

By theorems (4.3.2) and (4.3.3), $(x - z)(x - \bar{z})$ is a real polynomial. Since the multiplication of real polynomials stays real, by closure properties of real numbers, $((x - z)(x - \bar{z}))^n$ must be a real polynomial. $\square$

Lemma 4.4.2. *For any $p \in \mathbb{R}[x]$, if z is a root of p, then $\bar{z}$ is a root of the same multiplicity.*

Proof. The proof of this theorem becomes trivial if $z \in \mathbb{R}$. Hence, we only consider the case when $z \in \mathbb{C} \setminus \mathbb{R}$. We will proceed with this proof by contradiction. Assume there exists p with root z of multiplicity m and $\bar{z}$ of multiplicity n such that $m \neq n$. Without loss of generality, let $m > n$. Then by theorem (4.2.2),

$$p(x) = (x - z)^m (x - \bar{z})^n g(x) = (x - z)^n (x - \bar{z})^n [(x - z)^{m-n} g(x)]$$

for some $g \in \mathbb{C}[x]$. Then since $p(x), (x - z)^n (x - \bar{z})^n \in \mathbb{R}[x]$, the latter of which is implied by lemma (4.4.1), by theorem (4.2.1), we have $d, r \in \mathbb{R}[x]$ such that

$$p(x) = (x - z)^n (x - \bar{z})^n d(x) + r(x)$$

Since each real polynomial is also a complex polynomial, by the uniqueness portion of theorem (4.2.1), we have

$$r(x) = 0, \quad d(x) = (x - z)^{m-n} g(x).$$

Therefore

$$(x - z)^{m-n} g(x) \in \mathbb{R}[x].$$

But since we divided out all factors of $x - \bar{z}$ from g, it follows that $\bar{z}$ is not a root of $g(x)$ by theorem (4.4.2). Similarly, we conclude, that $\bar{z}$ is not a root of $(x - z)^{m-n}$ since $z \neq \bar{z}$ when $z \notin \mathbb{R}$. Therefore, we conclude by contradiction that $m = n$. $\square$

Theorem 4.4.3. *If $p \in \mathbb{R}[x]$, then, p can be factored up to the quadratic factors.*

Proof. Let $n := \deg(p)$. Then, suppose p has m complex roots (counting multiplicity). Moreover, we can list the roots as $z_1, z_2, ..., z_m$ (including duplicates). By lemma (4.4.2), $\bar{z}_1, \bar{z}_2, ..., \bar{z}_m$ are also roots of p. Then let

$w_1, w_2, ..., w_l$ be the remaining real roots of p. Then by theorem (4.2.2)

$$p(x) = \left[\prod_{k=1}^{m}(x - z_k)(x - \bar{z}_k)\right]\left[\prod_{k=1}^{l}(x - w_k)\right]$$ [15]

$$= \left[\prod_{k=1}^{m}(x^2 - z_k x - \bar{z}_k x - z_k \bar{z}_k)\right]\left[\prod_{k=1}^{l}(x - w_k)\right]$$

Since $x^2 - z_k x - \bar{z}_k x - z_k \bar{z}_k$ is a real polynomial of order two, this proves our theorem. □

Corollary 4.4.1. *Let $p \in \mathbb{R}[x]$ such that $\deg(p)$ is odd. Then p has at least one real root.*

Proof. We prove this corollary by contradiction. Suppose there exists p that doesn't have any real roots. Then by theorem (4.4.3) and theorem (4.2.2), it must be factored into a series of second-ordered polynomials. Hence

$$p(x) = \prod_{k} g_k(x)$$

where $g_k(x)$ are second-ordered polynomials. But since the degree on the right-hand side must be even, this contradicts the fact that p is odd, hence, by contradiction, p must have at least one real root. □

4.5 Exercies

1. Prove the following identities used in theorem (4.2.1):

$$\deg(a - b) \leq \max\{\deg(b), \deg(b)\}, \quad \deg(ab) = \deg(a) + \deg(b).$$

2. Find $q, r \in \mathbb{R}[x]$ as specified in theorem (4.2.1) for the following:

   a) $x^4 - tx^2 + 3 = (x + 2)q(x) + r(x)$

   b) $x^3 + 2x^2 - 3x + 4 = (x - 7)q(x) + r(x)$

   c) $2x^5 + x^4 - 6x + 9 = (x^2 - 3x + 1)q(x) + r(x)$

3. Factor the following polynomials:

a) $x^3 - 2x^2 - x + 2$

b) $x^4 + x^3 - x^2 + x - 2$

c) $x^4 + 6x^3 + 8x^2 + 10x - 25$

4. Prove the following theorem:

Theorem. *For any $p \in \mathbb{R}[x]$, for some $a \in \mathbb{R}$, $p(a) = b$ if and only if $p(x) = (x-a)g(x) + b$.*

Hint: This uses a similar argument as to theorem (4.2.2)

5. **(Challenge)** Let $\mathbb{F}$ be a set with two operations, $+, \cdot$, such that for any $a, b \in \mathbb{F}$, $a + b \in \mathbb{F}$ and $a \cdot b \in \mathbb{F}$. Let these operations also satisfy the following:

(A1) (Commutativity of Addition) For any $a, b \in \mathbb{F}$, $a + b = b + a$

(A2) (Associativity of Addition) For any $a, b, c \in \mathbb{F}$, $(a+b)+c = a+(b+c)$

(A3) (Additive Identity) There exists some element $0 \in \mathbb{F}$ such that for any $a \in \mathbb{F}$, $a + 0 = a$

(A4) (Additive Inverse) For any $a \in \mathbb{F}$, there exists $-a \in \mathbb{F}$ such that $a + (-a) = 0$.

(A5) (Commutativity of Multiplication) For any $a, b \in \mathbb{F}$, $a \cdot b = b \cdot a$

(A6) (Associativity of Multiplication) For any $a, b, c \in \mathbb{F}$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

(A7) (Multiplicative Identity) There exists some element $1 \in \mathbb{F}$ such that for any $a \in \mathbb{F}$, $a \cdot 1 = a$

(A8) (Multiplicative Inverse) For any $a \in \mathbb{F}$, $a \neq 0$, there exists $a^{-1} \in \mathbb{F}$ such that $a \cdot a^{-1} = 1$.

(A9) (Distributive Law) For any $a, b, c \in \mathbb{F}$, $ab + ac = a(b + c)$.[16]

Let $a, b \in \mathbb{F}$. Prove $a \cdot b = 0$ if and only if $a = 0$ or $b = 0$.

[16] What I've defined here is what one might call a *field* in math.

Chapter 5

Inequalities

5.1 Basic Inequalities

5.1.1 Increasing or Decreasing

In this section, we will introduce the notion of an increasing or decreasing function. Generally, this concept is pretty easy to visualize on a plot for example in figure (5.1.1), we can observe an increasing function. Then

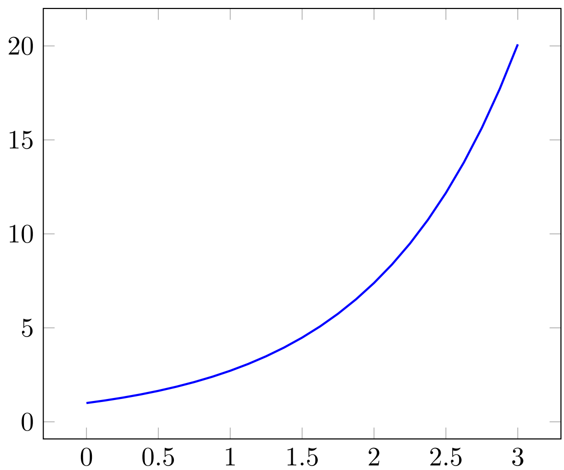


Figure 5.1.1: Plot of e^x.

in figure (5.1.2), we can observe $-\ln(x)$ is an example of a decreasing function. However, these declarations are only applied to a specific subset

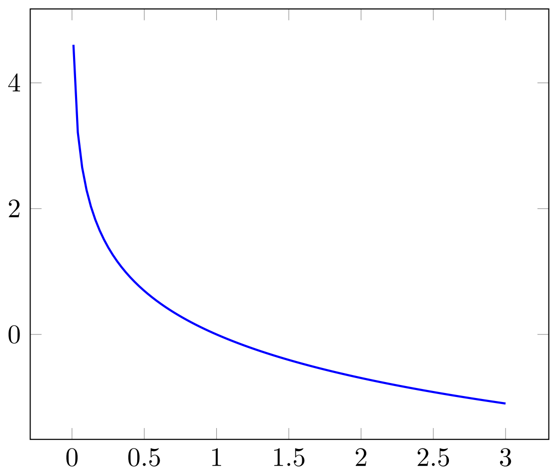


Figure 5.1.2: Plot of $-\ln(x)$.

of all possible functions. In figure (5.1.3), we can observe a function that

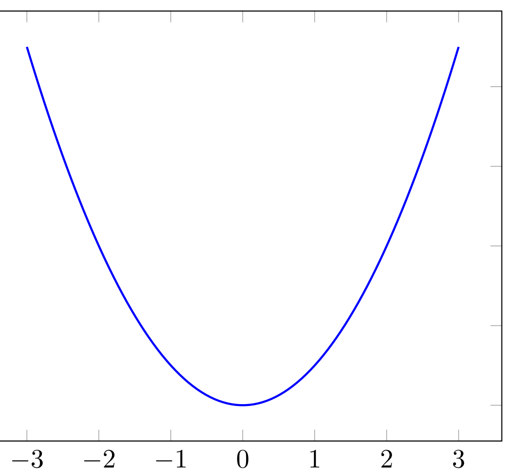


Figure 5.1.3: Plot of x^2.

is neither increasing nor decreasing. But it is here, that we note that whether a function is increasing or decreasing also heavily depends on what domain we define the function on. Had we restricted the domain to $x \geq 0$ in figure (5.1.3), we would get an increasing function.

Now with this basic intuition, let's define these ideas more rigorously with mathematics.

Definition 5.1.1. *Let $f : \mathbb{R} \to \mathbb{R}$[1] be an increasing function. Then $a < b$ implies $f(a) \leq f(b)$.*

Definition 5.1.2. *Let $f : \mathbb{R} \to \mathbb{R}$ be a decreasing function. Then $a < b$ implies $f(a) \geq f(b)$.*

Now had we exchanged the non-strict inequalities for strict equations, so $f(a) < f(b)$ and $f(a) > f(b)$ respectively, we would get the definition for a **strictly increasing** and **strictly decreasing** function. Now it must mentioned, that with the mathematical machinery that we have, it is rather difficult to prove any arbitrary real function is increasing or decreasing, so we will pick this up in a later lesson.

5.1.2 Solving Basic Inequalities

Example 5.1.1.

In this example, let's solve the inequality

$$12 > 7 - 2y$$

This might seem very obvious to many of you, but in this example, I will attempt to demonstrate a slightly different technique to solve this. Define a function $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x - 7$. By plotting the function on an axis, this f is increasing. Therefore

$$12 > 7 - 2y$$

implies

$$f(12) > f(7 - 2y)$$

$$(12) - 7 > (7 - 2y) - 7$$

$$-7 > -2y$$

[1] Technically, this is defined for any ordered set with a notion of "biggness", but for this introductory course, we will restrict ourselves just to $\mathbb{R}$.

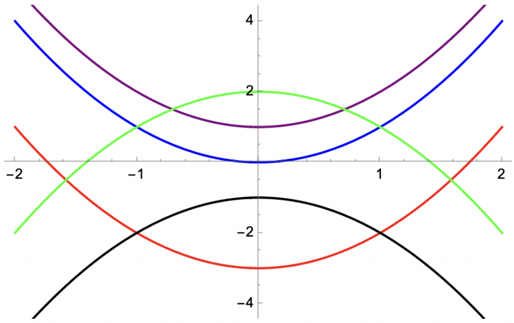


Figure 5.1.4

Then define a different function $g : \mathbb{R} \to \mathbb{R}$ such that $g(x) = -\frac{1}{2}x$. Since this is a decreasing function,

$$-7 > -2y$$

implies

$$g(-7) < g(-2y)$$
$$-\frac{1}{2}(-7) < -\frac{1}{2}(-2y)$$
$$\frac{7}{2} < y$$

hence we conclude every y such that

$$y \in \left(-\infty, \frac{7}{2}\right)$$

satisfies our inequality. ♠

Example 5.1.2.

Let's solve the inequality

$$x^2 + 3x \leq -2$$

We can by the same technique in example (5.1.1) show

$$x^2 + 3x + 2 \leq 0$$

we know by plotting the general quadratic as in figure (5.1.4), there are three options:

1) There are no solutions.

2) There is one solution.

3) There is a connected interval of solutions between the two zeros.

4) There is a disjoint interval of solutions around the two zeros.

5) Every real number satisfies the inequality.

Therefore, if the quadratic has one solution, then the inequality has no solutions, if the quadratic has no solutions, then the inequality has either no solutions or every real number is a solution. If there are exactly 2 solutions, then the quadric has either a connected or disjoint interval of solutions. Let's see which scenario we have here. Solving the quadric gives

$$x^2 + 3x + 2 = 0$$

$$(x + 2)(x + 1) = 0$$

hence

$$x = \{-2, -1\}.$$

Therefore our inequality has either a connected or disjoint interval. To figure out which one it is, we just need to test a point. I usually just use 0, but note this only works if 0 is not a zero of our polynomial. Substituting 0:

$$(0)^2 + 3(0) + 2 = 2 \nleq 0$$

Since 0 is in the disjoint portion of or quadratic, we conclude

$$x \in [-2, -1]$$

satisfies our inequality. ♠

5.1.3 Absolute Values

To help us build up to our big topic at the end, we're going to examine inequalities involving absolute values.

First, we are going to examine a couple of basic properties of the absolute value. We know, by just thinking about it if

$$|x| < c$$

obviously requiring $c > 0$, then

$$-c < x < c$$

The implication here goes both ways, so if

$$-c < x < c,$$

then that also implies

$$|x| < c$$

Then if we have

$$c < |x|$$

then instead of getting a connected interval, here we would get a disjoint one, with

$$x > c \quad \text{or} \quad x < -c$$

The implication here also goes both ways. Now with this, let's attempt to prove our first theorem.

Theorem 5.1.1 (Triangle Inequality)**.** *Suppose* $x, y \in \mathbb{R}$*, then*

$$|x| + |y| \geq |x + y|$$

Proof. By the definition of the absolute value,

$$-|x| \leq x \leq |x|$$

$$-|y| \leq y \leq |y|$$

Therefore adding the two inequalities gives

$$-(|x| + |y|) \leq x + y \leq |x| + |y|$$

This implies

$$|x + y| \leq ||x| + |y||$$

Since $|x| + |y| \geq 0$,

$$|x + y| \leq |x| + |y|$$

□

Corollary 5.1.1. *Suppose $x, y \in \mathbb{R}$, then*

$$||x| - |y|| \leq |x - y|$$

Proof. For this one, we are going to define the variable $a = x - y$. Therefore by theorem (5.1.1),

$$|a| + |y| \geq |a + y| = |x|$$

Therefore

$$|x - y| \geq |x| - |y|$$

Then letting $b = y - x$, by. theorem (5.1.1),

$$|x| + |b| \geq |x + b| = |y|$$

implying

$$|y - x| \geq |y| - |x|$$

$$|x - y| \geq -(|x| - |y|)$$

$$-|x - y| \leq |x| - |y|$$

Therefore

$$-|x - y| \leq |x| - |y| \leq |x - y|$$

implying

$$|x - y| \geq ||x| - |y||$$

□

5.2 Bounds

5.2.1 Bounded Functions

Definition 5.2.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be bounded above. Then there exists $M \in \mathbb{R}$ such that for any $x \in \mathbb{R}$, $M \geq f(x)$.*

Definition 5.2.2. *Let $f : \mathbb{R} \to \mathbb{R}$ be bounded below. Then there exists $M \in \mathbb{R}$ such that for any $x \in \mathbb{R}$, $M \leq f(x)$.*

Definition 5.2.3. *Let $f : \mathbb{R} \to \mathbb{R}$ be bounded. Then there exists $M > 0$ such that for any $x \in \mathbb{R}$, $M \geq |f(x)|$, or equivalently, f has an upper and lower bound.*

Now, to present these concepts more visually, we examine the function $f(x) = \frac{\sin(x)}{x}$. In figure (5.2.1), we can see I've bound the function by $y = 1.2, 1.5, 1.8$. Now since the definition of a bounded function only

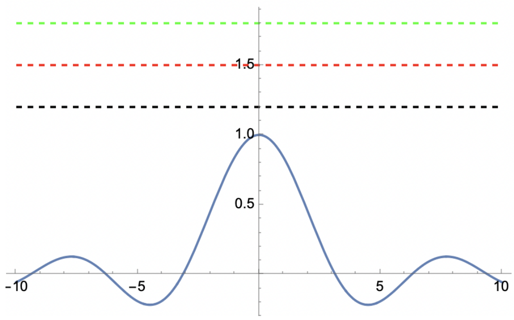


Figure 5.2.1

requires a bound to exist, not necessarily the least, we could have used any of these values for our M.[2]

This same goes for choosing a lower bound as in figure (5.2.2).

Now let's use this idea to prove some theorems.

Theorem 5.2.1. *Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be bounded functions. Then*

1) $f + g$ is bounded,

2) $f - g$ is bounded,

3) $f \cdot g$ is bounded.

[2]It turns out that because of our existence of a finite bound, the least upper bound must exist by the Axiom of Completeness since we are working in real numbers. This is a special property of the real numbers which is not true in general. This idea helps us define what we mean by the real numbers.

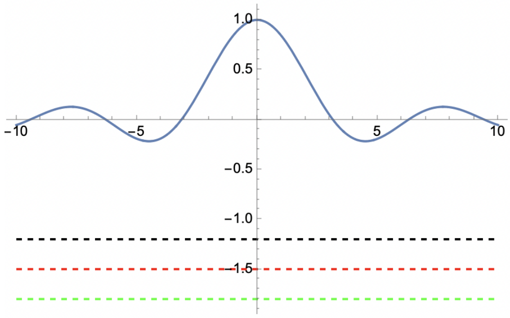


Figure 5.2.2

Proof. Since f and g are bounded, there exists $M, N > 0$ such that

$$|f(x)| \leq M \quad \text{and} \quad |g(x)| \leq N$$

Then

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq M + N$$

which proves (1). The proof of (2) follows exactly as (1).

Then since

$$|f(x)||g(x)| = |f(x)g(x)| \leq MN,$$

This implies $f \cdot g$ is bounded, proving (3). □

5.2.2 Limits

Limits are a concept that will be vital for your understanding of calculus. In precalc, we are going to introduce this idea so you can become familiar with it moving into calculus.

A limit put simply is just the value a function gets closer and closer to, even if the function doesn't reach this value. The limit allows us to give meaning to discontinuities that would otherwise be undefined. In figure (5.2.3), we see the plot of $f(x) = \frac{x-1}{(x-1)(x-2)}$. By finding the maximum possible domain from this function, we get the set

$$\mathcal{D} = \{x : x \neq 1, x \neq 2\}$$

For $x = 2$, the function is going to infinity, so we are going to ignore that. But for $x = 1$, the function clearly *wants* to output a value, but just can't

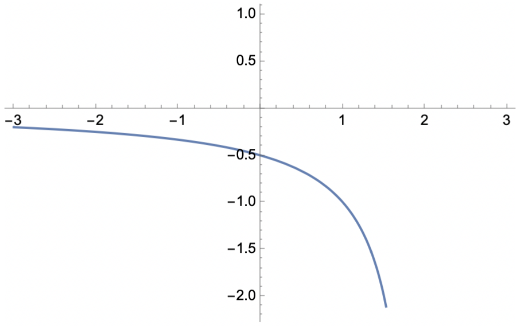


Figure 5.2.3

because of the divides by zero error. This is where the limit comes in handy because, with the limit, we can observe the behavior around $x = 1$ to assign a meaningful value to the function.

Now another way to characterize the limit is by computing values closer and closer to the problem. In figure (5.2.4), we observe as we get closer to $x = 1$, the value of the function gets closer and closer to -1, which is indeed what we get when we compute the limit of f as x approaches 1. Now this seems quite contrary to the rigor that we often associate with

x	$f(x)$
0.9	-0.909
0.95	-0.952
0.99	-0.990
0.999	-0.999
1.001	-1.001
1.01	-1.010
1.05	-1.053
1.1	-1.111

Figure 5.2.4

mathematics, so let's rigorously define this concept.

Definition 5.2.4. *Suppose* $f : \mathbb{R} \to \mathbb{R}$[3] *such that*

$$\lim_{x \to c} f(x) = L,$$

where L *is finite, then for every* $\epsilon > 0$*, there exists* $\delta > 0$ *such that* *for all* $x \neq c$, $|x - c| < \delta$ *implies* $|f(x) - L| < \epsilon$.

That might be very hard to take and might be our hardest concept yet, but let's explain the meaning behind this using the idea of bounds. When we write $|x - c|$, what we are computing here is just the distance between two points in the real numbers. This might seem obvious to some, but this is super important in understanding this definition. Therefore here, the δ and ϵ can be understood as the bounds of the distance between x and c, our point in question, and $f(x)$ and L respectively.

Now the important part is that the bound on $|f(x) - L|$ can be arbitrarily small, since $\epsilon > 0$, and for every ϵ, we can find an interval for x, with radius δ around c such that $|f(x) - L|$ on this interval is bounded by ϵ. In other words, we can bound $|f(x) - L|$ by an arbitrarily small number by shrinking our domain around c.

Now let's see what we mean by if the limit doesn't exist. If the limit doesn't exist, then there exists some ϵ which cannot be a bound for $|f(x) - L|$ no matter how small we shrink the interval around c for any L. We can see in figure (5.2.5) that no matter how much we shrink the interval

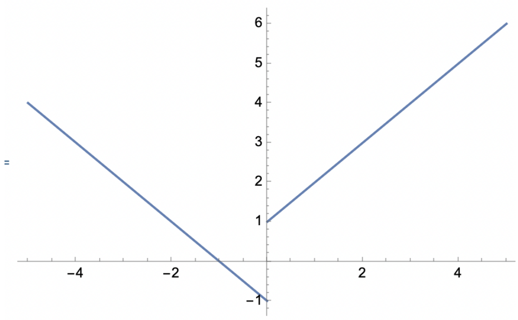


Figure 5.2.5

around 0, we cannot bound $|f(x) - L|$ by $\epsilon = 0.5$. Therefore, we conclude the limit around 0 doesn't exist for $f(x)$.

[3]There is no sense in talking about limits for anything other than $\mathbb{C}$ or $\mathbb{R}$, since a sense of connectedness must exist for this to make sense. There are other sets in which this sense of a limit exists, we call those sets topological spaces.

Example 5.2.1.

Let's use definition (5.2.4) to prove

$$\lim_{x\to 3} \frac{x^2-9}{x-3} = 6$$

To do this, we would need a way to show that every interval around 3 exists so we can bound a $|f(x)-3|$ for any ϵ. To do this, we typically try to find a function that outputs working δs for any given ϵ. After we find this function, all we have to do is prove that this function works within our definition of the limit.

This part because it's not part of the proof doesn't have to be as precise, so long as we can justify our findings later in a more rigorous proof. So let's try finding.

$$\epsilon = \left| \frac{x^2-9}{x-3} - 6 \right|.$$

Since when talking about the limit around 3, $x \neq 3$, let's divide it out of our expression.

$$\epsilon = |x+3-6| = |x-3|$$

Since we want a delta such that

$$|x-3| < \delta$$

therefore

$$\epsilon < \delta$$

Since we overestimated for ϵ anyway, it might be sensible to set

$$\epsilon = \delta.$$

Let's see if this choice works in a proof. Starting with $|x-3| < \delta$, since

$$|x-3| = |x+3-6| = \left| \frac{x^2-9}{x-3} - 6 \right| < \delta = \epsilon,$$

we have shown for every ϵ, we can find a delta such that the above is true, implying

$$\lim_{x\to 3} \frac{x^2-9}{x-3} = 6$$

♠

Example 5.2.2.

In this example, let's show

$$\lim_{x \to 2} x^2 + 2x = 8$$

Let's first find a suitable definition of ϵ in terms of δ. First, let's find

$$|x^2 + 2x - 8| = \epsilon$$

$$|x + 4||x - 2| = \epsilon$$

Since

$$|x - 2| < \delta$$

Now it's here that we must mention that we might need an extra condition on δ. Since we what our expression for δ to only depend on ϵ, we cannot let

$$\delta = \frac{\epsilon}{|x + 4|}$$

A remedy here is to set $\delta < 1$. This choice here is arbitrary, but sometimes a more careful choice of this bound is needed to prove more complicated limits. Let's examine the utility of our condition here. Since $x < 3$, by $\delta < 1$,

$$|x + 4||x - 2| = \epsilon < 7|x - 2|$$

Then it might be sensible to set

$$\delta = \frac{\epsilon}{7}$$

Let's see if our choice of δ works.

Without loss of generality, we can assume $\delta < 1$. Then since

$$|x - 2| < \delta = \frac{\epsilon}{7}$$

$$7|x - 2| < \epsilon$$

Since

$$|x + 4| < |x - 2| + 6 < \delta + 6 < 7,$$

chaining the two inequalities together, we get

$$|x + 4||x - 2| < \epsilon$$

$$|x^2 + 2x - 8| < \epsilon$$

hence proving our proposition. ♠

Now let's prove some generalized properties of limits.

Theorem 5.2.2. *Suppose*

$$\lim_{x \to c} f(x) = L \quad \text{and} \quad \lim_{x \to c} g(x) = M$$

where L and M are finite. Then

$$\lim_{x \to c} f(x) + \lim_{x \to c} g(x) = \lim_{x \to c} [f(x) + g(x)]$$

Proof. Since the limits for f and g exist, for every $\epsilon > 0$, there exists $\delta > 0$ such that for every x such that $0 < |x - c| < \delta$[4]

$$|f(x) - L| < \frac{\epsilon}{2} \quad \text{and} \quad |g(x) - M| < \frac{\epsilon}{2}$$

Therefore since by the theorem (5.1.1),

$$|f(x) + g(x) - (L + M)| < |f(x) - L| + |g(x) - M| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

hence proving our theorem. □

Theorem 5.2.3. *Suppose*

$$\lim_{x \to c} f(x) = L \quad \text{and} \quad \lim_{x \to c} g(x) = M$$

where L and M are finite. Then

$$\lim_{x \to c} f(x) \cdot \lim_{x \to c} g(x) = \lim_{x \to c} [f(x) \cdot (x)]$$

Proof. Since we know the limit exists for f around c, we know there must exist an interval $\lambda > 0$ such that $0 < |x - c| < \lambda$ where f is bounded.[5] We

[4] I combined the δs by taking the least required for the two functions. We can do this because δ all this does is shrink what the bound could have been for ϵ.

[5] Use the definition of the limit to show this.

set let N be this bound such that

$$|f(x)| < N.$$

Then without loss of generality, we let $\delta < \lambda$ so by the existence of the limit for f and g, for every $\epsilon > 0$, there exists $\delta > 0$ such that for every x such that $0 < |x - c| < \delta < \lambda$

$$|f(x) - L| < \frac{\epsilon}{2L} \quad \text{and} \quad |g(x) - M| < \frac{\epsilon}{2N}$$

Then we know

$$|f(x)g(x) - LM| = |f(x)g(x) - Mf(x) + Mf(x) - LM|$$

$$= |f(x)(g(x) - M) + M(f(x) - L|$$

whereby theorem (5.1.1),

$$\geq |f(x)(g(x) - M)| + |M(f(x) - L| = |f(x)||g(x) - M)| + |M||(f(x) - L|$$

whereby $|f(x)| < N$,

$$< N\left(\frac{\epsilon}{2N}\right) + L\left(\frac{\epsilon}{2L}\right) = \epsilon$$

hence proving the theorem. □

Theorem 5.2.4.

$$\lim_{x\to c} f(x) = L$$

where $L \neq 0$ *and finite. Then*

$$\frac{1}{\lim_{x\to c} f(x)} = \lim_{x\to c}\left[\frac{1}{f(x)}\right]$$

Proof. Since the limit exists for f,

$$|f(x) - L| < \frac{|L|}{2} \quad \text{for } 0 < |x - c| < \delta_1.$$

Since

$$|L| = |L + f(x) - f(x)| \geq |L - f(x)| + |f(x)|$$

$$= |f(x) - L| + |f(x)| < \frac{|L|}{2} + |f(x)|$$

hence

$$\frac{|L|}{2} < |f(x)|$$

$$\frac{1}{f(x)} < \frac{2}{|L|}$$

Since there also exists δ_2 such that

$$|f(x) - L| < \frac{|L|^2}{2}\epsilon \quad \text{for } 0 < |x - c| < \delta_2$$

If we let $\delta = \min(\delta_1, \delta_2)$, for $0 < |x - c| < \delta$ we have

$$|\frac{1}{f(x)} - \frac{1}{L}| = |\frac{L - f(x)}{Lf(x)}|$$

$$= \left|\frac{1}{Lf(x)}\right| |L - f(x)|$$

$$= \left|\frac{1}{Lf(x)}\right| |f(x) - L)| < \frac{1}{|L|}\frac{2}{|L|}|g(x) - L|$$

$$< \frac{2}{|L|^2}|g(x) - L| < \frac{2}{|L^2|}\frac{|L|^2}{2}\epsilon$$

$$= \epsilon.$$

Therefore

$$\frac{1}{\lim_{x\to c} f(x)} = \lim_{x\to c}\left[\frac{1}{f(x)}\right]$$

□

Chapter 6

Transcendental Functions

6.1 Exponentials

6.1.1 Exponentials for Reals

Before we dive into this topic, we must define the idea of taking an exponential, with the domain being the real numbers. Traditional exponentials have the disadvantage of only working with integer values or rational numbers. However, the problem is that the rational numbers are only a very small subset of the set of real numbers.[1] So for our purposes, we need something a little better. For our definition, we'd like to seek a nice connected function that outputs the same values as the discrete exponent and also satisfies all of the properties that the discrete case does. Unfortunately, this definition is quite a calculus heavily so I will leave that discussion for another class. For our class, we will only so we will only look over a few of its properties.

Theorem 6.1.1. *Let $x, y \in \mathbb{R}$, and $a \in \mathbb{R}^+$ then the following hold:*

1) $a^x b^y = a^{a+b}$

2) $(a^x)^y = a^{xy}$

3) for $b \in \mathbb{R}$, $(ab)^x = a^x b^x$

[1] The set of rational numbers is so small that the entire set doesn't even take up "space" in the real numbers. It is what's called a *null set*.

*(**Note**: the exponential is only defined for a that is positive)*

We can in fact verify these properties quite easily if we shrink the domain to just the rational numbers. I will leave this as an exercise for the reader.

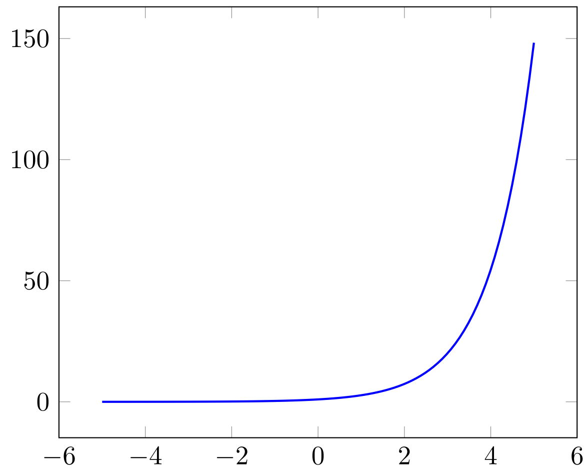


Figure 6.1.1

If we decide to plot the exponential function, we notice in figure (6.1.1) that the exponential is defined for all real numbers, but as the input gets smaller and smaller, it seems to *approach* zero. We call this a **horizontal asymptote** and claim there's a horizontal asymptote for the function a^x when x approaches $-\infty$. Since the function never reaches zero, we claim the domain and image for the exponential function is

$$\mathcal{D} = \mathbb{R} \quad \text{and} \quad a^{\mathcal{D}} = (0, +\infty).$$

6.1.2 Inverse of Exponential (Logarithm)

Definition 6.1.1. *Let $f(x) = a^x$. Then $f^{-1}(x) = log_a(x)$ read "log base a of x".*[2]

As you can see in figure (6.1.2), the logarithm is defined for all $x \in (0, \infty)$. We can observe there is a **vertical asymtote** at 0. Since as far as we

[2]When you see a log without a subscript, it can mean one of two things depending on context. In general, this will mean log_{10}, since this is used often when plotting data that grows exponentially. In a more mathematical context, this will generally mean natural log, which we will discuss later.

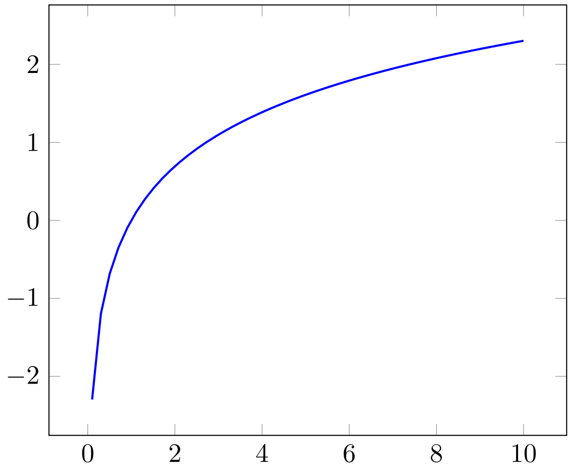


Figure 6.1.2

are concerned, the image of a^x doesn't include negative numbers, the logarithm is not defined for $(-\infty, 0]$.

Now we have the definition of the logarithm, let's look at some of its properties.

Theorem 6.1.2. *Let $x, y \in \mathbb{R}$, and $a \in \mathbb{R}^+$ then the following hold:*

1) $\log_a(x) + \log_a(y) = \log_a(xy)$

2) $y \log_a(x) = \log_a(x^y)$

Proof. Since

$$a^{\log_a(xy)} = xy = a^{\log_a(x)} a^{\log_a(y)} = a^{\log_a(x)+\log_a(y)}.$$

Then taking the logarithm of both sides gives

$$\log_a(a^{\log_a(xy)}) = \log_a(a^{\log_a(x)+\log_a(y)})$$

$$\log_a(xy) = \log_a(x) + \log_a(y).$$

hence proving (1).

Then since

$$a^{y \log_a(x)} = x^y = a^{\log(x^y)}$$

Taking the logarithm of both sides:

$$\log_a(a^{y \log_a(x)}) = \log_a(a^{\log(x^y)})$$

$$y \log_a(x) = \log(x^y).$$

Hence proving (2). □

6.1.3 Change of Basis

Theorem 6.1.3. *Let $a, b \in (0, \infty)$ and $x \in \mathbb{R}$. Then*

$$a^x = b^{\log_b(a)x}$$

Proof. Since $a = b^{\log_b(a)}$, using the properties the exponential,

$$a^x = (b^{\log_b(a)})^x = b^{\log_b(a)x}$$

hence proving the theorem. □

Theorem 6.1.4. *Let $a, b \in (0, \infty)$ and $x \in \mathbb{R}$. Then*

$$\log_a(x) = \frac{\log_b(x)}{\log_b(a)}$$

Proof. By the definition of the logarithm, $\log_a(x)$ implies

$$a^{\log_a(x)} = b^{\log_b(a) \log_a(x)} = x$$

since $a = b^{\log_b(a)}$. Taking the $\log_b$ on both sides gives

$$\log_b(b^{\log_b(a) \log_a(x)}) = \log_b(x)$$

hence

$$\log_b(a) \log_a(x) = \log_b(x)$$

implying

$$\log_a(x) = \frac{\log_b(x)}{\log_b(a)}$$

□

What theorems (6.1.3) and (6.1.4) tell us is that we can write any exponential as an exponential function of any other base.

The utility of theorems (6.1.3) and (6.1.4) might not be immediately clear, so let's discuss it here. Like π, there is a number $e \approx 2.71828$ which is also irrational that we often use in the context of exponentials. e can be thought of as a preferred basis when working with exponentials, due to its special properties. We will not discuss the special properties of this number, since it relates heavily to calculus, but we will introduce the notation to get you familiar with the idea. For example, using theorem (6.1.3), we can write the function

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

as

$$f(x) = e^{\log_e(x)x}$$

Because of how fundamental e^x is in the context of exponentials, we give it a special name: $\exp(x)$ (where exp stands for exponential). The advantage of this notation is that sometimes, with longer arguments, it's a little easier to read, but it also highlights that although similar, the $\mathbb{R}$ exponential and the discrete version are distinct mathematical operations. Additionally, we also define a special function for $\log_e(x)$ called $\ln(x)$, or natural logarithm (or *logarithme naturel* in French).

Although we started by defining an exponential with an arbitrary base, in modern mathematics, we first define $\exp(x)$ (or $\ln(x)$) since we view e^x and $\ln(x)$ as more fundamental than an exponential with an arbitrary base (hence why its called *natrual* log). Hopefully, in a future math class, the reason for this will become clearer, but for this class, we will end our discussion here.

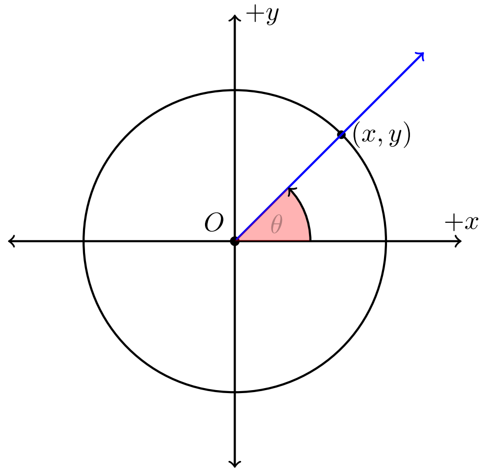
6.2 Trignometry

6.2.1 Definition

Now we have come to everyone's favorite topic, trig. But before, we get into this topic, we must define our functions.

Definition 6.2.1. *Using the figure below, let $\odot O$ be a unit circle (radius 1), then for any $\theta \in \mathbb{R}$ (measuring counterclockwise from $+x$ axis if θ is positive, and clockwise if θ is negative)*

$$x = \cos(\theta) \quad \textit{and} \quad y = \sin(\theta)^{3}$$



Then using definition (6.2.1), we can define the other trig functions in the conventional manner:

$$\tan(x) = \frac{\sin(x)}{\cos(x)} \quad \text{for } x \in \left\{x \in \mathbb{R} : x \neq \frac{\pi}{2} + \pi n,\ \forall n \in \mathbb{N}\right\}$$

$$\cot(x) = \frac{\cos(x)}{\sin(x)} \quad \text{for } \{x \in \mathbb{R} : x \neq \pi n,\ \forall n \in \mathbb{N}\}$$

$$\csc(x) = \frac{1}{\sin(x)} \quad \text{for } \{x \in \mathbb{R} : x \neq \pi n,\ \forall n \in \mathbb{N}\}$$

$$\sec(x) = \frac{1}{\cos(x)} \quad \text{for } x \in \left\{x \in \mathbb{R} : x \neq \frac{\pi}{2} + \pi n,\ \forall n \in \mathbb{N}\right\}$$

Then to find the image, trivially from the definition (6.2.1), $|\sin(x)|, |\cos(x) \leq 1$, hence the image of $\sin(x)$ and $\cos(x)$ is the interval $[-1, 1]$. Then immediately, since $\csc(x)$ and $\sec(x)$ are just the recipricals of $\sin(x)$ and $\cos(x)$ respectivly, $|\csc(x)|, |\sec(x)| \geq 1$ hence the image these functions is the compound interval $[-\infty, -1] \cup [1, \infty]$.

Finding the image of $\tan(x)$ and $\cot(x)$ can be a bit more tricky, but

[3] Check this is the same definition as *"SOHCAHTOA"*.

we will leverage a computer to help us do this. Using figures (6.2.1) and (6.2.2), we see the image of our two functions is $\mathbb{R}$.

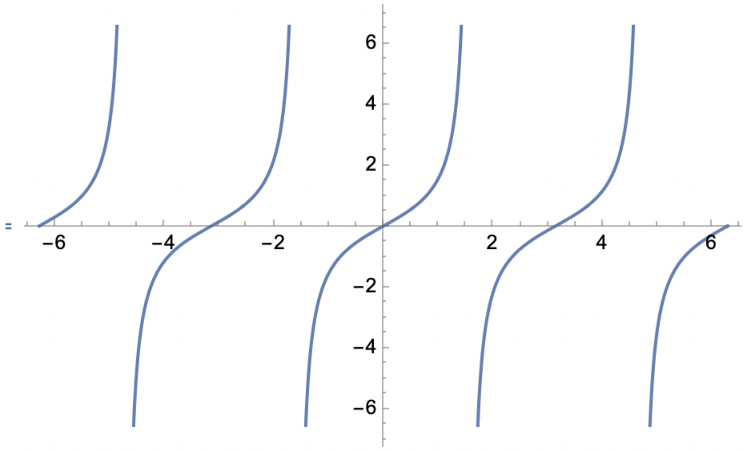


Figure 6.2.1: Plot of $\tan(x)$

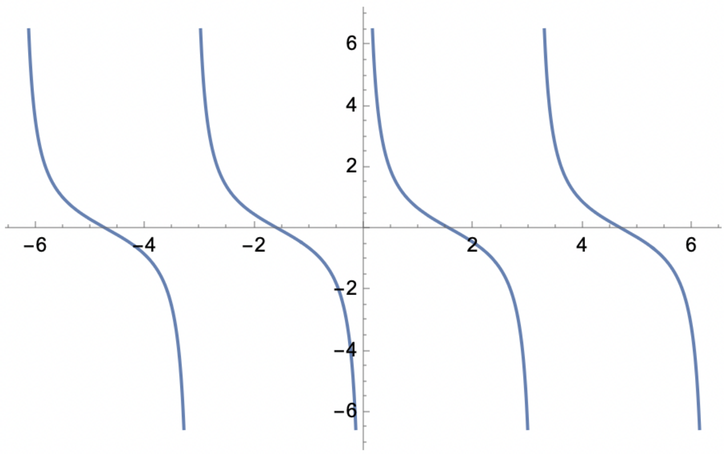


Figure 6.2.2: Plot of $\cot(x)$

6.2.2 Basic Identities

Trig Function Identities

Theorem 6.2.1. $\sin(x)$ *and* $\cos(x)$ *are* 2π *periodic, or equivalently,*

$$\sin(x) = \sin(x + 2\pi) \quad \textit{and} \quad \cos(x) = \cos(x + 2\pi)$$

Proof. Since $\sin(x)$ $\cos(x)$ are defined on the unit circle, the angle can be represented x can be more generally written as

$$x + 2\pi k \quad \text{for} \ \ k \in \mathbb{Z}$$

Hence

$$\sin(x + 2\pi k) = \sin(x) \quad \text{and} \quad \cos(x + 2\pi k) = \cos(x)$$

hence proving the theorem. □

Corollary 6.2.1. *sec*(x), $\csc(x)$ *are* 2π *periodic.*

Proof. This proof is trivial with theorem (6.2.1). □

Theorem 6.2.2. $\tan(x)$, $\cot(x)$ *are* π *periodic*

Proof. Postponed for a later section. □

Theorem 6.2.3. $\sin(x)$ *is odd.*

Proof. Let's prove this statement first only for the domain $[-\pi, \pi]$. For an angle θ, since $-\theta$ is the same angle reflected over the x-axis, by this reflection and definition (6.2.1),

$$\sin(x) = -\sin(-x)$$

Then by theorem (6.2.1),

$$\sin(x) = \sin(x + 2\pi k) = -\sin(-x) = -\sin(-x + 2\pi m)$$

for arbitrary $k, m \in \mathbb{Z}$. Then since these constants are arbitrary, let $k = -m$, hence

$$\sin(x + 2\pi k) = -\sin(-x - 2\pi k)$$

which implies

$$\sin(x + 2\pi k) = -\sin(-(x + 2\pi k))$$

Since $x + 2\pi k$ covers the entire set $\mathbb{R}$, we conclude that $\sin(x)$ is odd over $\mathbb{R}$. □

Theorem 6.2.4. $\cos(x)$ *is even.*

Proof. Following very closely to the previous proof, let's prove this statement first only for the domain $[-\pi, \pi]$. For an angle θ, since $-\theta$ is the same angle reflected over the x-axis, by this reflection and definition (6.2.1),

$$\cos(x) = \cos(-x)$$

Then by theorem (6.2.1),

$$\cos(x) = \cos(x + 2\pi k) = \cos(-x) = \cos(-x + 2\pi m)$$

for arbitrary $k, m \in \mathbb{Z}$. Then since these constants are arbitrary, let $k = -m$, hence

$$\cos(x + 2\pi k) = \cos(-x - 2\pi k)$$

which implies

$$\cos(x + 2\pi k) = \cos(-(x + 2\pi k))$$

Since $x + 2\pi k$ covers the entire set $\mathbb{R}$, we conclude that $\cos(x)$ is even over $\mathbb{R}$. □

Corollary 6.2.2. $\tan(x)$,$\cot(x)$,$\csc(x)$ *are odd, and sec*(x) *is even.*

Proof. Exercise. □

Pythagorean-esque Identies

Theorem 6.2.5 (Pythagorean Identity)**.** *Let* $x \in \mathbb{R}$*, then*

$$\cos^2(x) + \sin^2(x) = 1$$

Proof. By definition (6.2.1), the ordered pair $(\cos(x), \sin(x))$ lies on a unit circle. Therefore since a circle satisfies the relation

$$x^2 + y^2 = 1$$

it immediately follows that

$$\cos^2(x) + \sin^2(x) = 1$$

□

Corollary 6.2.3. *Let* $x \in \left\{x \in \mathbb{R} : x \neq \frac{\pi}{2} + \pi n,\ \forall n \in \mathbb{N}\right\}$, *then*

$$sec^2(x) = 1 + \tan^2(x)$$

Proof. Using theorem (6.2.5),

$$\cos^2(x) + \sin^2(x) = 1$$

Since $\cos(x) \neq 0$ by our domain restriction,

$$1 + \frac{\sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)} = \sec^2(x)$$

hence

$$1 + \tan^2(x) = \sec^2(x).$$

□

Corollary 6.2.4. *Let* $x \in \{x \in \mathbb{R} : x \neq \pi n,\ \forall n \in \mathbb{N}\}$, *then*

$$\csc^2(x) = 1 + \cot^2(x)$$

Proof. Using theorem (6.2.5),

$$\cos^2(x) + \sin^2(x) = 1$$

Since $\sin(x) \neq 0$ by our domain restriction,

$$\frac{\cos^2(x)}{\sin^2(x)} + 1 = \frac{1}{\sin^2(x)} = \csc^2(x)$$

hence

$$1 + \cot^2(x) = \csc^2(x).$$

□

Angle Identies

Theorem 6.2.6. *For $\alpha, \beta \in \mathbb{R}$,*

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$$

Proof. Postponed indefinitely. □

Corollary 6.2.5. *For $\alpha, \beta \in \mathbb{R}$,*

$$\cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta)$$

Proof. Using theorems (6.2.4) and (6.2.3),

$$\cos(\alpha-\beta) = \cos(\alpha)\cos(-\beta)-\sin(\alpha)\sin(-\beta) = \cos(\alpha)\cos(\beta)+\sin(\alpha)\sin(\beta)$$

□

Theorem 6.2.7. *For $\alpha, \beta \in \mathbb{R}$,*

$$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)$$

Proof. Postponed indefinitely. □

Corollary 6.2.6. *For $\alpha, \beta \in \mathbb{R}$,*

$$\sin(\alpha - \beta) = \sin(\alpha)\cos(\beta) - \cos(\alpha)\sin(\beta)$$

Proof. Using theorems (6.2.4) and (6.2.3),

$$\sin(\alpha-\beta) = \sin(\alpha)\cos(-\beta)+\cos(\alpha)\sin(-\beta) = \sin(\alpha)\cos(\beta)-\cos(\beta)\sin(\beta)$$

□

Corollary 6.2.7 (Double angle). *For $\theta \in \mathbb{R}$, the following are true:*

$$\sin(2\theta) = 2\sin(\theta)\cos(\theta)$$

$$\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$$

Proof. Trivial by theormes (6.2.6) and (6.2.7). □

Corollary 6.2.8 (Half angle). *For $\theta \in \mathbb{R}$, the following are true:*

$$\sin\left(\frac{\theta}{2}\right) = \pm\sqrt{\frac{1-\cos(\theta)}{2}}$$

$$\cos\left(\frac{\theta}{2}\right) = \pm\sqrt{\frac{1+\cos(\theta)}{2}}$$

Proof. Using corrolary (6.2.7) and letting $x = \frac{\theta}{2}$ gives

$$\cos(x) = \cos^2\left(\frac{x}{2}\right) - \sin^2\left(\frac{x}{2}\right).$$

Then by theorem (6.2.5),

$$\cos^2(\frac{x}{2}) + \sin^2(\frac{x}{2}) = 1$$

which implies

$$\sin^2\left(\frac{x}{2}\right) = 1 - \cos^2\left(\frac{x}{2}\right).$$

Therefore by substitution,

$$\cos(x) = \cos^2\left(\frac{x}{2}\right) - \left(1 - \cos^2\left(\frac{x}{2}\right)\right)$$

hence

$$\cos(x) + 1 = 2\cos^2\left(\frac{x}{2}\right)$$

$$\cos\left(\frac{x}{2}\right) = \pm\sqrt{\frac{\cos(x)+1}{2}}$$

Then again by theorem (6.2.5),

$$\cos^2\left(\frac{x}{2}\right) = 1 - \sin^2\left(\frac{x}{2}\right)$$

is true. Therefore by substitution,

$$\cos(x) = \left(1 - \sin^2\left(\frac{x}{2}\right)\right) - \sin\left(\frac{x}{2}\right)$$

implying

$$\cos(x) - 1 = -2\sin\left(\frac{x}{2}\right)$$

hence

$$\sin\left(\frac{x}{2}\right) = \pm\sqrt{\frac{1-\cos(x)}{2}}$$

thereby proving the theorem. □

Then using the sum and difference identities for $\sin(x)$ and $\cos(x)$, we can prove the following.

Theorem 6.2.8. *For any $x \in \mathbb{R}$,*

$$\sin(x) = \cos\left(\frac{\pi}{2} - x\right)$$

Proof. Using theorem (6.2.6),

$$\cos\left(\frac{\pi}{2} - x\right) = \cos\left(\frac{\pi}{2}\right)\cos(x) + \sin\left(\frac{\pi}{2}\right)\sin(x)$$

Since

$$\sin\left(\frac{\pi}{2}\right) = 1 \quad \text{and} \quad \cos\left(\frac{\pi}{2}\right) = 0,$$

it immediately follows that

$$\begin{aligned}\cos\left(\frac{\pi}{2} - x\right) &= \cos(x)\cdot 0 + \sin(x)\cdot 1 \\ &= \sin(x)\end{aligned}$$

hence proving the theorem. □

Corollary 6.2.9. *For any $x \in \mathbb{R}$,*

$$\cos(x) = \sin\left(\frac{\pi}{2} - x\right)$$

Proof. Let $\theta = \frac{\pi}{2} - x$. Then using theorem (6.2.8),

$$\sin(x) = \cos\left(\frac{\pi}{2} - x\right)$$

Substitution x for θ gives

$$\sin\left(\frac{\pi}{2} - \theta\right) = \cos(\theta)$$

hence proving the theorem. □

Corollary 6.2.10. *For any $x \in \mathcal{D}$, where $\mathcal{D}$ is the domain for which both $\tan(x)$ and $\cot(x)$ is defined, the following are true:*

$$\cot(x) = \tan\left(\frac{\pi}{2} - x\right)$$

$$\tan(x) = \cot\left(\frac{\pi}{2} - x\right)$$

Proof. Using the definition of $\cot(x)$ and $\tan(x)$:

$$\tan\left(\frac{\pi}{2} - x\right) = \frac{\sin\left(\frac{\pi}{2} - x\right)}{\cos\left(\frac{\pi}{2} - x\right)}$$

Then by theorem (6.2.8) and corollary (6.2.9),

$$\frac{\sin\left(\frac{\pi}{2} - x\right)}{\cos\left(\frac{\pi}{2} - x\right)} = \frac{\cos(x)}{\sin(x)} = \cot(x)$$

Then $\tan(x) = \cot\left(\frac{\pi}{2} - x\right)$ follows immediately. □

Corollary 6.2.11. *For any $x \in \mathcal{D}$, where $\mathcal{D}$ is the domain for which both $\csc(x)$ and $sec(x)$ is defined, the following are true:*

$$\csc(x) = sec\left(\frac{\pi}{2} - x\right)$$

$$sec(x) = \csc\left(\frac{\pi}{2} - x\right)$$

Proof. Using the definition of $\csc(x)$ and $\sec(x)$:

$$\sec\left(\frac{\pi}{2} - x\right) = \frac{1}{\cos\left(\frac{\pi}{2} - x\right)}$$

Then by theorem (6.2.8),

$$\frac{1}{\cos\left(\frac{\pi}{2}-x\right)} = \frac{\sin(x)}{\sin(x)} = \csc(x)$$

Then $\sec(x) = \csc\left(\frac{\pi}{2}-x\right)$ follows immediately. □

6.2.3 Proving Trig Identities

Now with these theorems, let's leverage them to prove a couple of identities in a few examples.

Example 6.2.1.

Let's verify

$$\cos^2(x) - \tan^2(x) = 2 - \sin^2(x) - \sec^2(x)$$

These types of problems are generally easiest if we establish which side we want to attack first. This is usually arbitrary, but for our purposes, I'm going to choose the left side. Using our Pythagorean-esque identities, we establish

$$\cos^2(x) = 1 - \sin^2(x) \tan^2(x) = \sec^2(x) - 1$$

Then by substitution, we get

$$\cos^2(x) - \tan^2(x) = 1 - \sin^2(x) - (\sec^2(x) - 1) = 2 - \sin^2(x) - \sec^2$$

hence verifying our desired identity. ♠

Example 6.2.2.

Let's verify

$$\tan(\alpha+\beta) = \frac{\tan(\alpha) + \tan\beta}{1+\tan(\alpha)\tan(\beta)}$$

Using the theorems (6.2.6) and (6.2.7) and the definition of $\tan(x)$,

$$\tan(\alpha+\beta) = \frac{\sin(\alpha+\beta)}{\cos(\alpha+\beta)} = \frac{\sin(\alpha)\cos(\beta)+\cos(\alpha)\sin(\beta)}{\cos(\alpha)\cos(\beta)-\sin(\alpha)\sin(\beta)}.$$

Then multiplying by

$$\frac{\frac{1}{\cos(\alpha)\cos(\beta)}}{\frac{1}{\cos(\alpha)\cos(\beta)}}$$

gives:

$$= \frac{\sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)}{\cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)} \cdot \frac{\frac{1}{\cos(\alpha)\cos(\beta)}}{\frac{1}{\cos(\alpha)\cos(\beta)}}$$

$$= \frac{\frac{\sin(\alpha)}{\cos(\alpha} + \frac{\sin(\beta)}{\cos(\beta)}}{1 - \frac{\sin(\alpha)}{\cos(\alpha)} \cdot \frac{\sin(\beta)}{\cos(\beta)}}$$

$$= \frac{\tan(\alpha) + \tan\beta}{1 + \tan(\alpha)\tan(\beta)}$$

hence verifying our identity. ♠

Using the identity that we proved in exercise (6.2.2), we actually can go back and prove theorem (6.2.2).

Proof. To prove the theorem, let's prove $\tan(x) = \tan(x + \pi)$. Using the identity in exercise (6.2.2),

$$\tan(x + \pi) = \frac{\tan(x) + \tan\pi}{1 + \tan(x)\tan(\pi)}.$$

Since $\tan(\pi) = \frac{\sin(\pi)}{\cos(\pi)} = 0$,

$$= \frac{\tan(x) + 0}{1 + \tan(x) \cdot 0}$$

$$= \tan(x)$$

hence $\tan(x) = \tan(x + \pi)$ proving $\tan(x)$ is π periodic. Then $\cot(x)$ is π periodic follows trivially. □

Chapter 7

Complex Numbers

7.1 Polor Coordinates

7.1.1 Polar Plane

In this section, I'm going to try to introduce to you a new way of plotting. Throughout your entire math career thus far, you may have only been exposed to one type of coordinate system, the cartesian plane. But here, we are going to explore ways we can plot things on a graph that don't have an x or a y coordinate.

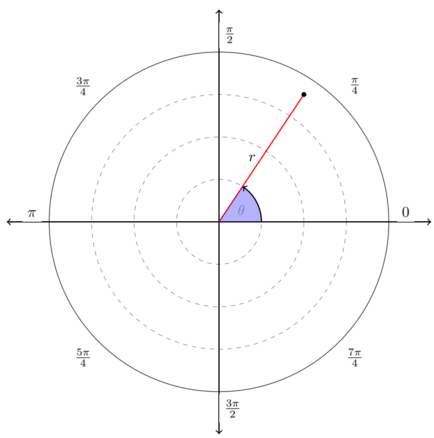


Figure 7.1.1

First, let's define polar coordinates. Like cartesian coordinates, a polar coordinate is also written as an ordered pair, except with the variables r and θ, where r is the distance from the **pole**, or origin, and θ is the angle that a ray connecting the pole and point makes with the polar axis, as seen in figure (7.1.1). Normally we set the polar axis as the positive x axis and let positive angles open counterclockwise. Since an angle of 2π is congruent to an angle of 0, the consequence of using an angle to label a point is that there are infinitely many ordered pairs that we can use to represent a position. For this reason, we typically restrict $\theta \in (-\pi, \pi]$. The choice we make here is arbitrary, so we could have chosen any interval we'd like, so long as every angle on the plane is reached. But even still, there are still multiple choices of ordered pairs that would label the point. Let's illustrate this fact in this next example.

Example 7.1.1.

For this example, let's find all the possible ways we can label the point $\left(2, \frac{\pi}{3}\right)$ given $\theta \in (-\pi, \pi]$. First, I've plotted this point in figure (7.1.2). Because of the way, we defined θ, we go around the circle once, so with a

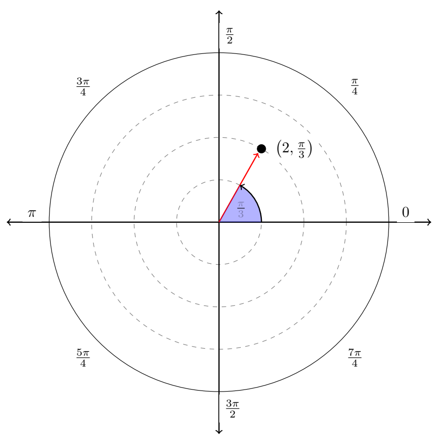


Figure 7.1.2

radius of 2, $\frac{\pi}{2}$ is the only time we hit our desired point. But since there aren't any restrictions on r, we can allow r to be negative. What this would mean is that instead of moving in the direction of the ray defined by the angle, we would go in the opposite direction. As in figure (7.1.3), if we draw a ray opposite of our original ray, and use that angle, we can use that angle to label our point. In this case, we notice the angle of our

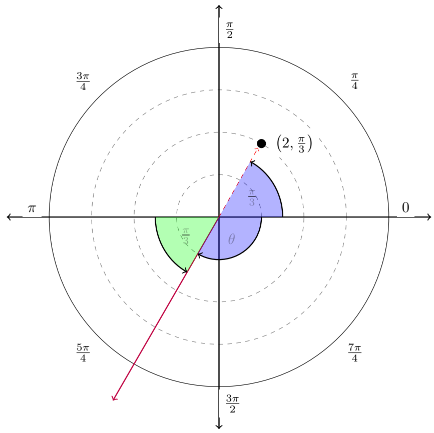


Figure 7.1.3

new ray with the $-x$-axis is $\frac{\pi}{3}$, since the red line and x-axis form vertical angles, by doing a bit of geometry, we find our angle is $-\frac{2\pi}{3}$, hence the other way we can label this point is

$$(-2, \frac{2\pi}{3}).$$

♠

7.1.2 Converting Between Polar and Cartesian

As you might have already suspected, all of this polar stuff is reminiscent of how trig functions. We can leverage trig to convert between cartesian coordinates and polar.

Theorem 7.1.1. *Let (r, θ) be a point in polar coordinates, then the corresponding point in cartesian coordinates can be given as*

$$(r\cos(\theta), r\sin(\theta))$$

This theorem comes directly from the definition of $\cos(x)$ and $\sin(x)$ by multiplying the circle by a factor of r. Using this, we can now try to convert equations plotted in polar, into equations we can plot in Cartesian or the other way around.

Example 7.1.2.

Let's convert $r = 2\sin\theta$ to cartesian to identify the shape of this curve. Multiplying r on both sides gives

$$r^2 = 2r\sin\theta.$$

The utility of this operation might not be immediately clear, but since we know

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta$$

having that extra r term allows us to substitute out the $r\sin\theta$ term fo y. Therefore

$$r^2 = 2y$$

Since we know r is simply the distance to the origin, using Pythagorean's theorem,

$$r = \sqrt{x^2 + y^2}$$

therefore our equation becomes

$$x^2 + y^2 = 2y$$

Then rearranging the equation and completing the square allows us to get

$$x^2 + y^2 - 2y = 0$$

$$x^2 + y^2 - 2y + 1 = 1$$

$$x^2 + (y-1)^2 = 1$$

which we see this equation simply represents a circle centered at $(0, 1)$. ♠

Example 7.1.3.

In this example, let's find the corresponding cartesian form of the equation

$$r = \frac{6}{\cos\theta + 3\sin\theta}.$$

Since the trig functions in the denominator lack the factor of r that allows us to easily substitute them for x and y, we can divide both sides by r to allow for this.[1] Therefore

$$1 = \frac{6}{r\cos\theta + 3r\sin\theta}$$

$$1 = \frac{6}{x + 3y}$$

$$x + 3y = 6$$

hence we conclude this is a line with a slope of $-\frac{1}{3}$ with a y-intercept of 2. ♠

7.2 Complex Numbers

7.2.1 Polar Form

Similar to coordinates in a plane, complex numbers also have a corresponding polar form. But before we examine that, we must define an operation that will be key for defining polar complex numbers.

I've mentioned that the complex numbers are not an ordered set, but similar to the real numbers, complex numbers have this notion of distance. To describe the distance, we define the following:

Definition 7.2.1. *Let the absolute value, (modulus) be defined as*

$$|z| = \sqrt{z\bar{z}}$$

Since we know that $z\bar{z}$ is always positive, from a previous lesson, so this value will always be positive. This definition, in reality, is arbitrary, there are many other ways we could have chosen to define this, as long as they satisfy a certain set of criteria. But we chose this definition since it allows our complex plane to be a plane and agrees with our other definitions of distance in a plane.

Now we have a method of finding the distance from any complex number to the origin, we also need an operation that measures the angle a complex number makes with the positive real axis.

[1] Since the only way for r to be 0 is if the numerator is 0, we know this is not possible since the numerator is a non-zero constant.

Definition 7.2.2. *Let* $\arg : \mathbb{C} \to \mathbb{R}$ *be defined as a multivalued function where* $\arg(z)$ *returns the argument (angle with respect to the positive real axis) of* z*.*

We technically could've used $\tan^{-1}\left(\frac{b}{a}\right)$, but due to this only being defined for $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, this is no good for numbers that are in the second or third quadrant. So in complex analysis, we generalize this operation with $\arg(z)$. Now notice this is a **multivalued** function, so this is not a function that we defined in Lesson 2. Instead, for any given complex number, it outputs an infinite number of angles.[2] For this reason, we typically restrict the function to only output on the interval $(-\pi, \pi]$. We call this the **principle argument** and typically denote this function as $\text{Arg}(z)$.

Now we can define polar complex numbers. Since we know we can label a complex number in polar form with the ordered pair $(|z|, \text{Arg}(z))$, using theorem (7.1.1),

$$z = |z|(\cos(\text{Arg}(z)) + i\sin(\text{Arg}(z)))$$

Sometimes we shorthand this notation and simply write

$$z = |z|\text{cis}(\text{Arg}(z))$$

where

$$\text{cis}(\theta) = \cos(\theta) + i\sin(\theta)$$

Example 7.2.1.

In this example, I'd like to show you how to multiply a complex number in polar form. We know the general form of a complex number in polar form is

$$z_1 = r_1\text{cis}(\theta_1)$$

$$z_2 = r_2\text{cis}(\theta_2)$$

where

$$r_1 = |z_1| \quad \text{and} \quad \theta_1 = \text{Arg}(z_1)$$

[2] This makes sense since as we talked about, if $\theta \in \mathbb{R}$, then there are an infinite number of ways to label an angle.

$$r_2 = |z_2| \quad \text{and} \quad \theta_2 = \text{Arg}(z_2)$$

Then

$$z_1 z_2 = (r_1 \text{cis}(\theta_1)) \cdot (r_2 \text{cis}(\theta_2))$$

$$= r_1 r_2 (\cos(\theta_1) + i\sin(\theta_1))(\cos(\theta_2) + i\sin(\theta_2))$$

$$= r_1 r_2 (\cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_1) + i\cos(\theta_1)\sin(\theta_2) + i\sin(\theta_1)\cos(\theta_2))$$

Then using the cosine and sine sum identities,

$$= r_1 r_2 (\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2))$$

$$= r_1 r_2 \text{cis}(\theta_1 + \theta_2)$$

♠

As you can see in the last example, cis(θ) sort of acts like an exponential, in the sense that multiplying them together will add its arguments. The reason for this is that it turns out cis(x) is just how we can define the complex exponential.

Theorem 7.2.1 (Euler's Theorem). *Let $\theta \in \mathbb{R}$, then*

$$e^{i\theta} = \cos(\theta) + i\sin(\theta)$$

Proof. Postponed indefinitely. □

Now, we see, we can write the polar form of a complex number as

$$z = |z|e^{i\arg(z)}.$$

Now with this, we can use the properties that we defined for exponentials to perform complex multiplication and division.

Theorem 7.2.2. *If $z \in \mathbb{C}$, then*

$$\bar{z} = |z|e^{-i\arg(z)}$$

Proof. Let $r = |z|$ and $\theta = \arg(z)$, where θ is any choice of $\arg(z)$. Then

$$z = re^{i\theta} = r(\cos(\theta) + i\sin(\theta))$$

For $\bar{z}$,

$$\bar{z} = r(\cos(\theta) - i\sin(\theta))$$

Then using the odd and even properties of trig functions:

$$\bar{z} = r(\cos(-\theta) + i\sin(-\theta))$$

$$= re^{-i\theta}$$

□

7.2.2 Roots of Numbers

Now using complex exponentials, let's define root operations in terms of complex numbers. As I've stated before, taking a root in complex numbers will result in multiple outputs, no matter if the power is odd or even. Let's find out how to compute these roots in general.

Example 7.2.2.

Let's compute in general the nth root of z. We know for r, θ,

$$z = re^{i\theta}$$

Therefore taking the nth root gives

$$\sqrt[n]{z} = (z)^{\frac{1}{n}} = r^{\frac{1}{n}}(e^{i\theta})^{\frac{1}{n}} = r^{\frac{1}{n}} \exp\left(\frac{i\theta}{n}\right)$$

Since angles 2π apart are congruent,

$$= r^{\frac{1}{n}} \exp\left(\frac{i\theta + 2i\pi k}{n}\right) \quad \text{for } k \in \mathbb{Z}.$$

Then taking only the positive result from $r^{\frac{1}{n}}$[3] and setting $s = r^{\frac{1}{n}}$

$$= s\exp\left(\frac{i\theta}{n} + \frac{2i\pi k}{n}\right)$$

Notice here, that although we defined $k \in \mathbb{Z}$, it turns out equivilently, we could have just defined $k \in \{0, 1, ..., n-1\}$, since any other $k \in \mathbb{Z}$ would

[3] Because turns out the others are redundant.

just produce overlapping angles and would be redundant. ♠

Example 7.2.3.

With the formula we computed before, let's compute the cube root of $z = 1 + i$. First, we compute the modulus of z:

$$|z| = \sqrt{(1+i)(1-i)} = \sqrt{2}$$

Then for the principle argument:

$$\text{Arg}(z) = \tan^{-1}(1) = \frac{\pi}{4}$$[4]

Therefore

$$z = \sqrt{2}\exp\left(\frac{i\pi}{4}\right)$$

Then using the formula in the last example:

$$\begin{aligned}\sqrt[3]{z} &= \sqrt[6]{2}\exp\left(\frac{i\pi}{4}\cdot\frac{1}{3} + \frac{2i\pi k}{3}\right) \quad \text{for } k \in \mathbb{Z} \\ &= \sqrt[6]{2}\exp\left(\frac{i\pi}{12} + \frac{2}{3}i\pi k\right).\end{aligned}$$

Then, the 3 unique solutions are

$$\left\{\sqrt[6]{2}\exp\left(\frac{i\pi}{12}\right), \sqrt[6]{2}\exp\left(\frac{3i\pi}{4}\right), \sqrt[6]{2}\exp\left(\frac{17i\pi}{12}\right)\right\}$$

♠

7.2.3 Complex Logarithms

Example 7.2.4.

In this example, I'm going to compute a complex logarithm in general. In terms of notation, we typically use $\ln(x)$ to define a complex natural logarithm, and in the context of complex numbers, we use $\log(z)$ to denote a complex natural logarithm.[5] Thus since $z = re^{\imath\theta}$, by using the definition

[4] Of course we check that $tan^{-1}(x)$ output in the right quadrant, or we would need to modify the output.

[5] We generally don't talk about logarithms for complex variables with arbitrary bases, since turns out, those are generally less useful, and can be replaced simply by

of the logarithm and its properties:

$$\log(z) = \log(re^{\imath\theta}) = \ln(r) + \log(e^{\imath\theta}) = \ln(r) + i\theta.$$

Since

$$\theta \cong \theta + 2\pi k \quad \text{for } k \in \mathbb{Z},$$

it turns out $\log(z)$ is a multivalued functions outputting

$$\log(z) = \ln(r) + i\theta + 2i\pi k.$$

More generally,

$$= \ln(r) + i\arg(z)$$

or if taking the principle value:

$$\operatorname{Log}(z) = \ln(r) + i\operatorname{Arg}(z)$$

♠

Example 7.2.5.

Now using the above formula, let's compute $\operatorname{Log}(-\sqrt{3} + i)$. First finding the modulus of z we get

$$|z| = \sqrt{z\bar{z}} = \sqrt{(-\sqrt{3} + i)(-\sqrt{3} - i)} = 2$$

Then computing the principle argument:

$$\operatorname{Arg}(z) = \tan^{-1}\left(-\frac{1}{\sqrt{3}}\right) = \frac{5\pi}{6}$$[6]

Therefore,

$$\operatorname{Log}(z) = \ln(2) + \frac{5i\pi}{6}$$

or if we expand to all outputs of $\log(z)$,

$$\log(z) = \ln(2) + \frac{5i\pi}{6} + 2i\pi k \quad \text{for } k \in \mathbb{Z}$$

dividing by the respective base.

[6] Make sure we chose the domain for $\tan^{-1}(\theta)$ that outputs in the right quadrant, don't just plug this into a calculator.

♠

Chapter 8

Rates of Change

8.1 Rates of Change

8.1.1 Linear Approximations

Today, we are going to discuss a pseudo-calculus topic. We are going to try to generalize the idea of a slope onto functions that aren't just straight lines.

As we know from geometry, we can find the equation of a line using only two points. This will become very important since we will use the idea of creating **linear approximations** of functions. Unfortunately, due to the curvy nature of general functions, we cannot directly apply the idea of the slope, and instead, we have to approximate its slope at a point with a line. Let's see how this works with an example.

Example 8.1.1.

In this example, I'd like to find the linear approximation for the function

$$f(x) = 5x^3 + 2x + 3$$

for the point $(1, f(1))$. We know from previous math classes, that if we had two points, we could generate a line through those points and find its slope. Therefore, we must find two points the point $(1, f(1))$, where we can draw a line through it.

For now, let's use the point $(0, f(0))$ and $(2, f(2))$. To find the line that goes through these points, we use the formula

$$\frac{f(b)-f(a)}{b-a} = \frac{f(2)-f(0)}{2-0} = \frac{47-3}{2} = 22$$

Then using point-slope form, we get

$$y - f(0) = y - 3 = 22x$$

Hence converting to slope-intercept form, we get the line

$$y = 22x + 3.$$

Then using figure (8.1.1), we can see we did a pretty good job here approximating, but I think we can do better.

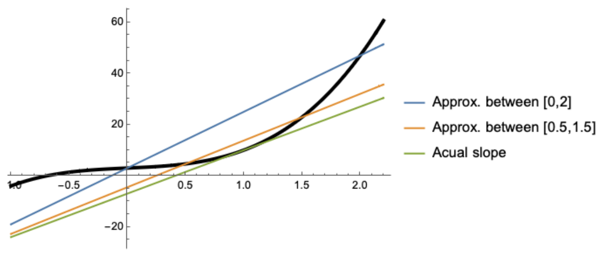


Figure 8.1.1

Let's shrink our interval so we take the line that runs between the points $\left(\frac{1}{2}, f\left(\frac{1}{2}\right)\right)$ and $\left(\frac{3}{2}, f\left(\frac{3}{2}\right)\right)$. Using the same formula as before, we get

$$\frac{f\left(\frac{3}{2}\right)-f\left(\frac{1}{2}\right)}{1.5-0.5} = \frac{\frac{183}{8}-\frac{37}{8}}{1} = \frac{73}{4}$$

Therefore our line can be represented as

$$y = \frac{73}{4}x - \frac{9}{2}$$

Then looking back at figure (8.1.1), the approximation is way better and a lot closer to what the actual slope is. ♠

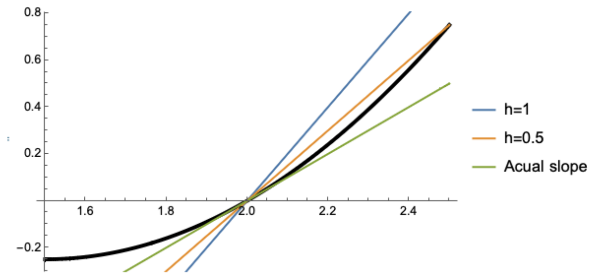


Figure 8.1.2

Example 8.1.2.

Now let's try a different example. Let's try to make linear approximations for the function

$$f(x) = x^2 - 3x + 2$$

for the point $(2, f(2))$. This time, however, instead of taking the interval around this point, let's anchor one point at $x = 2$ and set our second point h away from our initial. Writing the general formula for the slope of this line gives

$$\frac{f(x+h) - f(x)}{h}.$$ [1]

First, let's let $h = 1$. Then we find the slope is

$$\frac{f(3) - f(2)}{1} = 2$$

Therefore computing the line through these two points gives

$$y = 2x - 4$$

Plotting these lines like in figure (8.1.2), we see that our approximation is pretty, good, but we do better with smaller and smaller values of h. But if we notice, using $h = 0$ gives us a divide by zero so finding this value is going to be a bit tricky. Thankfully, we've learned a handy trick for computing values that are masked by dividing by zero errors. ♠

[1] This is the preferred method for finding linear approximations since very easily, we can take this approximation to 0, which will give us the actual slope of a point on an arbitrary function.

8.1.2 Tangent Line Approximations

In this section, let's try to take the approximation to 0. As seen in previous examples, as we shrink the interval of the approximation, the line gets closer and closer to its actual approximation, or its **tangent line** approximation. Let's try to create a tangent line approximation in the following example.

Example 8.1.3.

Let's find the tangent line approximation for the function

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

on the point $(1, f(1))$. We know the formula for the approximate slope of a point generally is given by

$$\frac{f(x+h) - f(x)}{h}.$$

Now let's compute this as $h \to 0$ in general in a limit.

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{((x+h)^2 - (x+h)) - (x^2 - x)}{h}$$

$$= \lim_{h\to 0} \frac{(x^2 + 2xh + h^2 - (x+h)) - (x^2 - x)}{h}$$

$$= \lim_{h\to 0} \frac{2xh + h^2 - h}{h}$$

Even though $h = 0$ causes a divide by zero, as we take the limit to a value, we eliminate that point from our domain, as in the definition of a limit. In this case, we eliminate $h = 0$ from our domain, hence dividing out the h here is valid. Therefore

$$= \lim_{h\to 0} 2x - h + 1$$

Now since this is a nice smooth function, by plotting it, we can easily tell the limit as $h \to 0$ is just whatever the function outputs, hence

$$= 2x + 1$$

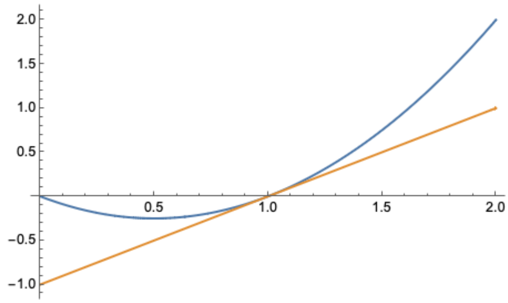


Figure 8.1.3

This is our tangent line approximation, and if we substitute in $x = 1$, we get the slope at this point is

$$2(1) - 1 = 1,$$

hence using the same technique as before, we get the tangent line is

$$y = x - 1$$

which we can see plotted in figure (8.1.3) ♠

Now, let's compute the slope of the tangent line approximation in general. But before doing so, we're going to need some help from the following theorem.

Theorem 8.1.1 (Binomial Theorem[2]). *Let* $n \in \mathbb{W}$, *then*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k y^{n-k}$$ [3]

Proof. Postponed indefinitely. □

Now armed with this, let's compute this formula in general.

[2] You may have seen this theorem with Pascal's triangle.
[3] $\binom{n}{k}$ this means combinations. It's computed by the formula $\frac{n!}{k!(n-k)!}$.

Example 8.1.4.

First, let's compute the formula for a monomial, ax^n, where $a \in \mathbb{R}$ and $n \in \mathbb{W}$. Using the limit use used in example (8.1.3),

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{a(x+h)^n - ax^n}{h}$$

Using some theorems we established about limits in Lesson 4, assuming both limits exist,

$$= \lim_{h\to 0} a \cdot \lim_{h\to 0} \frac{(x+h)^n - x^n}{h}$$

Since by plotting $y = a$ we know

$$\lim_{h\to 0} a = a$$

the only remaining question is if the right-hand limit exists. But can confirm that by continuing our computation. Then expanding the binomial with theorem (8.1.1), we get

$$a \lim_{h\to 0} \frac{(x+h)^n - x^n}{h} = a \lim_{h\to 0} \frac{\sum_{k=0}^{n} \binom{n}{k} x^k h^{n-k} - x^n}{h}$$

$$= a \lim_{h\to 0} \frac{\sum_{k=0}^{n-1} \binom{n}{k} x^k h^{n-k}}{h}$$

Then since here, everything is a factor of h and we can eliminate $h = 0$ from our domain, dividing out h gives

$$= a \lim_{h\to 0} \left(\sum_{k=0}^{n-1} \binom{n}{k} x^k h^{n-k-1} \right)$$

Then using the same idea as in example (8.1.3), we can substitute $h = 0$, therefore every term that is a factor of h vanishes in our sum. Therefore for $n > 0$,

$$= a \binom{n}{n-1} x^{n-1}$$

Since

$$\binom{n}{n-1} = \frac{n!}{(n-1)!(1!)} = n$$

Therefore the general formula for the slope of a tangent line approximation is

$$anx^{n-1}.$$

Then for $n = 0$, since we start our summation at 0, nothing is output by the sum, hence our rate of change is 0. [4] ♠

In example (8.1.4), we only computed the slope if the polynomial was one term. To generalize to all polynomials, we need to know how these slopes add.

Theorem 8.1.2. *Suppose f, g are both polynomials. If f', g' are functions that produce the slope of their tangent line approximations, then the slope of the tangent line approximation of their sum is given as*

$$(f+g)' = f' + g'.$$

Proof. Computing $(f+g)'$ gives

$$(f+g)' = \lim_{h \to 0} \frac{(f+g)(x+h) - (f+g)(x)}{h}$$

$$= \lim_{h \to 0} \frac{(f(x+h) - f(x)) + (g(x+h) - g(x))}{h}$$

Then by limits laws, we know

$$= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} + \lim_{h \to 0} \frac{g(x+h) - g(x)}{h}.$$

Of course, to make this conclusion, we need to know these limits exist, but since we know

$$f'(x) = \frac{f(x+h) - f(x)}{h} \quad \text{and} \quad g'(x) = \frac{g(x+h) - g(x)}{h}$$

this is the case, therefore

$$= f' + g'$$

□

Now with theorem (8.1.2), we can compute the slope of the tangent line with any arbitrary polynomial. Let's try this with the following example.

[4] This is called the power rule in calculus. We can generalize this to all rational exponents if we expand our binomial theorem to become the binomial series.

Example 8.1.5.

Let's compute a general formula for the tangent line for the polynomial

$$P(x) = x^5 + 2x^2 + x - 5$$

Using the formula we computed in example (8.1.4), we can compute the slope of the tangent line for each monomial. Doing so gives us

$$(x^5)' = 5x^4$$

$$(2x^2) = 4x$$

$$(x^1)' = 1$$

$$(-5x^0)' = 0$$

Then using theorem (8.1.2),

$$P'(x) = 5x^4 + 4x + 1$$

Therefore, for any point $(x, P(x))$, the tangent line in point-slope form is

$$y - P(x) = P'(x)(w - x)$$

where w is the independent variable. In figure (8.1.4), you can see a

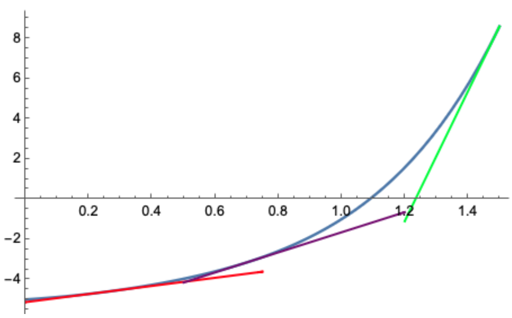


Figure 8.1.4

couple of them plotted. ♠

8.2 Plotting Polynomials

8.2.1 Local Extrema

Definition 8.2.1. *Define $f : \mathcal{D} \to \mathbb{R}$, where $\mathcal{D} \subseteq \mathbb{R}$. If there exists an open interval $(a, b) \subseteq \mathcal{D}$ such that $c \in (a, b)$, is the maximum value, then c is a* ***local maximum****.*

Definition 8.2.2. *Define $f : \mathcal{D} \to \mathbb{R}$, where $\mathcal{D} \subseteq \mathbb{R}$. If there exists an open interval $(a, b) \subseteq \mathcal{D}$ such that $c \in (a, b)$, is the minimum value, then c is a* ***local minimum****.*

The mathematical jargon might be difficult to navigate, but all the two definitions is saying is that if a point on a function function is a maximum on an interval, where this point is not on the boundary, then this point is a local maximum. Or as in figure (8.2.1)

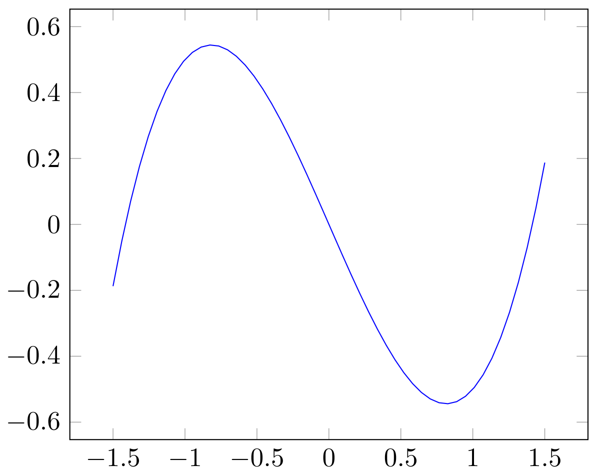


Figure 8.2.1