Mathematics Of Doing, Understand, Learning, and Educating Secondary Schools

MODULE(S^2): Algebra for Secondary Mathematics Teaching

Adapted for MODULE(S^2)

Version Spring 2018

The Mathematics Of Doing, Understand, Learning, and Educating Secondary Schools (MODULE(S^2)) project is partially supported by funding from a collaborative grant of the National Science Foundation under Grant Nos. DUE-1726707,1726804, 1726252, 1726723, 1726744, and 1726098. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Todo list

Contents

Part I

Introduction to Fields

1 Fields and Other Algebraic Structures (Week 1 2.5 hours)

Overview

Content

Field, implicitly defined as a relation which assigns elements of $\mathbb{N}$ to its factors; used to examine subsets, mathematical statements and their negations, properties of $\mathbb{R}$ and $\mathbb{Z}$, and to engage in mathematical practices.

Subfield, **superset**, **strict subset**, and **strict superset**; **equality of sets** A and B, defined as $A \subseteq B$ and $B \subseteq A$.

Field extension, defined as those which can be evaluated as true or false; and

additive identity of mathematical statement S, defined as a statement which is false if and only if S is true.

additive inverse of mathematical statement S, defined as a statement which is false if and only if S is true.

multiplicative identity of mathematical statement S, defined as a statement which is false if and only if S is true.

multiplicative inverse of mathematical statement S, defined as a statement which is false if and only if S is true.

ordered field

Summary

In this section we introduce the notion of a field. The emphasis is on justifying well known facts from high school algebra using the formal properties of fields. Etc.

Acknowledgements. Sources...

Materials.

- List handouts, needed materials.

Opening inquiry: Finite Fields

We are interested in mathematical structures with operations of addition $(+)$ and multiplication $(\cdot)$ that satisify the following properties:

(A) $a + b = b + a$ and $a \cdot b = b \cdot a$ (commutative laws)

(B) $(a + b) + c = a + (b + c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (associative laws)

(C) $a \cdot (b + c) = a \cdot b + a \cdot c$ (distributive law)

(D) There are distinct elements, called 0 and 1, such that $a + 0 = a$ and $a \cdot 1 = a$ for all a.

(E) For each a there is an element b such that $a + b = 0$ and if $a \neq 0$, there is an element c such that $a \cdot c = 1$.

Structures which satisfy these properties are called **<u>fields</u>**. We now define a field called $\mathbb{Z}_5$. We will avoid the complexities of a formal definition of $\mathbb{Z}_5$ and simply assert that $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$. The operations of addition $(+)$ and multiplication $(\cdot)$ are defined "modulo 5." That means we find the regular sum or product and then divide by 5 and take the remainder as our answer. For example, in $\mathbb{Z}_5$:

$$3 + 4 = 7$$
$$7 \div 5 = 1 \text{ R } 2.$$

Since the sum of 3 and 4 is 7 and the remainder when we divide 7 by 5 is 2, we conclude that in $\mathbb{Z}_5$, $3 + 4 = 2$. Using that procedure, fill in the following addition table:

+	0	1	2	3	4
0					
1					
2					
3					2
4					

1. Find a number $a \in \mathbb{Z}_5$ with the property that $a + 2 = 2 + a = 0$. The number a that you find is called the **<u>additive inverse</u>** of 2 in $\mathbb{Z}_5$. That is, it is the value of -2 in $\mathbb{Z}_5$.
   (a) $-2 =$ ______ in $\mathbb{Z}_5$ because $2 +$ ______ $=$ ______ $+ 2 =$ ______.
2. Find the value of -1, -3, and -4 in $\mathbb{Z}_5$.
   (a) $-1 =$ ______ in $\mathbb{Z}_5$ because $1 +$ ______ $=$ ______ $+ 1 =$ ______.
   (b) $-3 =$ ______ in $\mathbb{Z}_5$ because $3 +$ ______ $=$ ______ $+ 3 =$ ______.
   (c) $-4 =$ ______ in $\mathbb{Z}_5$ because $4 +$ ______ $=$ ______ $+ 4 =$ ______.

We follow the analogous process for multiplication. For example, $4 \cdot 4 = 16$. When we divide 16 by 5 we get 3 with a remainder of 1. So we conclude that in $\mathbb{Z}_5$, $4 \cdot 4 = 1$. Using that procedure, fill in the following multiplication table:

·	0	1	2	3	4
0					
1					
2					
3					
4					1

Since $4 \cdot 4 = 1$ in $\mathbb{Z}_5$, we conclude that 4 is its own **multiplicative inverse** in $\mathbb{Z}_5$. That is, $4^{-1} = 4$ in $\mathbb{Z}_5$!

3. Find the value of 2^{-1} in $\mathbb{Z}_5$ and the value of 3^{-1} in $\mathbb{Z}_5$.
   (a) $2^{-1} =$ ______ in $\mathbb{Z}_5$ because $2 \cdot$ ______ $=$ ______ $\cdot 2 =$ ______
   (b) $3^{-1} =$ ______ in $\mathbb{Z}_5$ because $3 \cdot$ ______ $=$ ______ $\cdot 3 =$ ______

One can see by inspection of the addition and multiplication tables for $\mathbb{Z}_7$ that every element of $\mathbb{Z}_7$ has an additive inverse, and every nonzero element of $\mathbb{Z}_7$ has a multiplicative inverse. That is, $\mathbb{Z}_7$ satisifies (E) above. It is also easy to see by inspecting those tables that for all $a \in \mathbb{Z}_7$ $0 + a = a + 0 = a$ and $a \cdot 1 = 1 \cdot a = a$. That is, property (D) above is true in $\mathbb{Z}_7$.

4. Verify that in $\mathbb{Z}_7$ the following number sentences are true. Be prepared to share you work with the class.
   - $3 \cdot (1 + 4) = 3 \cdot 1 + 3 \cdot 4$
   - $5 \cdot (2 + 6) = 5 \cdot 2 + 5 \cdot 7$
5. Make up two number sentences with the same form using numbers $0 - 6$ and verify that they are true in $\mathbb{Z}_7$.
6. Do you believe that the distributive law (C) holds in $\mathbb{Z}_7$? Have you proved it? Exactly how many different instances of the distributive law are there in $\mathbb{Z}_7$?
7. Verify that in $\mathbb{Z}_7$ the following number sentences are true. Be prepared to share you work with the class.
   - $3 + (1 + 4) = (3 + 1) + 4$ and $5 + (2 + 6) = (5 + 2) + 6$.
   - $3 \cdot (1 \cdot 4) = (3 \cdot 1) \cdot 4$ and $5 \cdot (2 \cdot 6) = (5 \cdot 2) \cdot 6$.
8. Do you believe that the associative laws (B) hold in $\mathbb{Z}_7$? Have you proved it? Exactly how many different instances of each of the two distributive laws are there in $\mathbb{Z}_7$?
9. Verify that in $\mathbb{Z}_7$ the following number sentences are true. Be prepared to share you work with the class.
   - $3 + 4 = 4 + 3$ and $5 + 6 = 6 + 5$.
   - $3 \cdot 4 = 4 \cdot 5$ and $5 \cdot 6 = 6 \cdot 5$.
10. Do you believe that the commutative laws (A) hold in $\mathbb{Z}_7$? Have you proved it? Exactly how many different instances of each of the two commutative laws are there in $\mathbb{Z}_7$?
11. Based on your work above, does $\mathbb{Z}_7$ satisfy properties (A)-(E)? Is it a field?
12. Create addition and multiplication tables for $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ and $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ and answer appropriate versions of questions 1-11 for these two structures. Once you have determined whether or not $\mathbb{Z}_4$ and $\mathbb{Z}_4$ are fields, form a conjecture for which values of n is $\mathbb{Z}_n$ a field?
13. Test your conjecture with another value of n and be prepared to discuss your conjecture with the class.

Instructor note. Distribute handout with this question. The activity should be done in groups of 3-4. As students work on it, circulate and listen to the questions and comments they make. They may say and do things that will lead into a discussion about additive inverses, multiplicative inverses and other properties of fields.

Groups may conjecture that $\mathbb{Z}_n$ is a field whenever n is odd. Be prepared to discuss this conjecture.

Introduction to Fields

In this section we will begin our study of **fields**. You've already encountered fields in your mathematical studies: the set of rational numbers $\mathbb{Q}$ and the set of real numbers $\mathbb{R}$ are fields, as is the set of complex numbers $\mathbb{C}$. In the opening inquiry to this section, you saw that $\mathbb{Z}_n$ is a field for some values of n. The sets $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{Z}_n$ are different in many ways, but here we will focus on the ways in which they are similar.

The study of fields is motivated by the desire to provide justification for the steps we use in solving equations with the two arithmetic operations of addition and multiplication. What rules do we need to follow in order to solve such equations? When can we guarantee that such an equation will always have a solution?

The goal of this section is to explore the foundations of the number systems we typically use for solving equations. This exploration will allow us to provide well grounded, thorough, and pedagogically appropriate justifications for the steps we use in algebra every day to solve equations. But it will also allow us to explore exciting extensions of our ordinary mathematical practices and allow us to connect equation solving to geometry in an intriguing way.

To begin, consider the equation

$$3x + 8 = 14.$$

It's not hard to see that the solution to this equation is $x = 2$: $3(2) + 8 = 14$. Let's us solve this equation step-by-step, justifying each step along the way. First we will subtract 8 from both sides:

$$(3x + 8) - 8 = 14 - 8.$$

(Note that we could also view this as adding -8 to both sides. The number -8 is known as the **additive inverse** of 8.) Applying the associative law on the left-hand side gives

$$3x + (8 - 8) = 6.$$

We know that $8 - 8 = 0$ so we have

$$3x + 0 = 6.$$

The number 0 is an **additive identity**. That means adding 0 returns the value we added it to. So we have

$$3x = 6.$$

We now multiply each side by $1/3$ to obtain

$$\frac{1}{3}(3x) = \frac{1}{3} \cdot 6.$$

Multiplication is associative, so we can write this as

$$\left(\frac{1}{3} \cdot 3\right) x = 2.$$

The number $1/3$ is the **multiplicative inverse** of 3, meaning that $\frac{1}{3} \cdot 3$ is equal to the **multiplicative identity**; that is, $\frac{1}{3} \cdot 3 = 1$. Thus we have

$$1x = 2.$$

The number 1 is the **multiplicative identity** meaning that $1x = x$. So we conclude that

$$x = 2.$$

Let us analyze this situation more carefully. First note that the equation $3x + 8 = 14$ uses two operations, called addition and multiplication. (Subtraction can always be defined in terms of addition, and division can be defined in terms of multiplication.) We used some familiar properties of addition and multiplication such as associativity of addition and multiplication.

Above we multiplied by $1/3$ at point in the solution. Since $1/3$ is a rational number, we say that we solved this equation "over the rationals." But, notice that in this example we didn't really need to do this. Next we give a solution to the equation $3x + 8 = 14$ "over the integers." We begin the same way:

$$\begin{aligned} 3x + 8 &= 14 \\ (3x + 8) - 8 &= 14 - 8 \\ 3x + (8 - 8) &= 6 \\ 3x + 0 &= 6 \\ 3x &= 6. \end{aligned}$$

Next we observe that $6 = 3(2)$ so we have

$$3x = 3(2).$$

One can prove that in the integers that if a, b, and c are integers and $ab = ac$ then $b = c$. Using just that fact, we can conclude that

$$x = 2.$$

- Prove that if a, b, and c are integers and $ab = ac$ then $b = c$. Remember - division is not allowed, we want to do this proof entirely in the integers.
- Can you solve the equation $3x + 8 = 14$ over the natural numbers? (Here you're not allowed to use additive inverses!)

The equation $3x + 8 = 14$ can be solved over the rationals, integers, or natural numbers, but notice that the equation $3x + 8 = 10$ cannot be solved over the integers or natural numbers. The solution $x = 2/3$ is a rational number and is not a natural number or integer. Notice that so long as a, b and c are always rational numbers, $ax + b = c$ will always have a rational solution. The same goes for equations with real or complex coefficients. On the other hand, if a, b, and c are integers, that does not guarantee that $ax + b = c$ will have an integer solution. We want to determine all of the properties necessary on a set of numbers for an equation such as $ax + b = c$ to always have a solution in that set. That is, we want to figure out what makes a set of numbers like $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ in this regard. We will call such a set of numbers a field.

We begin with some terminology.

We call 0 an additive identity because for any number n, $n + 0 = 0 + n = n$. The number 0 is also an additive identity in the set of complex numbers, although more formally it is $0 + 0i$. A corresponding notion for multiplication exists - the multiplicative identity.

- Consider the collection of all 2×2 matrices whose entries are real numbers. Write down the additive identity of this set.
- How would you define the general notion of a **multiplicative identity**? What is a multiplicative identity in $\mathbb{Q}$?
- Is there a multiplicative identity for the set of all 2×2 matrices with real entries?

Once we have a notion of an additive identity, we can define the notion of an additive inverse. We say that a number b is an additive inverse of a number a if and only if $a + b = b + a = 0$.

How would you define the notion of a **multiplicative inverse**? Give an example of a number a and its multiplicative inverse b.

A **field** $\mathbb{F}$ is a collection of mathematical objects (possibly numbers, matrices, functions, etc.) with two operations, called addition ($+$) and multiplication ($\cdot$), in which we can always solve an equation of the form

$$ax + b = c$$

where $a, b, c\mathbb{F}$ and $a \neq 0$. The properties we need to make this happen are given in the following definition.

Definition 1.1. A field $\mathbb{F}$ is a nonempty set together with two operations addition $+$ and multiplication $\cdot$ which satisfy the following properties, called the field axioms:

(A) $a + b = b + a$ and $a \cdot b = b \cdot a$ (commutative laws)

(B) $(a + b) + c = a + (b + c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (associative laws)

(C) $a \cdot (b + c) = a \cdot b + a \cdot c$ (distributive law)

(D) There are distinct elements, called 0 and 1, such that $a + 0 = a$ and $a \cdot 1 = a$ for all a.

(E) For each a there is an element b such that $a + b = 0$ and if $a \neq 0$, there is an element c such that $a \cdot c = 1$.

Of course, you have seen fields before: the rational numbers $\mathbb{Q}$ and the real numbers $\mathbb{R}$ are both fields under their usual operations of addition and multiplication.

More on Identities and Inverses

We all know that in the rational numbers there is only one additive identity: the number 0. But could it be that there is a field with more than one additive identity? We have the following proposition:

Proposition 1.2. *In any field $\mathbb{F}$, then additive identity is unique.*

Proof. Suppose that we have additive identities 0 and z in $\mathbb{F}$. Since 0 is an additive identity, we know that

$$0 + z = z.$$

But since z is also an additive identity, we also know that

$$0 + z = 0.$$

So, we have that

$$z = 0 + z = 0.$$

This proves that the additive identity in any field is unique. □

There are a couple of observations to make about this proof. First, a good general strategy for proving that something is unique is to assume that there are two of them and then prove that they are equal. If needed, you can also assume that your two proposed objects are not equal and derive a contradiction, but notice that we did not need to do that in the proof above. Second, observe that besides using the definition of an additive identity, the only other property we used to prove the proposition above is that addition is commutative.

Since the additive identity in any field is unique, we will always use the usual symbol 0 to represent it, unless we have a good reason not to.

Use the proof above as a model to show that in any field the multiplicative identity is unique.

Similarly, since the multiplicative identity in any field is unique, we will almost always use the usual symbol 1 to represent it.

There is a similar fact to observe with respect to additive and multiplicative inverses. For example, there is only one rational number whose sum with $-\frac{1}{2}$ is 0, namely $\frac{1}{2}$. Similarly, there is only one rational number whose product with $-\frac{1}{2}$ is 1, namely -2. Above you may have noticed that we said "an additive inverse" instead of "the additive inverse," and "a multiplicative inverse" instead of "the multiplicative inverse." We didn't want to suggest that they are unique, and were hoping that a reader might notice our strange locution and question it. But now we are at a point where we are happy to admit that additive and multiplicative inverses are, in a sense, unique:

Proposition 1.3. *If $\mathbb{F}$ is a field and $a \in \mathbb{F}$, then its additive inverse is unique to it.*

Proof. Suppose that $\mathbb{F}$ is a field and that $a \in \mathbb{F}$. We want to prove that there is only one element $b \in \mathbb{F}$ so that

$$a + b = b + a = 0.$$

To this end, suppose that there are two such elements $b, c \in \mathbb{F}$. Then we have both:

$$a+b=b+a=0$$
$$a+c=c+a=0$$

Consider the sum $b+a+c$. On one hand we have

$$b+a+c=(b+a)+c=0+c=c.$$

On the other hand we have

$$b+a+c=b+(a+c)=b+0=b.$$

Thus we conclude that $c=b$ and that every element in a field has a unique additive inverse. □

If b is the additive inverse of a we write $b=-a$. Note that $-a$ may be positive or negative. For example, the additive inverse of 4 is -4, but the additive inverse of -5 is 5. This brings up an important point. When people see "$-a$" it is common to read it as "minus a," or "negative a." The least common thing for people to say is "the additive inverse of a." But, that's what we want you to do because it really helps to keep things straight as, for example, in the following proposition.

Proposition 1.4. *Suppose $a, b \in \mathbb{F}$. Then*

(a) $-(-a)=a$

(b) $-a=(-1)a$

(c) $-(a+b)=(-a)+(-b)$

(d) $-(a \cdot b)=(-a) \cdot b=a \cdot(-b)$.

Proof. The proof of item (a) is really an exercise in understanding the definition of the additive inverse. The expression $-a$ means "the additive inverse of a." So the expression "$-(-a)$ means the additive inverse of $-a$. What is the additive inverse of $-a$? It's a of course. That's because

$$a+(-a)=(-a)+a=0.$$

To prove (b), we want to so that $(-1)a$ is the additive inverse of a. How would we do that? Well, we must show that

$$a+(-1)a=0.$$

Here it goes:

$$\begin{aligned} a+(-1)a &= 1a+(-1)a \\ &= (1+(-1))a \\ &= 0 \cdot a \\ &= 0 \end{aligned}$$

Thus $a+(-1)a=0$. Since addition is commutative we know that $a+(-1)a=(-1)a+a=0$. So $(-1)a$ fits the definition of an additive inverse of a. Since additive inverses are unique we conclude that $(-1)a=-a$. The proofs of parts (c) and (d) are left as exercises. □

It's a good idea to translate the statements in Proposition 1.4 into statements in ordinary language:

Mathematical Statement	**English Statement**
$-(-a)=a$	The additive inverse of the additive inverse of a is a itself.
$-a=(-1)a$	The additive inverse of a is -1 times a.
$-(a+b)=(-a)+(-b)$	The additive inverse of a sum is the sum of the additive inverses.
$-(a \cdot b)=(-a) \cdot b$	The additive inverse of a times b is b times the additive inverse of a.

Notice that in the proof above we had a nice, if slightly tricky, application of the distributive property. That trick is really helpful. Here's another application of it.

Proposition 1.5. *If $a \in \mathbb{F}$, then $a \cdot 0 = 0 \cdot a = 0$.*

Proof. We have

$$\begin{aligned} a &= a \cdot 1 \\ &= a \cdot (1 + 0) \\ &= a \cdot 1 + a \cdot 0 \\ &= a + a \cdot 0. \end{aligned}$$

Thus, $a = a + a \cdot 0$. Now add $-a$ to both sides:

$$\begin{aligned} (-a) + a &= (-a) + (a + a \cdot 0) \\ 0 &= (-a + a) + a \cdot 0 \\ 0 &= 0 + a \cdot 0 \\ 0 &= a \cdot 0 \end{aligned}$$

□

> In the proof above we only used field axioms, but did not identify which ones we used as we went along. For each step in the proof above, identify the field axioms that justify the step.

Now let's discuss multiplicative inverses. The fundamental facts about multiplicative inverses largely parallel the fundamental facts about multiplicative inverses. In a field, there is a unique multiplicative identity. We'll always call it 1 unless we have a good reason not to. And, in a field every element except the additive identity has a multiplicative inverse which is unique to it. We will denote the multiplicative inverse of a as a^{-1}. Because our experience with fields is mostly limited to $\mathbb{Q}$ and $\mathbb{R}$, it is common to reflexively think that

$$a^{-1} = \frac{1}{a}.$$

For example, $2^{-1} = 1/2$. And it is true that in $\mathbb{Q}$ and $\mathbb{R}$ (and even in $\mathbb{C}$), $a^{-1} = 1/a$ for nonzero a. However, as we saw in the opening inquiry, it is not true in every field that $a = 1/a$.

Example 1.6. In the opening inquiry, we defined a field called $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$. Recall that the operations of addition $(+)$ and multiplication $(\cdot)$ are defined "modulo 5." That means we find the regular sum or product and then divide by 5 and take the remainder as our answer. For example, in $\mathbb{Z}_5$:

$$\begin{aligned} 3 + 4 &= 7 \\ 7 \div 5 &= 1 \text{ R } 2. \end{aligned}$$

Since the sum of 3 and 4 is 7 and the remainder when we divide 7 by 5 is 2, we conclude that in $\mathbb{Z}_5$, $3 + 4 = 2$. Using that procedure, we developed the following addition table:

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

Thus, in $\mathbb{Z}_5$ we have that

- $-1 = 4$

- $-2 = 3$
- $-3 = 2$
- $-4 = 1$

We followed the analogous process for multiplication. For example, $4 \cdot 4 = 16$. When we divide 16 by 5 we get 3 with a remainder of 1. So we conclude that in $\mathbb{Z}_5$, $4 \cdot 4 = 1$. Using that procedure, we developed the following multiplication table:

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Since $4 \cdot 4 = 1$ in $\mathbb{Z}_5$, we conclude that 4 is its own multiplicative inverse in $\mathbb{Z}_5$. That is, $4^{-1} = 4$ in $\mathbb{Z}_5$! Similarly, in $\mathbb{Z}_5$, $1^{-1} = 1$, $2^{-1} = 3$ and $3^{-1} = 2$.

The main point of the previous example is that $-a$ does not always mean "the negative of a," and a^{-1} does not always mean $1/a$. In every context it's safe to read $-a$ as "the additive inverse of a" and to read a^{-1} as "the multiplicative inverse of a."

We have the following facts about multiplicative inverses. The proof of this proposition is left as an exercise.

Proposition 1.7. *Suppose that $\mathbb{F}$ is a field and that $a, b \in \mathbb{F}$ and $a, b \neq 0$.*

(a) $(a^{-1})^{-1} = a$

(b) $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$

(c) $(-a)^{-1} = -a^{-1}$

> Again, it's helpful here to state the parts of this proposition in ordinary language. For example, part (a) asserts that "the multiplicative inverse of the multiplicative inverse of a is a itself." Do the same for parts (b) and (c).

Field Extensions

Sometimes we will be working with more than one field at a time. If we have two fields, say $\mathbb{F}$ and $\mathbb{G}$ and

- $\mathbb{F} \subseteq \mathbb{G}$, and
- The operations on $\mathbb{F}$ are the operations on $\mathbb{G}$ restricted to $\mathbb{F}$

then we say that $\mathbb{G}$ is an **extension** of $\mathbb{F}$. If $\mathbb{G}$ is an **extension** of $\mathbb{F}$, then we say that $\mathbb{F}$ is a **subfield** of $\mathbb{G}$.

As an example, $\mathbb{Q}$ is a subfield of $\mathbb{R}$. Similarly, $\mathbb{C}$ is an extension of $\mathbb{R}$.

INQUIRY: $\mathbb{Q}(\sqrt{n})$

In this section we develop the idea of a **quadratic extension** of the field of rational numbers. This construction will be very important to us later in this module. We begin with a task in which students convince themselves that $\mathbb{Q}(\sqrt{2})$ is a field.

Consider the set $\mathbb{Q}(\sqrt{2})$ defined as
$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}.$$
Given $a + b\sqrt{2}$ and $c + d\sqrt{2}$ in $\mathbb{Q}(\sqrt{2})$ we define

- $(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}$, and
- $(a + b\sqrt{2}) \cdot (c + d\sqrt{2}) = (ac + 2bd) + (ad + bd)\sqrt{2}$.

The definition of addition here is quite natural, but the definition of multiplication might seem confusing until you realize that it is just the result of the distributive law:

$$\begin{aligned}(a + b\sqrt{2}) \cdot (c + d\sqrt{2}) &= a(c + d\sqrt{2}) + b\sqrt{2}(c + d\sqrt{2}) \\ &= ac + ad\sqrt{2} + bc\sqrt{2} + bd\sqrt{2}\sqrt{2} \\ &= ac + ad\sqrt{2} + bc\sqrt{2} + 2bd \\ &= (ac + 2bd) + (ad + bd)\sqrt{2}\end{aligned}$$

With these definitions, explore the following questions in small groups.

1. Does $\mathbb{Q}(\sqrt{2})$ satisfy the commutative laws? Convince yourself that the commutative law for multiplication is true in $\mathbb{Q}(\sqrt{2})$ by computing $(a + b\sqrt{2}) \cdot (c + d\sqrt{2})$ and $(c + d\sqrt{2}) \cdot (a + b\sqrt{2})$ and showing that they have the same value.
2. Does $\mathbb{Q}(\sqrt{2})$ satisfy the associative laws? Convince yourself that the associative law for multiplication is true in $\mathbb{Q}(\sqrt{2})$ by computing $((a + b\sqrt{2}) \cdot (c + d\sqrt{2})) \cdot (e + f\sqrt{2})$ and $(a + b\sqrt{2}) \cdot ((c + d\sqrt{2}) \cdot (e + f\sqrt{2}))$ and showing that they have the same value.
3. What is the additive identity in $\mathbb{Q}(\sqrt{2})$? What is the mutiplicative identity in $\mathbb{Q}(\sqrt{2})$?
4. If $a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ what is its additive inverse is clearly $-a - b\sqrt{2}$ and this is in $\mathbb{Q}(\sqrt{2})$. Suppose we know that $a + b\sqrt{2} \neq 0$. What is the multiplicative inverse of $a + b\sqrt{2}$? That is, what is the value of $(a + b\sqrt{2})^{-1}$? Is the multiplicative inverse of $a + b\sqrt{2}$ an element of $\mathbb{Q}(\sqrt{2})$?

In the inquiry above, we showed that $\mathbb{Q}(\sqrt{2})$ is a field. The field $\mathbb{Q}(\sqrt{2})$ is called a **quadratic extension** of $\mathbb{Q}$. First, $\mathbb{Q}(\sqrt{2})$ is an extention of $\mathbb{Q}$ because $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2})$ and the operations of addition and multiplication on $\mathbb{Q}$ are the same as those on $\mathbb{Q}(\sqrt{2})$, just restricted to $\mathbb{Q}$. We call $\mathbb{Q}(\sqrt{2})$ a quadratic extension because we extened $\mathbb{Q}$ by **adjoining** the root of a quadratic polynomial over $\mathbb{Q}$. In this case, we adjoined a root of

$$x^2 - 2$$

to $\mathbb{Q}$.

Let n be a positive integer and define $\mathbb{Q}(\sqrt{n})$. Show that $\mathbb{Q}(\sqrt{n})$ is a field. When does $\mathbb{Q}(\sqrt{n}) = \mathbb{Q}$?

Ordered Fields

Above we gave the axioms that define a field. We add to those the following **order axioms** to create an **ordered field**.

Definition 1.8. Suppose that $\mathbb{F}$ is a field. Then $\mathbb{F}$ is an **ordered field** if there is an order relation $<$ on $\mathbb{F}$ that satisfies the following properties for any $a, b, c \in \mathbb{F}$:

- (Trichotomy) Exactly one of the following holds: $a < b$, $a = b$ or $b < a$.
- (Transitivity) $a < b$ and $b < c$ implies that $a < c$.
- (Addition) If $a < b$, then $a + b < a + c$.
- (Multiplication) If $a < b$, $0 < c$, then $ac < bc$.

Definition 1.9. Suppose that $\mathbb{F}$ is an ordered field. An element $a \in \mathbb{F}$ is positive if $0 < a$ and a is negative if $a < 0$.

Proposition 1.10. *Suppose that $\mathbb{F}$ is an ordered field. Then*

(a) *$x \in \mathbb{F}$ is positive if and only if $-x$ is negative.*

(b) *If $x, y \in \mathbb{F}$, then $x + y$ and xy are positive.*

(c) *If $x \neq 0$, then x^2 is positive.*

(d) *1 is positive.*

Proof. For part (a) we proceed as follows. Suppose that $x \in \mathbb{F}$ is positive. By definition that means $0 < x$. Since $x \in \mathbb{F}$ it's additive inverse $-x$ is also in $\mathbb{F}$. By the addition axiom for ordering, since $0 < x$ we have

$$-x + 0 < -x + x.$$

We know that $-x + 0 = -x$ and that $-x + x = 0$. So we have $-x < 0$. Thus, if x is positive, then $-x$ is negative. The proofs of the remaining items are left as exercises. $\square$

Use the field axioms and the ordering axioms to prove parts (b)-(d) of the previous proposition.

Proposition 1.11. *Suppose that $\mathbb{F}$ is an ordered field. Then for all $a, b, c \in \mathbb{F}$ if $a < b$ and $c < 0$, then $ac > bc$.*

Proof. To see this, suppose that $a < b$ and that $c < 0$. Then $0 < -c$. So multiplication by $-c$ preserves order:

$$(-c)a < (-c)b.$$

But $(-c)a = -ca$ and $(-c)b = -cb$. Thus, we have $-ca < -cb$. But then

$$(ca + cb) - ca < (ca + cb) - cb.$$

Then by commutativity and associativity of addition

$$(ca - ca) + cb < ca + (cb - cb).$$

But $ca - ca = 0$ and $cb - cb = 0$ so we have $cb < ca$. By commutativity of multiplication we can restate this as

$$ac > bc.$$

$\square$

Corollary 1.12. $\mathbb{C}$ *is not an ordered field.*

Proof. This statement means that it is impossible to define an ordering on $\mathbb{C}$ which satisfies the four ordering axioms above. To see this, suppose by way of contradiction that we have found an ordering $<$ on $\mathbb{C}$ that satisfies all four of the ordering axioms. Then either $0 < i$ or $i < 0$. If $0 < i$, then $i \cdot 0 < i \cdot i$. So $0 < -1$. But since 1 is positive, we know that -1 must be negative, and so we have a contradiction. Now suppose that $i < 0$. Then since we are multiplying by a negative element, the inequality is reversed. So $i \cdot i > i \cdot 0$. That is, $-1 > 0$. Again, this is a contradiction. $\square$

Proposition 1.13. *In an ordered field*

(a) *the product of a positive element and a negative element is negative.*

(b) *the product of a negative element and a negative element is positive.*

Proof. To prove (a), let us suppose that $x > 0$ and $y < 0$. We claim that $xy < 0$. Suppose not. Then $xy \geq 0$. Since $x > 0$ we know that $x^{-1} > 0$. Thus $x^{-1}(xy) \geq x^{-1}(0) = 0$. Then by associativity, $(x^{-1}x)y \geq 0$. That is $y \geq 0$. This contradicts our assumption that $y < 0$.

We leave part (b) as an exercise. $\square$

Homework

1. Prove that if x is positive, then so is x^{-1}, and that if x is negative then x^{-1} is negative. Then use this fact to prove the next corollary.
2. If x and y are positive, then xy^{-1} and yx^{-1} are positive.
3. Prove that the product of two negative elements of a field is a positive element.

In-Class Resources

Opening Inquiry: Finite Fields

We are interested in mathematical structures with operations of addition ($+$) and multiplication ($\cdot$) that satisify the following properties:

(A) $a + b = b + a$ and $a \cdot b = b \cdot a$ (commutative laws)

(B) $(a + b) + c = a + (b + c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (associative laws)

(C) $a \cdot (b + c) = a \cdot b + a \cdot c$ (distributive law)

(D) There are distinct elements, called 0 and 1, such that $a + 0 = a$ and $a \cdot 1 = a$ for all a.

(E) For each a there is an element b such that $a + b = 0$ and if $a \neq 0$, there is an element c such that $a \cdot c = 1$.

Structures which satisfy these properties are called **fields**. We now define a field called $\mathbb{Z}_5$. We will avoid the complexities of a formal definition of $\mathbb{Z}_5$ and simply assert that $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$. The operations of addition ($+$) and multiplication ($\cdot$) are defined "modulo 5." That means we find the regular sum or product and then divide by 5 and take the remainder as our answer. For example, in $\mathbb{Z}_5$:

$$3 + 4 = 7$$
$$7 \div 5 = 1 \text{ R } 2.$$

Since the sum of 3 and 4 is 7 and the remainder when we divide 7 by 5 is 2, we conclude that in $\mathbb{Z}_5$, $3 + 4 = 2$. Using that procedure, fill in the following addition table:

$+$	0	1	2	3	4
0					
1					
2					
3					2
4					

1. Find a number $a \in \mathbb{Z}_5$ with the property that $a + 2 = 2 + a = 0$. The number a that you find is called the **additive inverse** of 2 in $\mathbb{Z}_5$. That is, it is the value of -2 in $\mathbb{Z}_5$.
   (a) $-2 =$ ______ in $\mathbb{Z}_5$ because $2 +$ ______ $=$ ______ $+ 2 =$ ______.
2. Find the value of -1, -3, and -4 in $\mathbb{Z}_5$.
   (a) $-1 =$ ______ in $\mathbb{Z}_5$ because $1 +$ ______ $=$ ______ $+ 1 =$ ______.
   (b) $-3 =$ ______ in $\mathbb{Z}_5$ because $3 +$ ______ $=$ ______ $+ 3 =$ ______.
   (c) $-4 =$ ______ in $\mathbb{Z}_5$ because $4 +$ ______ $=$ ______ $+ 4 =$ ______.

We follow the analogous process for multiplication. For example, $4 \cdot 4 = 16$. When we divide 16 by 5 we get 3 with a remainder of 1. So we conclude that in $\mathbb{Z}_5$, $4 \cdot 4 = 1$. Using that procedure, fill in the following multiplication table:

$\cdot$	0	1	2	3	4
0					
1					
2					
3					
4					1

Since $4 \cdot 4 = 1$ in $\mathbb{Z}_5$, we conclude that 4 is its own **multiplicative inverse** in $\mathbb{Z}_5$. That is, $4^{-1} = 4$ in $\mathbb{Z}_5$!

3. Find the value of 2^{-1} in $\mathbb{Z}_5$ and the value of 3^{-1} in $\mathbb{Z}_5$.

   (a) $2^{-1} =$ ______ in $\mathbb{Z}_5$ because $2 \cdot$ ______ $=$ ______ $\cdot 2 =$ ______

   (b) $3^{-1} =$ ______ in $\mathbb{Z}_5$ because $3 \cdot$ ______ $=$ ______ $\cdot 3 =$ ______

One can see by inspection of the addition and multiplication tables for $\mathbb{Z}_7$ that every element of $\mathbb{Z}_7$ has an additive inverse, and every nonzero element of $\mathbb{Z}_7$ has a multiplicative inverse. That is, $\mathbb{Z}_7$ satisifies (E) above. It is also easy to see by inspecting those tables that for all $a \in \mathbb{Z}_7$ $0 + a = a + 0 = a$ and $a \cdot 1 = 1 \cdot a = a$. That is, property (D) above is true in $\mathbb{Z}_7$.

4. Verify that in $\mathbb{Z}_7$ the following number sentences are true. Be prepared to share you work with the class.

   - $3 \cdot (1 + 4) = 3 \cdot 1 + 3 \cdot 4$
   - $5 \cdot (2 + 6) = 5 \cdot 2 + 5 \cdot 7$

5. Make up two number sentences with the same form using numbers $0 - 6$ and verify that they are true in $\mathbb{Z}_7$.

6. Do you believe that the distributive law (C) holds in $\mathbb{Z}_7$? Have you proved it? Exactly how many different instances of the distributive law are there in $\mathbb{Z}_7$?

7. Verify that in $\mathbb{Z}_7$ the following number sentences are true. Be prepared to share you work with the class.

   - $3 + (1 + 4) = (3 + 1) + 4$ and $5 + (2 + 6) = (5 + 2) + 6$.
   - $3 \cdot (1 \cdot 4) = (3 \cdot 1) \cdot 4$ and $5 \cdot (2 \cdot 6) = (5 \cdot 2) \cdot 6$.

8. Do you believe that the associative laws (B) hold in $\mathbb{Z}_7$? Have you proved it? Exactly how many different instances of each of the two distributive laws are there in $\mathbb{Z}_7$?

9. Verify that in $\mathbb{Z}_7$ the following number sentences are true. Be prepared to share you work with the class.
   - $3 + 4 = 4 + 3$ and $5 + 6 = 6 + 5$.
   - $3 \cdot 4 = 4 \cdot 5$ and $5 \cdot 6 = 6 \cdot 5$.

10. Do you believe that the commutative laws (A) hold in $\mathbb{Z}_7$? Have you proved it? Exactly how many different instances of each of the two commutative laws are there in $\mathbb{Z}_7$?

11. Based on your work above, does $\mathbb{Z}_7$ satisfy properties (A)-(E)? Is it a field?

12. Create addition and multiplication tables for $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ and answer appropriate versions of questions 1-11 for this structure.

13. Create addition and multiplication tables for $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ and answer appropriate versions of questions 1-11 for this structure.

14. Once you have determined whether or not $\mathbb{Z}_4$ and $\mathbb{Z}_5$ are fields, form a conjecture about which values of n make $\mathbb{Z}_n$ a field. Test your conjecture with another value of n and be prepared to discuss your conjecture with the class.

Inquiry: $\mathbb{Q}(\sqrt{2})$

Consider the set $\mathbb{Q}(\sqrt{2})$ defined as
$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}.$$
Given $a + b\sqrt{2}$ and $c + d\sqrt{2}$ in $\mathbb{Q}(\sqrt{2})$ we define

- $(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}$, and
- $(a + b\sqrt{2}) \cdot (c + d\sqrt{2}) = (ac + 2bd) + (ad + bd)\sqrt{2}$.

The definition of addition here is quite natural, but the definition of multiplication might seem confusing until you realize that it is just the result of the distributive law:
$$\begin{aligned}(a + b\sqrt{2}) \cdot (c + d\sqrt{2}) &= a(c + d\sqrt{2}) + b\sqrt{2}(c + d\sqrt{2})\\ &= ac + ad\sqrt{2} + bc\sqrt{2} + bd\sqrt{2}\sqrt{2}\\ &= ac + ad\sqrt{2} + bc\sqrt{2} + 2bd\\ &= (ac + 2bd) + (ad + bd)\sqrt{2}\end{aligned}$$

With these definitions, explore the following questions in small groups.

1. Does $\mathbb{Q}(\sqrt{2})$ satisfy the commutative laws? Convince yourself that the commutative law for multiplication is true in $\mathbb{Q}(\sqrt{2})$ by computing $(a + b\sqrt{2}) \cdot (c + d\sqrt{2})$ and $(c + d\sqrt{2}) \cdot (a + b\sqrt{2})$ and showing that they have the same value.
   - $(a + b\sqrt{2}) \cdot (c + d\sqrt{2}) =$
   - $(c + d\sqrt{2}) \cdot (a + b\sqrt{2}) =$

   Now start with the result of $(a + b\sqrt{2}) \cdot (c + d\sqrt{2})$ and manipulate it algebraically to derive the result of $(c + d\sqrt{2}) \cdot (a + b\sqrt{2})$.

2. Does $\mathbb{Q}(\sqrt{2})$ satisfy the associative laws? Convince yourself that the associative law for multiplication is true in $\mathbb{Q}(\sqrt{2})$ by computing $((a + b\sqrt{2}) \cdot (c + d\sqrt{2})) \cdot (e + f\sqrt{2})$ and $(a + b\sqrt{2}) \cdot ((c + d\sqrt{2}) \cdot (e + f\sqrt{2}))$ and showing that they have the same value.
   - $(a + b\sqrt{2}) \cdot (c + d\sqrt{2}) =$
   - $((a + b\sqrt{2}) \cdot (c + d\sqrt{2})) \cdot (e + f\sqrt{2}) =$
   - $(c + d\sqrt{2}) \cdot (e + f\sqrt{2}) =$
   - $(a + b\sqrt{2}) \cdot ((c + d\sqrt{2}) \cdot (e + f\sqrt{2})) =$

   Now start with the result of $((a + b\sqrt{2}) \cdot (c + d\sqrt{2})) \cdot (e + f\sqrt{2})$ and manipulate it algebraically to derive the result of $(a + b\sqrt{2}) \cdot ((c + d\sqrt{2}) \cdot (e + f\sqrt{2}))$.

3. What is the additive identity in $\mathbb{Q}(\sqrt{2})$? What is the mutiplicative identity in $\mathbb{Q}(\sqrt{2})$?

4. If $a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ what is its additive inverse is clearly $-a - b\sqrt{2}$ and this is in $\mathbb{Q}(\sqrt{2})$. Suppose we know that $a + b\sqrt{2} \neq 0$. What is the multiplicative inverse of $a + b\sqrt{2}$? That is, what is the value of $(a + b\sqrt{2})^{-1}$? Is the multiplicative inverse of $a + b\sqrt{2}$ an element of $\mathbb{Q}(\sqrt{2})$?
Hint: Start with the equation $(a + b\sqrt{2})x = 1$ and solve for x. Yes, of course, after one step you get $x = \frac{1}{a+b\sqrt{2}}$. The crucial part though is showing that you can put this into the form $c + d\sqrt{2}$ for some $c, d \in \mathbb{Q}$.

Inquiry: $\mathbb{Q}(x)$

In $\mathbb{Q}$ and $\mathbb{R}$ we know that we have the integers as a subset. And, each positive integer can be built up by repeatedly adding 1 to itself. That is,

$$1+1=2$$
$$1+1+1=3$$

and so on. Then, once we have the positive integers, we know we have the negative integers because in a field every element has an additive inverse.

We can do exactly the same construction in any field. That's because any arbitrary field $\mathbb{F}$ has a multplitive identity, which we call 1. (As we have seen, it may actually be quite different from the number 1.) Then we can form special elements by repeatedly adding the multiplicative identity to itself over and over. So for example, in any field we can define

$$2=1+1$$
$$3=1+1+1$$

and so on. We call these elements **integral elements**. In this inquiry we want to study the integral elements of a somewhat strange field - the field of rational functions with rational coefficients. We begin by defining this field, which we denote $\mathbb{Q}(x)$.

Defining $\mathbb{Q}(x)$

First we define the set of objects $\mathbb{Q}(x)$:

$$\mathbb{Q}(x) = \left\{ \frac{p(x)}{q(x)} \mid p(x) \text{ and } q(x) \text{ are polynomials with rational coefficients} \right\}.$$

1. Which of the following are elements of $\mathbb{Q}(x)$?
   - $r(x) = x^2 + 1$
   - $s(x) = \frac{3x^2+\pi x+1}{4x^3+2x^2+x+2}$
   - $t(x) = \frac{x^{1/2}+2x+1}{\frac{3}{2}x+5}$
   - $u(x) = \frac{\frac{2}{3}x^2+2x+1}{\frac{3}{2}x+5}$

We define the equality of two rational functions, as well as their sums and products, like we do with fractions. That is, given $r_1(x), r_2(x) \in \mathbb{Q}(x)$ with $r_1(x) = \frac{p_1(x)}{q_1(x)}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)}$ we define

- $r_1(x) = r_2(x)$ if and only if $p_1(x)q_2(x) - q_1(x)p_2(x) = 0$
- $(r_1 + r_2)(x) = \frac{p_1(x)q_2(x)+q_1(x)p_2(x)}{q_1(x)q_2(x)}$
- $(r_1 \cdot r_2)(x) = \frac{p_1(x)p_2(x)}{q_1(x)q_2(x)}$

2. Suppose that
$$r_1(x) = \frac{x^2+\frac{1}{2}x+2}{x+1}, \quad r_2(x) = \frac{x+2}{x^2+4}, \text{ and } r_3(x) = \frac{2x^2+x+4}{2x+2}.$$
Is $r_1(x) = r_2(x)$? Is $r_1(x) = r_3(x)$? Is $r_2(x) = r_3(x)$?

3. Suppose that
$$r_1(x) = \frac{3x^2+2x+1}{x+5} \text{ and } r_2(x) = \frac{8x+2}{4x^2+\frac{1}{2}}.$$
Compute $(r_1+r_2)(x)$ and $(r_1 \cdot r_2)(x)$.

Next we define an ordering on $\mathbb{Q}(x)$ as follows: Given $r_1(x), r_2(x) \in \mathbb{Q}(x)$ with $r_1(x) = \frac{p_1(x)}{q_1(x)}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)}$, suppose that the leading coefficients of $q_1(x)$ and $q_2(x)$ are both positive. Then
$$r_1(x) < r_2(x)$$
if and only if the leading coefficient of
$$p_2(x)q_1(x) - p_1(x)q_2(x)$$
is positive. This ordering can be difficult to get a handle on. In practice you want to (a) make sure to leading coefficients in the denominator are positive, and if not multiply by $-1/-1$ to make it so and then (b) cross multiply and subtract. Here are two examples, followed by some exercises for you to try.

Example 1.14. Let us compare $r_1(x) = \frac{2x^2+11x+1}{x+5}$ and $r_2(x) = 2x+1$. First, we observe that as a rational function $r_2(x) = \frac{2x+1}{1}$ and that both $r_1(x)$ and $r_2(x)$ have positive leading coefficients in the denominator. Now we cross multiply and obtain
$$(x+5)(2x+1) = 2x^2+11x+5$$
$$(1)(2x^2+11x+1) = 2x^2+11x+1$$
We see that $(2x^2+11x+5) - (2x^2+11x+1) = 4$ while subtracting in the other direction gives a negative. So we conclude that
$$\frac{2x^2+11x+1}{x+5} < 2x+1.$$

Example 1.15. Let us compare $r_1(x) = \frac{-x^2+3}{-2x+1}$ and $r_2(x) = \frac{2x+1}{2x-1}$. First, we multiply $r_1(x)$ by $-1/-1$ to rewrite it as $r_1(x) = \frac{x^2-3}{2x-1}$. Now we cross multiply and obtain
$$(2x-1)(x^2-3) = 2x^3 - x^2 - 6x + 4$$
$$(2x-1)(2x+1) = 4x^2 - 1$$
We see that $(2x^3 - x^2 - 6x + 4) - (4x^2 - 1) = 2x^3 - 5x^2 - 6x + 3$ has positive leading coefficient. So we conclude that
$$r_2(x) = \frac{2x+1}{2x-1} < r_1(x) = \frac{-x^2+3}{-2x+1}.$$

4. It is important to play with this ordering a bit to get used to it. Try to work through these examples in small groups. For each pair $r_1(x)$ and $r_2(x)$ determine whether $r_1(x) < r_2(x)$, $r_1(x) = r_2(x)$ or $r_2(x) < r_1(x)$.
   (a) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x+5}{3x^2+x-2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x+6}{3x^2+x-2}$.
   (b) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x^2+5}{3x^2+x-2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x+5}{3x^2+x-2}$.
   (c) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x^2+5}{3x-2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x^2+5}{3x^2-2}$.
   (d) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x^2+5}{-3x+2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x^2+5}{3x-2}$.

The purpose of this opening task is to give you a sense of how the field $\mathbb{Q}(x)$ works. The fact that these operations make it an ordered field is a theorem which is not difficult to prove, but is somewhat tedious. So here we simply state it without proof.

Theorem 1.16. *The set of rational functions over the rationals,*

$$\mathbb{Q}(x) = \left\{ \frac{p(x)}{q(x)} \mid p(x) \text{ and } q(x) \text{ are polynomials with rational coefficients} \right\}.$$

together with the operations of addition (+) and multiplication (·) defined above form a field. If we add in the ordering $<$ defined above, $\mathbb{Q}(x)$ is an ordered field.

We opened this inquiry by talking about integral elements. Here, in $\mathbb{Q}(x)$, the integral elements are of the form

$$q_n(x) = n$$

for $n \in \mathbb{N}$. (Notice that these are degree 0 polynomials, and are also rational functions since $q_n(x) = \frac{n}{1}$.) As your final task in this inquiry, prove the following facts:

1. Suppose that $r(x) = x^2 + 2x + 3$. Prove that $q_{500}(x) < r(x)$
2. Suppose that $r(x) = 2x + 3$. Prove that for any $n \in \mathbb{N}$, $q_n(x) < r(x)$.
3. Suppose that $r(x) = a_m x^m + a_{m-1} x_{m-1} + \cdots a_1 x + a_0$ is an element of $\mathbb{Q}(x)$ with $m \geq 1$. Prove that if $a_m > 0$, then $q_{500}(x) < r(x)$
4. Suppose that $r(x) = a_m x^m + a_{m-1} x_{m-1} + \cdots a_1 x + a_0$ is an element of $\mathbb{Q}(x)$ with $m \geq 1$. Prove that for any $n \in \mathbb{N}$, if $a_m > 0$, then $q_n(x) < r(x)$

You've just shown that in $\mathbb{Q}(x)$ every polynomial with degree at least 1 and with positive leading coefficient is greater than every integral element.

Part II

Semi-Advanced Topics in Ordered Fields

Opening Inquiry: $\mathbb{Q}(x)$

In $\mathbb{Q}$ and $\mathbb{R}$ we know that we have the integers as a subset. And, each positive integer can be built up by repeatedly adding 1 to itself. That is,

$$1 + 1 = 2$$
$$1 + 1 + 1 = 3$$

and so on. Then, once we have the positive integers, we know we have the negative integers because in a field every element has an additive inverse.

We can do exactly the same construction in any field. That's because any arbitrary field $\mathbb{F}$ has a multplitive identity, which we call 1. (As we have seen, it may actually be quite different from the number 1.) Then we can form special elements by repeatedly adding the multiplicative identity to itself over and over. So for example, in any field we can define

$$2 = 1 + 1$$
$$3 = 1 + 1 + 1$$

and so on. We call these elements **integral elements**. In this inquiry we want to study the integral elements of a somewhat strange field - the field of rational functions with rational coefficients. We begin by defining this field, which we denote $\mathbb{Q}(x)$.

DEFINING $\mathbb{Q}(x)$

First we define the set of objects $\mathbb{Q}(x)$:

$$\mathbb{Q}(x) = \left\{ \frac{p(x)}{q(x)} \mid p(x) \text{ and } q(x) \text{ are polynomials with rational coefficients} \right\}.$$

1. Which of the following are elements of $\mathbb{Q}(x)$?
   - $r(x) = x^2 + 1$
   - $s(x) = \frac{3x^2+\pi x+1}{4x^3+2x^2+x+2}$
   - $t(x) = \frac{x^{1/2}+2x+1}{\frac{3}{2}x+5}$
   - $u(x) = \frac{\frac{2}{3}x^2+2x+1}{\frac{3}{2}x+5}$

We define the equality of two rational functions, as well as their sums and products, like we do with fractions. That is, given $r_1(x), r_2(x) \in \mathbb{Q}(x)$ with $r_1(x) = \frac{p_1(x)}{q_1(x)}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)}$ we define

- $r_1(x) = r_2(x)$ if and only if $p_1(x)q_2(x) - q_1(x)p_2(x) = 0$
- $(r_1 + r_2)(x) = \frac{p_1(x)q_2(x)+q_1(x)p_2(x)}{q_1(x)q_2(x)}$
- $(r_1 \cdot r_2)(x) = \frac{p_1(x)p_2(x)}{q_1(x)q_2(x)}$

2. Suppose that
$$r_1(x) = \frac{x^2 + \frac{1}{2}x + 2}{x+1}, \quad r_2(x) = \frac{x+2}{x^2+4}, \text{ and } r_3(x) = \frac{2x^2+x+4}{2x+2}.$$
Is $r_1(x) = r_2(x)$? Is $r_1(x) = r_3(x)$? Is $r_2(x) = r_3(x)$?
3. Suppose that
$$r_1(x) = \frac{3x^2+2x+1}{x+5} \text{ and } r_2(x) = \frac{8x+2}{4x^2+\frac{1}{2}}.$$
Compute $(r_1 + r_2)(x)$ and $(r_1 \cdot r_2)(x)$.

Next we define an ordering on $\mathbb{Q}(x)$ as follows: Given $r_1(x), r_2(x) \in \mathbb{Q}(x)$ with $r_1(x) = \frac{p_1(x)}{q_1(x)}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)}$, suppose that the leading coefficients of $q_1(x)$ and $q_2(x)$ are both positive. Then

$$r_1(x) < r_2(x)$$

if and only if the leading coefficient of

$$p_2(x)q_1(x) - p_1(x)q_2(x)$$

is positive. This ordering can be difficult to get a handle on. In practice you want to (a) make sure to leading coefficients in the denominator are positive, and if not multiply by $-1/-1$ to make it so and then (b) cross multiply and subtract. Here are two examples, followed by some exercises for you to try.

Example 1.17. Let us compare $r_1(x) = \frac{2x^2+11x+1}{x+5}$ and $r_2(x) = 2x+1$. First, we observe that as a rational function $r_2(x) = \frac{2x+1}{1}$ and that both $r_1(x)$ and $r_2(x)$ have positive leading coefficients in the denominator. Now we cross multiply and obtain

$$(x+5)(2x+1) = 2x^2 + 11x + 5$$
$$(1)(2x^2 + 11x + 1) = 2x^2 + 11x + 1$$

We see that $(2x^2 + 11x + 5) - (2x^2 + 11x + 1) = 4$ while subtracting in the other direction gives a negative. So we conclude that

$$\frac{2x^2+11x+1}{x+5} < 2x+1.$$

Example 1.18. Let us compare $r_1(x) = \frac{-x^2+3}{-2x+1}$ and $r_2(x) = \frac{2x+1}{2x-1}$. First, we multiply $r_1(x)$ by $-1/-1$ to rewrite it as $r_1(x) = \frac{x^2-3}{2x-1}$. Now we cross multiply and obtain

$$(2x-1)(x^2-3) = 2x^3 - x^2 - 6x + 4$$
$$(2x-1)(2x+1) = 4x^2 - 1$$

We see that $(2x^3 - x^2 - 6x + 4) - (4x^2 - 1) = 2x^3 - 5x^2 - 6x + 3$ has positive leading coefficient. So we conclude that

$$r_2(x) = \frac{2x+1}{2x-1} < r_1(x) = \frac{-x^2+3}{-2x+1}.$$

4. It is important to play with this ordering a bit to get used to it. Try to work through these examples in small groups. For each pair $r_1(x)$ and $r_2(x)$ determine whether $r_1(x) < r_2(x)$, $r_1(x) = r_2(x)$ or $r_2(x) < r_1(x)$.
   (a) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x+5}{3x^2+x-2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x+6}{3x^2+x-2}$.
   (b) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x^2+5}{3x^2+x-2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x+5}{3x^2+x-2}$.
   (c) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x^2+5}{3x-2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x^2+5}{3x^2-2}$.
   (d) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x^2+5}{-3x+2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x^2+5}{3x-2}$.

The purpose of this opening task is to give you a sense of how the field $\mathbb{Q}(x)$ works. The fact that these operations make it an ordered field is a theorem which is not difficult to prove, but is somewhat tedious. So here we simply state it without proof.

Theorem. The set of rational functions over the rationals,

$$\mathbb{Q}(x) = \left\{ \frac{p(x)}{q(x)} \mid p(x) \text{ and } q(x) \text{ are polynomials with rational coefficients} \right\}.$$

together with the operations of addition $(+)$ and multiplication $(\cdot)$ defined above form a field. If we add in the ordering $<$ defined above, $\mathbb{Q}(x)$ is an ordered field.

We opened this inquiry by talking about integral elements. Here, in $\mathbb{Q}(x)$, the integral elements are of the form

$$q_n(x) = n$$

for $n \in \mathbb{N}$. (Notice that these are degree 0 polynomials, and are also rational functions since $q_n(x) = \frac{n}{1}$.) As your final task in this inquiry, prove the following facts:

5. Suppose that $r(x) = x^2 + 2x + 3$. Prove that $q_{500}(x) < r(x)$
6. Suppose that $r(x) = 2x + 3$. Prove that for any $n \in \mathbb{N}$, $q_n(x) < r(x)$.
7. Suppose that $r(x) = a_m x^m + a_{m-1} x_{m-1} + \cdots a_1 x + a_0$ is an element of $\mathbb{Q}(x)$ with $m \geq 1$. Prove that if $a_m > 0$, then $q_{500}(x) < r(x)$
8. Suppose that $r(x) = a_m x^m + a_{m-1} x_{m-1} + \cdots a_1 x + a_0$ is an element of $\mathbb{Q}(x)$ with $m \geq 1$. Prove that for any $n \in \mathbb{N}$, if $a_m > 0$, then $q_n(x) < r(x)$

You've just shown that in $\mathbb{Q}(x)$ every polynomial with degree at least 1 and with positive leading coefficient is greater than every integral element.

2 Upper and Lower Bounds

Once we have a notion of order in a field, we can talk about upper and lower bounds.

Definition 2.1. Let $\mathbb{F}$ be an ordered field and suppose that $S \subseteq \mathbb{F}$. Then

- u is an upper bound of S if $x \leq u$ for all $x \in S$.
- l is a lower bound of S if $l \leq x$ for all $x \in S$.

Example 2.2. Consider the set $S = \{x \in \mathbb{Q} \mid 2 \leq x < 3\}$. Notice that 3 is an upper bound of S, but so is 3.5, 4, and indeed any number larger than 3. Similarly, 2 is a lower bound of $\mathbb{Q}$ as is $1\frac{3}{4}$ and any number less than 2. Upper bounds and lower bounds may or not be members of the sets they bound.

Definition 2.3. Let $\mathbb{F}$ be an ordered field and suppose that $S \subseteq \mathbb{F}$. Then

u is a least upper bound of S if

1. $x \leq u$ for all $x \in S$, and
2. If b is an upper bound of S, then $u \leq b$.

l is a greatest lower bound of S

1. if $l \leq x$ for all $x \in S$, and
2. if b is a lower bound of S, then $b \leq l$.

Example 2.4. Consider the set $S = \{x \in \mathbb{Q} \mid 2 \leq x < 3\}$. Notice that 3 is an upper bound of S, and any other upper bound of S is greater than or equal to 3. Thus, 3 is the least upper bound of S. Similarly, 2 is a lower bound of $\mathbb{Q}$ and any other lower bound of S must be less than or equal to 2. Thus, 2 is the greatest lower bound of S. Note that a greatest lower bound or least upper bound of a set may be in the set or not in the set.

The Archimedean Property and Completeness

Definition 2.5. An ordered field $\mathbb{F}$ is Archimedean if and only if for each positive $x \in \mathbb{F}$ there is an integral element $k \in \mathbb{F}$ such that $x < k$.

Theorem 2.6. *For each positive element x in an Archimedean field $\mathbb{F}$ there is a unique integral element n such that*

$$n \leq x < n + 1.$$

Definition 2.7. An ordered field $\mathbb{F}$ is complete if and only if every subset of $\mathbb{F}$ that is bounded above has a least upper bound.

TODO: Alternate completeness axiom (with greatest lower bound)

TODO: Notice that $\mathbb{Q}$ shows that Archimedean does not imply complete.

Theorem 2.8. *Any complete ordered field is Archimedean.*

WHAT ARE THE REAL NUMBERS, REALLY?

1. There are fields that are not ordered fields.
2. There are ordered fields that are not Archimedean.
3. There are Archimedean fields that are not complete.
4. Every complete ordered field is Archimedean.

The real numbers $\mathbb{R}$ is a complete ordered field. In fact, it is the unique complete ordered field.

Theorem 2.9. *Any complete ordered field is isomorphic to the ordered field of real numbers.*

Homework

1. Homework problems.

In-Class Resources

OPENING INQUIRY: $\mathbb{Q}(x)$

In $\mathbb{Q}$ and $\mathbb{R}$ we know that we have the integers as a subset. And, each positive integer can be built up by repeatedly adding 1 to itself. That is,

$$1+1=2$$
$$1+1+1=3$$

and so on. Then, once we have the positive integers, we know we have the negative integers because in a field every element has an additive inverse.

We can do exactly the same construction in any field. That's because any arbitrary field $\mathbb{F}$ has a multplitive identity, which we call 1. (As we have seen, it may actually be quite different from the number 1.) Then we can form special elements by repeatedly adding the multiplicative identity to itself over and over. So for example, in any field we can define

$$2=1+1$$
$$3=1+1+1$$

and so on. We call these elements **integral elements**. In this inquiry we want to study the integral elements of a somewhat strange field - the field of rational functions with rational coefficients. We begin by defining this field, which we denote $\mathbb{Q}(x)$.

DEFINING $\mathbb{Q}(x)$

First we define the set of objects $\mathbb{Q}(x)$:

$$\mathbb{Q}(x) = \left\{ \frac{p(x)}{q(x)} \mid p(x) \text{ and } q(x) \text{ are polynomials with rational coefficients} \right\}.$$

1. Which of the following are elements of $\mathbb{Q}(x)$?
   - $r(x) = x^2 + 1$
   - $s(x) = \frac{3x^2+\pi x+1}{4x^3+2x^2+x+2}$
   - $t(x) = \frac{x^{1/2}+2x+1}{\frac{3}{2}x+5}$
   - $u(x) = \frac{\frac{2}{3}x^2+2x+1}{\frac{3}{2}x+5}$

We define the equality of two rational functions, as well as their sums and products, like we do with fractions. That is, given $r_1(x), r_2(x) \in \mathbb{Q}(x)$ with $r_1(x) = \frac{p_1(x)}{q_1(x)}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)}$ we define

- $r_1(x) = r_2(x)$ if and only if $p_1(x)q_2(x) - q_1(x)p_2(x) = 0$

- $(r_1 + r_2)(x) = \frac{p_1(x)q_2(x)+q_1(x)p_2(x)}{q_1(x)q_2(x)}$
- $(r_1 \cdot r_2)(x) = \frac{p_1(x)p_2(x)}{q_1(x)q_2(x)}$

2. Suppose that
$$r_1(x) = \frac{x^2 + \frac{1}{2}x + 2}{x+1}, \quad r_2(x) = \frac{x+2}{x^2+4}, \text{ and } r_3(x) = \frac{2x^2+x+4}{2x+2}.$$
Is $r_1(x) = r_2(x)$? Is $r_1(x) = r_3(x)$? Is $r_2(x) = r_3(x)$?

3. Suppose that
$$r_1(x) = \frac{3x^2+2x+1}{x+5} \text{ and } r_2(x) = \frac{8x+2}{4x^2+\frac{1}{2}}.$$
Compute $(r_1 + r_2)(x)$ and $(r_1 \cdot r_2)(x)$.

Next we define an ordering on $\mathbb{Q}(x)$ as follows: Given $r_1(x), r_2(x) \in \mathbb{Q}(x)$ with $r_1(x) = \frac{p_1(x)}{q_1(x)}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)}$, suppose that the leading coefficients of $q_1(x)$ and $q_2(x)$ are both positive. Then
$$r_1(x) < r_2(x)$$
if and only if the leading coefficient of
$$p_2(x)q_1(x) - p_1(x)q_2(x)$$
is positive. This ordering can be difficult to get a handle on. In practice you want to (a) make sure to leading coefficients in the denominator are positive, and if not multiply by $-1/-1$ to make it so and then (b) cross multiply and subtract. Here are two examples, followed by some exercises for you to try.

Example 2.10. Let us compare $r_1(x) = \frac{2x^2+11x+1}{x+5}$ and $r_2(x) = 2x+1$. First, we observe that as a rational function $r_2(x) = \frac{2x+1}{1}$ and that both $r_1(x)$ and $r_2(x)$ have positive leading coefficients in the denominator. Now we cross multiply and obtain
$$(x+5)(2x+1) = 2x^2+11x+5$$
$$(1)(2x^2+11x+1) = 2x^2+11x+1$$
We see that $(2x^2+11x+5) - (2x^2+11x+1) = 4$ while subtracting in the other direction gives a negative. So we conclude that
$$\frac{2x^2+11x+1}{x+5} < 2x+1.$$

Example 2.11. Let us compare $r_1(x) = \frac{-x^2+3}{-2x+1}$ and $r_2(x) = \frac{2x+1}{2x-1}$. First, we multiply $r_1(x)$ by $-1/-1$ to rewrite it as $r_1(x) = \frac{x^2-3}{2x-1}$. Now we cross multiply and obtain

$$(2x-1)(x^2-3) = 2x^3 - x^2 - 6x + 4$$
$$(2x-1)(2x+1) = 4x^2 - 1$$

We see that $(2x^3 - x^2 - 6x + 4) - (4x^2 - 1) = 2x^3 - 5x^2 - 6x + 3$ has positive leading coefficient. So we conclude that

$$r_2(x) = \frac{2x+1}{2x-1} < r_1(x) = \frac{-x^2+3}{-2x+1}.$$

4. It is important to play with this ordering a bit to get used to it. Try to work through these examples in small groups. For each pair $r_1(x)$ and $r_2(x)$ determine whether $r_1(x) < r_2(x)$, $r_1(x) = r_2(x)$ or $r_2(x) < r_1(x)$.

   (a) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x+5}{3x^2+x-2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x+6}{3x^2+x-2}$.

   (b) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x^2+5}{3x^2+x-2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x+5}{3x^2+x-2}$.

   (c) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x^2+5}{3x-2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x^2+5}{3x^2-2}$.

   (d) $r_1(x) = \frac{p_1(x)}{q_1(x)} = \frac{x^2+5}{-3x+2}$ and $r_2(x) = \frac{p_2(x)}{q_2(x)} = \frac{x^2+5}{3x-2}$.

The purpose of this opening task is to give you a sense of how the field $\mathbb{Q}(x)$ works. The fact that these operations make it an ordered field is a theorem which is not difficult to prove, but is somewhat tedious. So here we simply state it without proof.

Theorem. The set of rational functions over the rationals,

$$\mathbb{Q}(x) = \left\{ \frac{p(x)}{q(x)} \mid p(x) \text{ and } q(x) \text{ are polynomials with rational coefficients} \right\}.$$

together with the operations of addition $(+)$ and multiplication $(\cdot)$ defined above form a field. If we add in the ordering $<$ defined above, $\mathbb{Q}(x)$ is an ordered field.

We opened this inquiry by talking about integral elements. Here, in $\mathbb{Q}(x)$, the integral elements are of the form

$$q_n(x) = n$$

for $n \in \mathbb{N}$. (Notice that these are degree 0 polynomials, and are also rational functions since $q_n(x) = \frac{n}{1}$.) As your final task in this inquiry, prove the following facts:

5. Suppose that $r(x) = x^2 + 2x + 3$. Prove that $q_{500}(x) < r(x)$

6. Suppose that $r(x) = 2x + 3$. Prove that for any $n \in \mathbb{N}$, $q_n(x) < r(x)$.

7. Suppose that $r(x) = a_m x^m + a_{m-1} x_{m-1} + \cdots a_1 x + a_0$ is an element of $\mathbb{Q}(x)$ with $m \geq 1$. Prove that if $a_m > 0$, then $q_{500}(x) < r(x)$

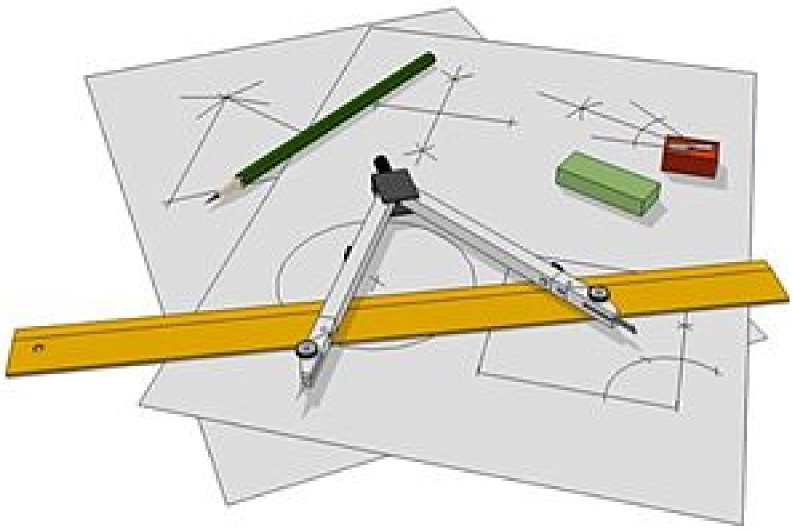
8. Suppose that $r(x) = a_m x^m + a_{m-1} x_{m-1} + \cdots a_1 x + a_0$ is an element of $\mathbb{Q}(x)$ with $m \geq 1$. Prove that for any $n \in \mathbb{N}$, if $a_m > 0$, then $q_n(x) < r(x)$

You've just shown that in $\mathbb{Q}(x)$ every polynomial with degree at least 1 and with positive leading coefficient is greater than every integral element.

Part III

Three Famous Problems about Constructible Numbers

Now we turn to a collection of problems considered by ancient Greek mathematicians, namely compass-and-straightedge constructions. Here, by the word “compass” we do not mean the navigation tool! The way we’re using the word, a compass is a drawing instrument that can be used for inscribing circles or arcs. By a “straightedge” we basically mean a ruler with no markings on it.



The ancient Greeks were curious about the following question which gave rise to these problems: Given a line segment of length 1 in the plane, for what values of a can we construct a line segment of length a using compass and straightedge constructions?

The Greeks discovered how to construct sums, differences, products, ratios, and square roots of given lengths. They could also construct half of a given angle; that is, they could “bisect” any angle. We will see these constructions below, and there are many other interesting constructions they figured out.

However, there were some constructions they could not figure out. For example, they could not figure out how to trisect an arbitrary angle. That is, they could not construct one third of a given angle except in some particular cases. They could not “square a circle.” That is, they could not figure out how to construct a square with the same area as a given circle. Nor could they “double a cube.” That is, they could not figure out how to construct the side of a cube whose volume would be twice the volume of a cube with a given side.

The interest in these problems from our point of view is that it took the development of the theory of fields to solve these problems. It was not until the early 19th century that it was proved to be impossible to trisect an arbitrary angle or of double the volume of a cube using straightedge and compass constructions. Then in the late 19th century it was shown that π is a transcendental number, and so it is impossible to use straightedge and compass to construct a square with the same area as a given circle. These proofs are beautiful uses of the theory of fields, and we will explore them below.

The Rules of the Game

In the past, compasses were used in mathematics, drafting, navigation and other purposes. Physical compasses are usually made of metal or plastic, and consist of two parts connected by a hinge which can be adjusted to allow the changing of the radius of the circle drawn. Typically one part has a spike at its end, and the other part a pencil or pen. Unlike physical compasses, we will assume our compasses can be opened arbitrarily wide. We will also assume that our compasses have no markings on them to measure angles or anything else.

We will assume our straightedges are infinitely long, and has no markings on them. Our straightedges can only be used to draw a line segment between two points or to extend an existing segment.

Our constructions must be exact. Sometimes you'll be tempted to "eyeball" it by looking at the construction and guessing at its accuracy, or using some form of measurement, such as the units of measure on a ruler. This is forbidden, and getting close does not count as a solution.

Finally, compass and striaightedge construction must have a finite number of steps.

The Fundamental Constructions

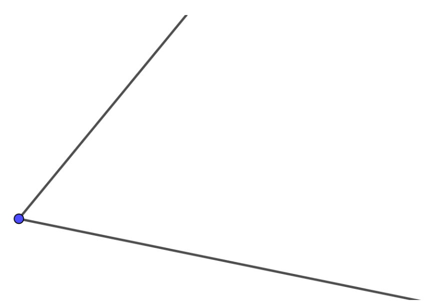
We begin with the three fundamental constructions with which we will build all of our more sophisticated constructions. We will assume that these are the only operations we can perform with our straightedge and compass.

Definition 2.12 (Fundamental Constructions)**.** The following compass and straightedge constructions are known as our three fundamental constructions.

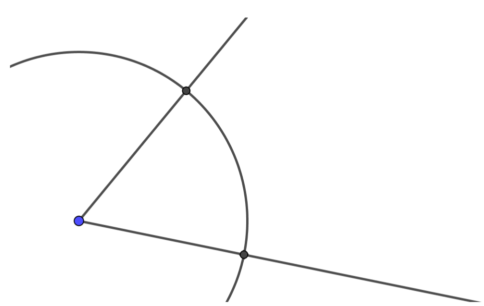
1. Given two points, we may draw a line through them, extending it indefinitely in each direction.
2. Given two points, we may draw the line segment connecting them.
3. Given a point and line segment, we may draw a circle with center at the point and radius equal to the length of the line segment.

We now build up some important basic constructions using the fundamental constructions.

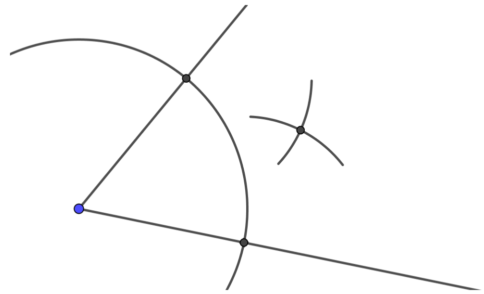
Example 2.13. Using the fundamental constructions, we can bisect any angle.



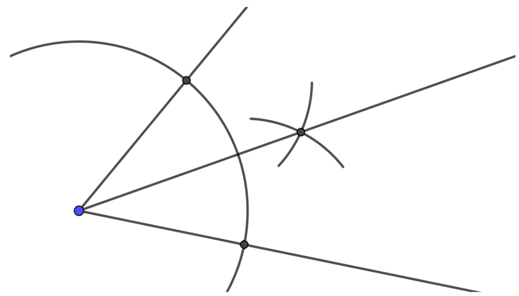
1. We begin by opening the compass to some fixed length, putting the point of the compass at the vertex of the angle, and drawing an arc through the two rays formed by the angle. This determines a point on each ray, equidistant from the vertex. (This uses the third fundamental construction.)



2. Keeping the compass at a fixed length (but perhaps different from the length in step 1), we put the compass point at each point formed in step 1 and create two intersecting arcs between the two rays of our angle. (This uses the third fundamental construction.)



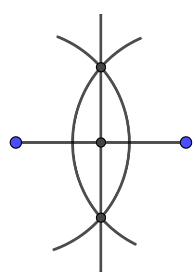
3. Using the straightedge, we draw a ray from the vertex of the angle through the point of intersection of the two arcs. This ray bisects the given angle. (This uses the first fundamental construction.)



Inquiry: Bisect a line segement

To bisect a line segment, follow these steps:

1. Open the compass to a width greater than half the length of the segment.
2. Place the point of the compass at one endpoint of the segment and draw an arc which passes over the midpoint of the segment.
3. Do the same with the point at the other endpoint of the segment without changing the width of the compass.
4. The arcs should intersect at two points. The line between those points bisects the segment.



Inquiry: Construct Angles of 30° and 60°

Follow these steps to construct angles of 30° and 60° using only the fundamental constructions.

1. Using the straightedge, draw a line segment of some length.
2. Open the compass to the length of this segment.
3. Put the point of the compass at one end of the segment and draw an arc over the center of the line

segment.

4. Put the point of the compass at the other end of the segment and draw another arc over the center of the line segment.
5. Draw a line segment from one end of the original segment to the intersection point of the two arcs.

The angle between the two line segments is 60° because the triangle formed by the two endpoints of the original line segment and the intersection point of the arcs is an equilateral triangle. You can now construct a 30° angle by bisecting the 60° angle.

Inquiry: Construct a line parallel to a given line

Using the fundamental constructions, we can draw a line parallel to a given line through any point not on the given line.

1. Begin with a line ℓ and a point P not on the line.
2. Mark an arbitrary point A on the given line.
3. Draw a line through the new point A and the original given point P.
4. Place the point of the compass at point A and draw an arc from line ℓ which intersects the new line, say at point B.
5. Without changing the width of the compass, place the point of the compass at point P and draw a (rough) copy of the first arc but centered at P. Be sure it intersects the new line. Let's call that point C.
6. Open the compass to the distance from point P to point A.
7. Place the point of the compass at point A and draw an arc which intersects
8. FIx this!

Inquiry: Construct a line perpendicular to a given line

Using the fundamental constructions, we can draw a line perpendicular to a given line at a point on the line.

1. TODO

Constructible Lengths

Now we consider the question of what segment lengths we can construct, assuming we have in hand some given segment lengths. The next lemma essentially says that all rational numbers are constructible.

Before stating and proving an important lemma, we have an inquiry which asks you to do a crucial step in one part of the proof.

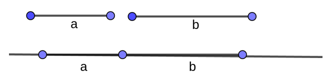
Inquiry: Using Similar Triangles

Consider the following image:

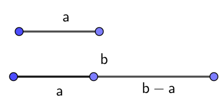
Here one long side of the triangle has length a, and a portion of that side is marked as having length x. Another side of the triangle has length b, and a portion of that side is marked as having length 1. Use similar triangles to prove that in this situation $x = \frac{a}{b}$.

Theorem 2.14. *Given segments of length 1, a and b, it is possible to construct segments of lengths $a + b$, $b - a$ (when $b > a$), ab, and a/b.*

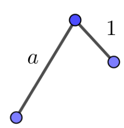
Proof. It is easy to see how to construct segments of length $a + b$ and $a - b$, given segments of length a and b. For the sum, suppose we have segments of length a and b. Extend the segment of length a to a line using the first fundamental construction. Then set the compass with to the length b, put the point at one endpoint of the segment of length a, and use the compass mark a point on the line a distance of b from that endpoint in the direction opposite the other endpoint. The resulting long segment has length $a + b$.



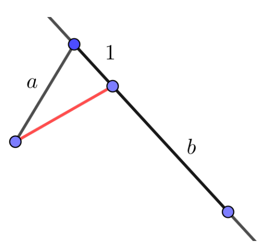
For $b - a$ when $b > a$, draw a segement of length a starting from one endpoint of the segment of length b, toward the other endpoint of that segment. What's left is a segment of length $b - a$.

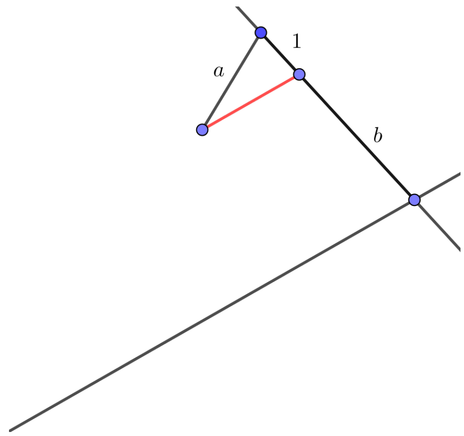


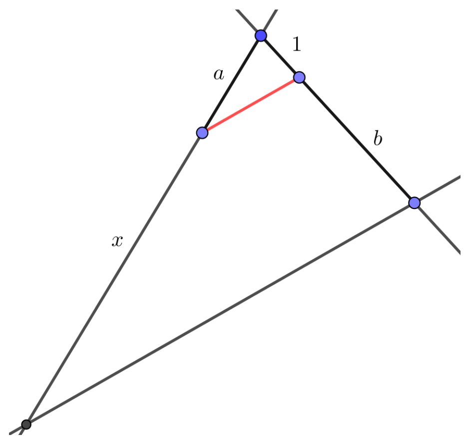
Now suppose that we have given segments of lengths a and b. How can we construct a segment of length ab? This is a little more involved and we will do it in multiple steps. First, we pick a point and draw segments of length a and length 1 emanating from that point with some angle θ between them with $0^\circ < \theta < 180^\circ$



Now we construct two segments. One segment joints the two non-overlapping endpoints of our segment of length a and our segment of length 1. We also construct a segment of length b adjacent to the segment of length 1 and along the line determined by that segment.







This shows that a segment of length of ab can be constructed if we are given segments of length a, b, and 1. As for a/b, the diagram in the inquiry above shows that a segment of length a/b can be constructed.

To complete the proof for this case, you must tie up two loose ends. First, explain how to construct this diagram using the fundamental constructions. Second, the diagram assumes that $b > 1$. We need a diagram which handles the case $b \leq 1$. The tieing of these loose ends is left to the reader. □

Definition 2.15. A real number a is constructible if given initially a segment of length 1, it is possible to construct a segment of length $|a|$.

This definition together with Theorem 2.14 allow us to conclude that every rational number is constructible. How is this? It should be clear that if 1 is constructible, then so is 2. After all, it's just $1 + 1$, and if we can construct a segment of length one, then we can construct two of them side-by-side. And similarly, we can construct segments with the length of any positive natural number. This allows us to conclude that every integer is constructible. But ratios of constructible numbers are constructible, so we get that every rational number is construcitble. So, we have that if 1 is constructible, then so is every natural number. Well, is 1 a constructible length? Well, recall that our straightedge is not a ruler - there are no distance marks. In particular, the length of "1" is not given to us. We are allowed to decide it. That is, we are free to just create a segment of any length and call it our unit length. Then every other length is measured relative to that. We can create a segment twice its length, three times its length,etc. Then we're off to the races and we can construct any rational length.

This is all great, but suggests the question of whether or not some non-rational numbers are constructible. The answer is yes, but remember that to show that a particular non-rational number is constructible, we have to construct a line segment of that length.

INQUIRY: REMEMBER THE EQUILATERAL TRIANGLE

Above we showed that we could construct angles of $60°$ and $30°$. We did that by

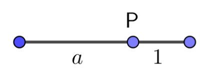
1. constructing an equilateral triangle, and then
2. bisecting one of the angles in the triangle.

Can we generate any constructible numbers from this construction? Use your knowledge of compass and straightedge constructions and trigonometry to find some non-rational constructible numbers from this example.

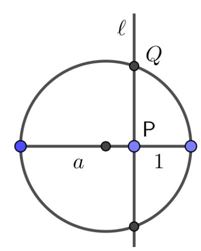
INQUIRY: CONSTRUCTING SQUARE ROOTS

Here we suppose that we are given segments of length 1 and length a, and we construct a segment of length $\sqrt{a}$. To do so, follow these steps.

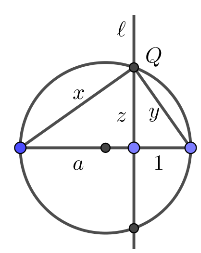
1. Construct a segment of length $a + 1$. Let P be the point at which the segment of length a meets the segment of length 1.



2. At point P construct a line ℓ perpendicular to the segment.
3. Bisect the segment, and mark its midpoint.
4. Place the point of the compass at the midpoint of the segment, open the compass to half the width of the segment, and draw a circle centerted at the midpoint whose radius is half the length of the segment.
5. The circle intersects the perpendicular bisector ℓ at two points. Pick one of them to work with, and let's call it Q.



6. Let x be the distance from one endpoint of the segment to Q and let y be the distance from Q to the other endpoint. Let z be the distance from the segment to Q.



7. There are three right triangles in your picture. Write down the corresponding three instances of the Pythagorean theorem.
8. Use these equations to show that $z = \sqrt{a}$.

Lemma 2.16. *Given segments of length* 1 *and* a*, a segment of length* $\sqrt{a}$ *may be constructed.*

Proof. You've just proved this in the inquiry above. $\square$

Notice that, for example, since 1 is constructible, so is 2. This theorem allows us to conclude that $\sqrt{2}$ is constructible. So is $\sqrt{3}$, $\sqrt{5}$, etc.

Inquiry: Am I Construcible?

In this task, you will be able to conclude based on our previous work that some of the numbers listed are constructible. For others, you will not be able to draw that conclusion. Below, we'll present a theorem that exactly characterizes which numbers are construcible and which are not. Right now you can only identify if something is constructible. Showing something is not constructible is harder. For each number listed below indicate "Constructible" or "No Conclusion." If the number is constructible, explain how you would construct it given our previous constructions.

- 100
- $\frac{3}{2}$
- $\sqrt{7}$
- $\sqrt[3]{7}$
- $3+\sqrt{7}$
- $\sqrt{3+\sqrt{7}}$
- $\sqrt{\sqrt{\sqrt{3+\sqrt{7}}}}$
- $\sqrt[4]{3+\sqrt{7}}$
- $\sqrt[5]{3+\sqrt{7}}$
- $\sqrt[6]{3+\sqrt{7}}$
- $\sqrt[7]{3+\sqrt{7}}$
- $\sqrt[n]{3+\sqrt{7}}$, where n is even.
- $\sqrt[n]{3+\sqrt{7}}$, where n is odd.
- π

Quadratic Extensions

Now we come to the important relationship between constructible numbers and field extensions, particularly so called "quadratic extensions." First, recall the following result.

Theorem 2.17. *If* $\mathbb{F}$ *is a field, then so is* $\mathbb{F}(\sqrt{k})$*.*

Proof. This was proved as homework problem XXXX. $\square$

First we give a sufficient condition for a number to be constructible.

Theorem 2.18. *A number* a *is constructible if there is a finite sequence of fields* $\mathbb{Q}=\mathbb{F}_0\subseteq\mathbb{F}_1\subseteq\ldots\subseteq\mathbb{F}_N$ *with* $a\in\mathbb{F}_N$ *and such that for each* j*,* $0\le j\le N-1$*,* $\mathbb{F}_{j+1}$ *is a quadratic extension of* $\mathbb{F}_j$*.*

Proof. The proof of this is by induction on N. First consider the case where $N=0$. Since $\mathbb{F}_0=\mathbb{Q}$ and we know that every rational number is constructible, it follows that the theorem is true for $N=0$. Suppose now that we have a sequence of field extensions

$$\mathbb{Q}=\mathbb{F}_0\subseteq\mathbb{F}_1\subseteq\ldots\subseteq\mathbb{F}_N$$

and that each j, $0 \le j \le N-1$, $\mathbb{F}_{j+1}$ is a quadratic extension of $\mathbb{F}_j$. Suppose further that every number in $\mathbb{F}_N$ is constructible, and that $\mathbb{F}_{N+1}$ is a quadratic extension of $\mathbb{F}_N$. In particular, let us suppose that $F_{N+1} = \mathbb{F}_N(\sqrt{k})$ where $k \in \mathbb{F}_N$. Suppose that $a \in \mathbb{F}_{N+1}$. Then

$$a = b + c\sqrt{k}$$

where $b, c \in \mathbb{F}_N$. By hypothesis, every number in $\mathbb{F}_N$ is contructible. So b and c are construcible. Since $k \in \mathbb{F}_N$ it is also constructible. We know that we can construct the square root of any constructible number, so it follows that $\sqrt{k}$ is constructible. Products of constructible numbers are constructible, so $c\sqrt{k}$ is constructible. Sums of constructible numbers are constructible, so

$$a = b + c\sqrt{k}$$

is constructible. It follows then that every number in $\mathbb{F}_{N+1}$ is constructible.

Therefore it follows by induction that for any N if there is a finite sequence of fields $\mathbb{Q} = \mathbb{F}_0 \subseteq \mathbb{F}_1 \subseteq \ldots \subseteq \mathbb{F}_N$ such that for each j, $0 \le j \le N-1$, $\mathbb{F}_{j+1}$ is a quadratic extension of $\mathbb{F}_j$ and $a \in \mathbb{F}_N$, then a is constructible. □

Definition 2.19. If $\mathbb{F}$ is a field, the plane of $\mathbb{F}$ will denote the set of all points (x, y) in the Cartesian plane so that x and y are in $\mathbb{F}$. By a line in $\mathbb{F}$ we mean a line passing through two points in the plane of $\mathbb{F}$. By a circle in $\mathbb{F}$ we mean a circle with both its center and some point on its circumference in the plane of $\mathbb{F}$.

Note that any fundamental construction using only points in the plane of a field $\mathbb{F}$ involves the construction of a line, line segment, or a circle in $\mathbb{F}$. To see this, recall the three fundamental constructions:

1. Given two points, we may draw a line through them, extending it indefinitely in each direction.
2. Given two points, we may draw the line segment connecting them.
3. Given a point and line segment, we may draw a circle with center at the point and radius equal to the length of the line segment.

If we are using only points in the plane of $\mathbb{F}$, then, in particular, the two points we start with in the first two fundamental constructions must be points in the plane of $\mathbb{F}$. Thus the line or line segment constructed is in $\mathbb{F}$, as defined above. What about the third fundamental construction? In this case, the center of the circle (x, y) must be in the plane of $\mathbb{F}$ and the length of the line segment r that determines the radius must be in $\mathbb{F}$. But then the point $(x + r, y)$ is on the circle and is also in the plane of $\mathbb{F}$.

Lemma 2.20. *Every line in $\mathbb{F}$ can be represented by an equation of the form $ax + by + c = 0$ with $a, b, c \in \mathbb{F}$*

Proof. Suppose that (x_1, y_1) and (x_2, y_2) are points on a line in $\mathbb{F}$ and that these two points are in the plane of $\mathbb{F}$ so that x_1, x_2, y_1, y_2 are all in $\mathbb{F}$. Then, if the line is not vertical, then the slope of the line is

$$m = \frac{y_2 - y_1}{x_2 - x_1}.$$

That number m is in $\mathbb{F}$ if all of x_1, x_2, y_1, and y_2 are. The point-slope equation of the line is then

$$y - y_1 = m(x - x_1).$$

We can rewrite this as $ax + by + c = 0$ where $a = m$, $b = -1$, and $c = y_1 - x_1$, and these numbers a, b, c are all in $\mathbb{F}$ if x_1, x_2, y_1, and y_2 are. If the line is vertical, then both of the points are of the form (c, y_1), (c, y_2) for some fixed c in the x-coordinate and the equation of the line is $x = c$, or, if you like, $x - c = 0$. □

Lemma 2.21. *Every circle in $\mathbb{F}$ can be represented by an equation of the form $x^2 + y^2 + ax + by + c = 0$ with $a, b, c \in \mathbb{F}$.*

Proof. The proof of this is the same idea as the proof of Lemma 2.20. We start with the center of the circle (x_0, y_0) which is assumed to be in the plane of $\mathbb{F}$ and a point on the circumference of the circle (x_1, y_1) which is also assumed to be in the plane of $\mathbb{F}$. Then the radius of the circle is

$$r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

which is in not necessarily in $\mathbb{F}$, but r^2 definitely is. The equation of the circle is then

$$(x - x_0)^2 + (y - y_0)^2 = r^2.$$

Expanding the terms on the left hand side and rearranging can clearly get us an equation of the form $x^2 + y^2 + ax + by + c = 0$ with $a, b, c \in \mathbb{F}$. □

Theorem 2.22.

1. *The point of intersection of two distinct, nonparallel lines in $\mathbb{F}$ is in the plane of $\mathbb{F}$.*
2. *The points of intersection of a line in $\mathbb{F}$ and a circle in $\mathbb{F}$ are either in the plane of $\mathbb{F}$ or in the plane of some quadratic extension of $\mathbb{F}$.*
3. *The points of intersection of two circles in $\mathbb{F}$ are either in the plane of $\mathbb{F}$ or in the plane of some quadratic extension of $\mathbb{F}$.*

Proof. You can probably imagine how this is proved. For item (1) we start with two distinct, nonparallel lines. As above, these have equations

$$a_1x + b_1y + c_1 = 0$$
$$a_2x + b_2y + c_2 = 0$$

$a_1, b_1, c_1, a_2, b_2, c_2 \in \mathbb{F}$. Now we solve for the intersection point. The lines must have exactly one intersection point because they are distinct, nonparallel lines, and that point will have coordinates which use only sums, differences, products and quotients of the coefficients. Since the coefficinets are in $\mathbb{F}$ and $\mathbb{F}$ is a field, it follows that each of the coordinates of the intersection point will also be in $\mathbb{F}$.

The situation is similar for item (2), except now we have a line and a circle so our equations look like

$$a_1x + b_1y + c_1 = 0$$
$$x^2 + y^2 + a_2x + b_2y + c_2 = 0$$

with $a_1, b_1, c_1, a_2, b_2, c_2 \in \mathbb{F}$. If we simultaneously solve these two equations we'll use all $(+, -, \cdot, /)$ of our arithemtic operations but we will also likely have to do a square root. (You can either imagine this or really try it!) If no square roots are needed, then our intersection point(s) are in the plane of $\mathbb{F}$. If we do need to take a square root, then our intersection points will be in the plane of a quadratic extension of $\mathbb{F}$.

In case (3) we now have two circles and our equations look like

$$x^2 + y^2 + a_1x + b_1y + c_1 = 0x^2 + y^2 + a_2x + b_2y + c_2 \qquad = 0$$

with $a_1, b_1, c_1, a_2, b_2, c_2 \in \mathbb{F}$. It's not so bad to simultaneously solve these two equations. We start by subtracting the second equation from the first and we get that

$$(a_1 - a_2)x + (b_1 - b_2)y + (c_1 - c_2) = 0.$$

So,

$$y = \frac{c_2 - c_1}{a_1 - a_2} + \frac{b_2 - b_1}{c_1 - c_2}x.$$

Now we plug this into one of our original equations to get a quadratic in x. We can use the quadratic formula to find either zero, one or two values of x. That means the circles intersect in zero points, one point, or two points. After we find the x values we determine the y values. Again, to find the x-values we just take one square root. As a result, our intersection points will either be in the plane of $\mathbb{F}$ or the plane of a quadtric extension of $\mathbb{F}$. □

Theorem 2.23. *The following statements are equivalent:*

1. *The number a is constructible.*
2. *There is a finite sequence of fields $\mathbb{Q} = \mathbb{F}_0 \subseteq \mathbb{F}_1 \subseteq \ldots \subseteq \mathbb{F}_N$ with $a \in \mathbb{F}_N$ and such that for each j, $0 \leq j \leq N-1$, $\mathbb{F}_{j+1}$ is a quadratic extension of $\mathbb{F}_j$.*

Proof. We've already proved that (2) implies (1) above. And, we've done most of the hard work for (1) implies (2). To see that (1) implies (2), suppose that a is constructible. That means there was some finite sequence of fundamental constructions that resulted in the construction of a segment of length a. We start with a segment of length 1 whose endpoints we assume to have coordinates $(0,0)$ and $(1,0)$. After the first fundamental construction we produce a line segment, line, or circle in the plane of $\mathbb{Q}$ or in the plane of some quadratic extension of $\mathbb{Q}$. At each stage we produce new points by intersecting our existing constructions with our new constructed objects. These intersection points are either in the plane of the field we are working over, or in the plane of some quadtratic extension of that field. Since our construction of a segment of length a must terminate after a finite number of steps, it follows that if a is constructible, then there is a finite sequence of fields $\mathbb{Q} = \mathbb{F}_0 \subseteq \mathbb{F}_1 \subseteq \ldots \subseteq \mathbb{F}_N$ with $a \in \mathbb{F}_N$ and such that for each j, $0 \leq j \leq N-1$, $\mathbb{F}_{j+1}$ is a quadratic extension of $\mathbb{F}_j$. □

3 Three Famous Problems

Doubling the Cube

Given a line segment representing the edge of a cube, is it possible to construct another line segment representing the edge of a cube with exactly twice the volume of the first cube?

Without loss of generality we will take the length of a side of the original cube to be 1. Then the desired line segment must have length $\sqrt[3]{2}$.

Theorem 3.1. *Let $\mathbb{F}(\sqrt{k})$ be a quadratic extension of a field $\mathbb{F}$. If $\sqrt[3]{2}$ is in $\mathbb{F}(\sqrt{k})$, then $\sqrt[3]{2}$ must be in $\mathbb{F}$ itself.*

Theorem 3.2. *It is impossible to double the cube.*

Trisecting an Angle

TODO: We show that it is impossible to trisect an angle of 60°. If this were possible, it would be possible to construct a 20° angle. This implies that $\cos(20^\circ)$ is constructible. This implies that a root of $x^3 - 3x - 1 = 0$ is constructible.

Theorem 3.3. *If $\mathbb{F}(\sqrt{k})$ contains a root of $x^2 - 3x - 1 = 0$, then so does $\mathbb{F}$.*

Theorem 3.4. *It is not possible to trisect and arbitrary angle.*

Squaring a Circle