

An introduction to finite fields

N J Wildberger

October 28, 2005

Abstract

This note is a short introduction to the prime finite fields, emphasizing the analogies with the arithmetic of fractions. It is meant for high school students.

1 What is a field?

A **field** is a framework in which we may do arithmetic. It consists of *elements* which may be added and multiplied subject to certain laws such as the *commutative*, *associative* and *distributive* laws. These are based on familiar properties of the usual operations $+$ and \times on fractions. Any field is required to contain special elements which will always be called 0 , 1 and -1 and which have particular properties. So a field is just an algebraic framework for the arithmetic learnt in primary school. As such, fields ought to be studied much earlier than they are. Currently students of mathematics learn about them in Algebra courses in their second or third year of university study. But a field is a simple and important enough notion to warrant study in high schools. I believe that many high school students would find the study of finite fields particularly pleasant and interesting.

There are many examples of fields. The most basic and important example is the field of rational numbers. This is the model on which all other fields are based. Other fields include the field of decimal numbers, the field of complex numbers and the prime fields F_p where p is a prime number. These latter are examples of finite fields, containing only a finite number of elements, and they are computationally much simpler than other fields. There are many other interesting, and somewhat more complicated examples of fields. This note tries to give a precise but brief introduction to the finite prime fields, often denoted F_p .

[For those with a standard background: I happen not to believe in the mysticism of ‘infinite set theory’. So the usual framework is slightly recast to avoid these. I also happen to believe that the ‘field of order two’ is special and warrants a separate treatment; it should not be considered a field in the same sense as those of odd prime order.]

2 What is the field of rational numbers?

Recall that an **integer** is an element of the doubly infinite sequence

$$\dots, -3, -2, -1, 0, 1, 2, 3, \dots$$

A **rational number** is just a fraction, in other words an expression of the form

$$\frac{a}{b}$$

where a and b are integers with b non-zero. This fraction is also denoted a/b . There is an important notion of equality between fractions:

$$a/b = c/d \quad \text{precisely when} \quad ad - bc = 0.$$

Thus for example

$$\frac{2}{3} = \frac{4}{6} = \frac{14}{21} = \frac{-2}{-3}.$$

The fraction $a/1$ is often written simply as a , so that every integer is regarded as a rational number. In particular the three integers $0, 1$ and -1 are rational numbers. The operations of addition $+$ and multiplication \times of rational numbers are defined as follows:

$$\begin{aligned} \frac{a}{b} + \frac{c}{d} &\equiv \frac{ad + bc}{bd} \\ \frac{a}{b} \times \frac{c}{d} &\equiv \frac{ac}{bd}. \end{aligned}$$

The operation of multiplication is often written without the multiplication sign, so that for example

$$\frac{2}{3} \times \frac{5}{7} = \left(\frac{2}{3}\right) \left(\frac{5}{7}\right) = \frac{10}{21}$$

or using symbols

$$x \times y = xy.$$

Another convention is that multiplication takes precedence over addition, so that for example

$$\begin{aligned} xy + z &= (xy) + z \\ xy + z &\neq x(y + z). \end{aligned}$$

One ought to check that the operations of addition and multiplication are well-defined, in other words that the results of an operation do not depend on which of two equivalent expressions for a fraction are used. For example since $2/3 = 4/6$ and $5/7 = 15/21$, it follows that

$$\frac{2}{3} + \frac{5}{7} = \frac{4}{6} + \frac{15}{21}.$$

The reader should convince themselves that this indeed works.

The operations of addition and multiplication of fractions obey the following laws. Here x, y and z represent arbitrary rational numbers.

Commutative laws

$$\begin{aligned}x + y &= y + x \\xy &= yx\end{aligned}$$

Associative laws

$$\begin{aligned}(x + y) + z &= x + (y + z) \\(xy)z &= x(yz)\end{aligned}$$

Distributive laws

$$\begin{aligned}x(y + z) &= xy + xz \\(x + y)z &= xz + yz.\end{aligned}$$

The numbers 0, 1 and -1 give rise to special laws also.

Identity laws For any x

$$\begin{aligned}0 + x &= x + 0 = x \\1 \times x &= x \times 1 = x \\1 + (-1) &= 0.\end{aligned}$$

Inverse law For any non-zero number x there is a number z with the property that

$$xz = zx = 1.$$

This number z , which depends on x , is denoted

$$z = x^{-1}.$$

Having these laws, it is also useful to define additional secondary operations of subtraction $-$ and division \div as follows:

$$\begin{aligned}x - y &\equiv x + (-1)y \\x \div y &\equiv x \times y^{-1} \text{ for } y \neq 0.\end{aligned}$$

You may then check that

$$\begin{aligned}\frac{a}{b} - \frac{c}{d} &= \frac{ad - bc}{bd} \\ \frac{a}{b} \div \frac{c}{d} &= \frac{ad}{bc} \text{ for } c \neq 0.\end{aligned}$$

There are many additional facts that can be deduced using these basic definitions and laws. It is an excellent exercise to carefully deduce the following rules from the basic laws mentioned above.

Cancellation rules i) If $a + b = a + c$ then $b = c$.

ii) If $a \times b = a \times c$ and $a \neq 0$ then $b = c$.

Zero divisor rule If $a \times b = 0$ then either $a = 0$ or $b = 0$.

Additive inverses For any number a there is exactly one number b such that $a + b = 0$. In fact $b = (-1) \times a$.

Algebraic identities i) $(a + b)^2 = a^2 + 2ab + b^2$

ii) $(a - b)^2 = a^2 - 2ab + b^2$

iii) $(a + b) \times (a - b) = a^2 - b^2$

(Of course there are many more important algebraic identities.)

That is a pretty concise summary of facts that every student in grade school should learn. But even more crucial than this abstract knowledge is a solid grasp of the mechanics of working with fractions. This is the most important basic mathematical skill. Students who learn to rely on calculators to deal with arithmetic are significantly handicapped. Let me restate it as follows: *the ability to understand and perform arithmetic with fractions using pen and paper is the most fundamental and important of all mathematical skills.*

3 What is a general field?

A **field** is an arithmetical framework consisting of the following ingredients, modelled directly from the previous example.

First there must be a notion of **number**, possibly involving a concept of equality. There must be three distinguished numbers called 0, 1 and -1 . A number is called an **element** of the field.

Secondly there must be operations of addition $+$ and multiplication \times defined on pairs of numbers. If there is a notion of equality between numbers, then these operations should be well-defined, that is independent of which representative(s) are used.

Thirdly these two operations must satisfy all the laws stated above, namely the commutative, associative, distributive, identity and inverse laws.

Then all the additional facts, such as the cancellation rules, additive inverses and algebraic identities hold automatically, since they are just consequences of the basic laws.

4 What is the smallest example of a field?

The smallest example of a field has the minimum number of elements: the numbers 0, 1 and -1 . Here are addition and multiplication tables for such a

field (it turns out to be uniquely determined- a good exercise).

+	0	1	-1	×	0	1	-1
0	0	1	-1	0	0	0	0
1	1	-1	0	1	0	1	-1
-1	-1	0	1	-1	0	-1	1

There are also degenerate fields; one of them has only two elements 0 and $1 = -1$, and the other has only one element $0 = 1 = -1$. Unfortunately the two element degenerate field is usually called a field, causing considerable awkwardness in higher mathematics.

5 What are the finite fields F_p ?

If p is an prime number (no positive factors other than 1 and itself) which is odd, such as

$$p = 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, \dots$$

then it turns out there is a finite field with exactly p elements. This field can be described by a simple modification of the description of the rational number field. An element of F_p is an expression of the form a/b where a and b are integers with b not divisible by p , and with the notion of equality

$$\frac{a}{b} = \frac{c}{d}$$

precisely when $ad - bc$ is divisible by p . Arithmetic in such a finite field might seem strange, but it has many applications (including internet security encryption!)

Note that the usual integers are still elements of F_p , since we may identify $a/1$ with the integer a . However they are no longer distinct, as

$$\begin{aligned} \dots - 2p &= -p = 0 = p = 2p = 3p = \dots \\ \dots - 2p + 1 &= -p + 1 = 1 = p + 1 = 2p + 1 = 3p + 1 = \dots \\ \dots - 2p + 2 &= -p + 2 = 2 = p + 2 = 2p + 2 = 3p + 2 = \dots \end{aligned}$$

and so on. Thus we may replace any integer at any time with another obtained by adding or subtracting a multiple of p , enabling us to replace any integer with one of the p possibilities

$$0, 1, 2, \dots, p - 1.$$

Suppose $p = 7$ for example. Then $24/37 = 3/2$ since

$$24 \times 2 - 37 \times 3 = -63$$

is divisible by 7. Alternatively $24 = 3$ and $37 = 2$ since the differences are multiples of 7. Furthermore $3/2 = 5/1$ since $3 \times 1 - 2 \times 5$ is divisible by 7. So $24/37 = 5$ in F_7 .

Now it turns out to be a beautiful fact that *every* element of F_p is equal to exactly one of the integers

$$0, 1, 2, \dots, p-1.$$

So the field F_p contains exactly p distinct elements. although any one of these has many different equivalent formulations as a fraction. The reason for this beautiful fact amounts to this. It turns out that if b is any integer not divisible by p , then there is another integer c with the property that $b \times c$ is exactly one more than a multiple of p . This implies that in the field F_p

$$b^{-1} = c.$$

That means that an expression of the form a/b can be rewritten as the integer ac . It further implies that all the arithmetic in F_p can be done with integers, as long as we know how to invert integers appropriately.

Here is an example. Suppose you want to simplify the fraction $4/3$ in F_7 . Now since $3 \times 5 = 15 = 1 + 2 \times 7$ it follows that $3^{-1} = 5$ in F_7 . Thus

$$\frac{4}{3} = 4 \times 5 = 20 = 6.$$

In this way any fraction in F_7 can be converted to one of the seven numbers

$$0, 1, 2, 3, 4, 5, 6.$$

Let's look at the addition and multiplication tables for F_7 . They have pleasant properties which you may investigate.

+	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
3	3	4	5	6	0	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

×	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	1	3	5
3	0	3	6	2	5	1	4
4	0	4	1	5	2	6	3
5	0	5	3	1	6	4	2
6	0	6	5	4	3	2	1

Note that $-1 = 6$ is not a square, in other words not of the form a^2 for some a . You may now check various special cases of the laws and rules mentioned above.

Here are the addition and multiplication tables in F_5 .

+	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

×	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

Note that $4 = -1 = 2^2$ is a square in F_5 . This difference between F_7 and F_5 turns out to be crucial in applications. Fields F_p where p is of the form $4k+3$ are like F_7 : the number -1 is not a square. Fields F_p where p is of the form $4k+1$ are like F_5 : the number -1 is a square. Generally speaking the former fields ($p = 4k+3$) are much easier to work with than the latter fields ($p = 4k+1$).

It's a good exercise to create similar tables for F_{11} and F_{13} .

6 Fermat's theorem in F_p

It is a consequence of the (multiplicative) Cancellation rule and the Zero divisor rule that for any non-zero a the numbers $a \times 1, a \times 2, \dots, a \times (p-1)$ in F_p are all distinct and non-zero. That means that these are just the $(p-1)$ numbers $1, 2, \dots, p-1$ in some possibly different order. Thus

$$1 \times 2 \times \dots \times (p-1) = (a \times 1) \times (a \times 2) \times \dots \times (a \times (p-1))$$

and so

$$1 \times 2 \times \dots \times (p-1) = a^{p-1} \times 1 \times 2 \times \dots \times (p-1).$$

But by the Zero divisor rule, $1 \times 2 \times \dots \times (p-1) \neq 0$. Thus applying the (multiplicative) Cancellation rule shows that

$$1 = a^{p-1}$$

for any non-zero a . This important result is called **Fermat's theorem**. As a consequence

$$a = a^p$$

for *any* number a in F_p .

7 Additional exercises:

1. Show that in F_7

$$\begin{aligned} \frac{534}{311} &= 3 \\ \frac{421}{32} \times \frac{411}{25} &= 6 \\ \frac{82}{97} - \frac{34}{55} &= 1 \\ 1^3 + 2^3 + 3^3 + 4^3 + 5^3 + 6^3 &= 0. \end{aligned}$$

3. Find solutions to the following equations (if they exist) in F_7 .

$$\begin{array}{lll} \text{i) } 3x = 5 & \text{ii) } x^2 = 2 & \text{iii) } x^2 = 3 \\ \text{iv) } x^2 = 2 & \text{v) } 4x^2 - 3x = 2 & \text{vi) } 2x^3 - x^2 + 4x = 6. \end{array}$$