

---

---

The University of New South Wales  
School of Mathematics and Statistics

**MATH3521**

**ALGEBRAIC TECHNIQUES IN NUMBER THEORY**

**CHAPTERS 5–8**

by THOMAS BRITZ\*

2008

---

---

**ABOUT THE NOTES.**

These course notes were written during April and May of 2008, for the second half of the 3rd-year course *MATH3521 Algebraic Techniques in Number Theory* at the School of Mathematics and Statistics of the University of New South Wales, Sydney, Australia. The notes were written in  $\text{\TeX}$  using a very nice layout and set of macros by David Angell. The contents were initially inspired in part by similar course notes by P. Brown and by A. An Heuf, and by the book *Elementary Theory of Numbers* by Harriet Griffin (McGraw-Hill, New York, 1954).

Please let me know if you spot any mistakes!

Best regards, Thomas  
(Sydney, June 2008).

---

\* [britz@math.unsw.edu.au](mailto:britz@math.unsw.edu.au)

---

---

The University of New South Wales  
School of Mathematics and Statistics

MATH3521

ALGEBRAIC TECHNIQUES IN NUMBER THEORY

---

---

5. QUADRATIC RECIPROCITY.

Some non-linear congruence equations modulo  $m$  were presented in earlier chapters. We now consider one of the simplest families of such equations, namely the family of quadratic equations having the form

$$x^2 \equiv a \pmod{p}$$

where  $p$  is a prime. As the following example shows, many quadratic congruence equations may be reduced to such simple equations.\*

**Problem.** Solve  $x^2 + 4x + 2 \equiv 0 \pmod{7}$ .

**Solution.** Re-write this congruence equation as  $(x + 2)^2 \equiv 2 \pmod{7}$ . Now solve  $y^2 \equiv 2 \pmod{7}$  with  $y = x + 2$  by simply testing each  $y \in \mathbb{Z}_7$ ; we find that  $y \equiv 3$  or  $4 \pmod{7}$ . That is,  $x \equiv 1$  or  $2 \pmod{7}$ .

The next theorem describes the number of solutions to our equation.

**Theorem.** If  $p$  is an odd prime and  $\gcd(a, p) = 1$  (so  $a \not\equiv 0 \pmod{p}$ ), then  $x^2 \equiv a \pmod{p}$  has exactly 2 solutions or it has no solutions.

**Proof.** Suppose that  $x$  is a solution; then  $-x$  is also a solution. Since  $p$  is odd and  $\gcd(a, p) = 1$ ,  $2x \not\equiv 0 \pmod{p}$ , so  $x \not\equiv -x \pmod{p}$ . Therefore,  $-x$  and  $x$  are two distinct solutions. Suppose that  $y$  is also a solution. Then  $(x + y)(x - y) = x^2 - y^2 \equiv 0 \pmod{p}$ . But  $\mathbb{Z}_p$  has no zero divisors, so  $y \equiv \pm x \pmod{p}$ ; thus,  $-x$  and  $x$  are the only solutions.

The above theorem is not always true for non-primes; for example,  $x^2 \equiv 1 \pmod{8}$  has 4 solutions.

---

\* In fact, all quadratic congruence equations may be reduced in this way if one is allowed to change the modulus to a non-prime.

**Quadratic Residues.** Every positive integer is either a square or a non-square. This is of course also true for integers modulo  $p$  but squares in  $\mathbb{Z}_p$  are often called something different in order to avoid confusing the two sorts of squares:

**Definition.** An integer  $a \neq 0$  is a **quadratic residue** if  $x^2 \equiv a \pmod{p}$  has a solution; otherwise,  $a$  is a **quadratic non-residue**.

**Example.** Consider the squares of each element of  $\mathbb{U}_{11}$ :

$x$	1	2	3	4	5	6	7	8	9	10
$x^2$	1	4	9	5	3	3	5	9	4	1

The quadratic residues modulo 11 are  $\{1, 4, 9, 5, 3\}$ , and the quadratic non-residues are  $\{2, 6, 7, 8, 10\}$ . We see that there are just as many quadratic residues as non-quadratic residues. This is always true:

**Theorem.** If  $p$  is an odd prime, then exactly half of the integers  $1, 2, \dots, p - 1$  are quadratic residues modulo  $p$ ; furthermore, these are the even powers of any primitive root modulo  $p$ .

**Proof.** Let  $g$  be a primitive root modulo  $p$ . Then  $\mathbb{U}_p = \{g, g^2, \dots, g^{p-1}\}$ . The even powers of  $g$  are clearly squares and comprise half of the integers  $1, 2, \dots, p - 1$ , so we must show that no square in  $\mathbb{U}_p$  is an odd power of  $g$  modulo  $p$ . For each square  $g^{2s}$  with  $2s < p$ , write  $g^{2s} \equiv g^N \pmod{p}$  for some integer  $N \geq 0$ . Then  $g^{N+(p-1-2s)} \equiv g^{p-1} \equiv 1 \pmod{p}$ , so  $\phi(p)$ , the order of  $g$ , divides  $N + (p - 1 - 2s)$ . Since  $\phi(p) = p - 1$  and  $p - 1 - 2s$  are even,  $N$  is also even.

**Example.** The element 2 is a primitive root modulo 11 with even powers  $\{2^2, 2^4, 2^6, 2^8, 2^{10}\} \equiv \{4, 5, 9, 3, 1\} \pmod{11}$ . By theorem above, the squares in  $\mathbb{U}_{11}$  are  $\{1, 3, 4, 5, 9\}$ , as found in the preceding example.

**Euler's Criterion.** The quadratic congruence equation  $x^2 \equiv a \pmod{p}$  has a solution if and only if  $a$  is a quadratic residue modulo  $p$ . Therefore, we want to determine when an integer is a quadratic residue modulo  $p$ . For small numbers, we can simply calculate the squares modulo  $p$  but this is not easy for big numbers; for example, is 3127 a square in  $\mathbb{Z}_{12713}$ ? L. P. Euler gave the following criteria for quadratic residues:

**Euler's Criterion.** If  $p$  is an odd prime that does not divide  $a$ , then

- $x^2 \equiv a \pmod{p}$  has a solution  $\Leftrightarrow a^{\frac{p-1}{2}} \equiv 1 \pmod{p}$ ; equivalently,
- $x^2 \equiv a \pmod{p}$  has no solution  $\Leftrightarrow a^{\frac{p-1}{2}} \equiv -1 \pmod{p}$ .

**Proof.** Let  $g$  be a primitive root modulo  $p$  and write  $a \equiv g^N \pmod{p}$  for some integer  $N \geq 0$ . By the preceding theorem,  $a$  is a quadratic residue modulo  $p$  if and only if  $N$  is even. If  $N$  is even, then

$$a^{\frac{p-1}{2}} \equiv (g^{2k})^{\frac{p-1}{2}} \equiv (g^{p-1})^k \equiv 1 \pmod{p},$$

by Fermat's Little Theorem (page 87). Conversely, if  $a^{\frac{p-1}{2}} \equiv 1 \pmod{p}$ , then  $g^{N\frac{p-1}{2}} \equiv 1 \pmod{p}$ , so  $p-1$  divides  $N\frac{p-1}{2}$ , and  $N$  must be even.

Finally,  $(a^{\frac{p-1}{2}})^2 \equiv a^{p-1} \equiv 1 \pmod{p}$ . Thus,  $a^{\frac{p-1}{2}} \equiv \pm 1 \pmod{p}$ ; this shows the equivalence of the two statements in the theorem.

**Example.** To determine whether 3 is a quadratic residue modulo 23, we can use the criteria of Euler's theorem: since  $3^{\frac{23-1}{2}} \equiv 3^{11} \equiv 1 \pmod{23}$ , we see that 3 is in fact a quadratic residue modulo 23.

**Legendre's Symbol.** Euler's criterion is still fairly difficult to use on large numbers, so we need more practical methods for dealing with quadratic residues. To this end, we introduce some easy notation to avoid long and awkward phrases like "a is a quadratic residue modulo p".

**Definition.** The **Legendre symbol**  $\left(\frac{a}{p}\right)$  is given by

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue mod } p; \\ -1 & \text{if } a \text{ is a quadratic non-residue mod } p; \\ 0 & \text{if } a \text{ is neither (i.e., if } a \equiv 0 \pmod{p} \text{).} \end{cases}$$

Euler's Criterion can then be expressed as follows:

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod{p}.$$

**Example.** By the preceding example,  $\left(\frac{3}{23}\right) = 1$ .

We summarise our results so far:

**Theorem.** Let  $p$  be an odd prime and  $a$  be an integer with  $\gcd(a, p) = 1$ . Then the following statements are equivalent:

- $x^2 \equiv a \pmod{p}$  has a solution;
- $x^2 \equiv a \pmod{p}$  has precisely two solutions;
- $a$  is a quadratic residue modulo  $p$ ;
- $a$  is an even power of some primitive root modulo  $p$ ;
- $a$  is not an odd power of any primitive root modulo  $p$ ;
- $a^{\frac{p-1}{2}} \not\equiv -1 \pmod{p}$ ;
- $a^{\frac{p-1}{2}} \equiv 1 \pmod{p}$ ;
- $\left(\frac{a}{p}\right) \neq -1$ ;
- $\left(\frac{a}{p}\right) = 1$ .

Apart from providing simple notation, the Legendre symbol is useful due to several simple properties that we can use to evaluate it quickly:

**Theorem.** Let  $p$  be an odd prime and  $a$  and  $b$  be any integers. Then

- if  $a \equiv b \pmod{p}$ , then  $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$ ;
- $\left(\frac{a^2}{p}\right) = 1$  unless  $a \equiv 0 \pmod{p}$ ;
- $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$ ;
- $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$ .

**Proof.** Exercise.

**Examples.** Some Legendre symbols are calculated by the rules above:

- $\left(\frac{47}{17}\right) = \left(\frac{-4}{17}\right) = \left(\frac{4}{17}\right)(-1)^8 = \left(\frac{2^2}{17}\right) = 1$ .
- $\left(\frac{6}{11}\right) = \left(\frac{2}{11}\right)\left(\frac{3}{11}\right) = \left(\frac{2}{11}\right)\left(\frac{-8}{11}\right) = \left(\frac{2}{11}\right)\left(\frac{8}{11}\right)(-1)^5 = -\left(\frac{4^2}{11}\right) = -1$ .
- $\left(\frac{5}{23}\right) = \left(\frac{-64}{23}\right) = \left(\frac{8^2}{23}\right)(-1)^{11} = -1$ .

The following result follows from the theorem on page 126:

**Proposition.** If  $p$  is an odd prime, then  $\left(\frac{1}{p}\right) + \left(\frac{2}{p}\right) + \cdots + \left(\frac{p-1}{p}\right) = 0$ .

We can now determine when there are solutions and what these solutions are for a special case of the equation  $x^2 \equiv a \pmod{p}$ :

**Theorem.**  $x^2 \equiv -1 \pmod{p}$  has solutions iff  $p \equiv 1 \pmod{4}$  or  $p = 2$ ; if  $p \equiv 1 \pmod{4}$ , then  $\pm\left(\frac{p-1}{2}\right)!$  are the solutions.

**Proof.** The  $p = 2$  case is trivial so consider  $p > 2$ . Then there are solutions if and only if  $(-1)^{\frac{p-1}{2}} = 1$ , i.e., if and only if  $p \equiv 1 \pmod{4}$ .

Suppose that  $p \equiv 1 \pmod{4}$ . Then  $(-1)^{\frac{p-1}{2}} = 1$ . Noting that  $p - i \equiv -i \pmod{p}$  for each  $i = 1, \dots, \frac{p-1}{2}$ , we apply Wilson's Theorem:

$$\begin{aligned} \left(\pm\left(\frac{p-1}{2}\right)!\right)^2 &\equiv \left(\left(\frac{p-1}{2}\right)!\right)^2 (-1)^{\frac{p-1}{2}} \\ &\equiv 1 \cdot 2 \cdots \left(\frac{p-1}{2}\right) \left(-\frac{p-1}{2}\right) \cdots (-2)(-1) \pmod{p} \\ &\equiv 1 \cdot 2 \cdots \left(\frac{p-1}{2}\right) \left(\frac{p+1}{2}\right) \cdots (p-2)(p-1) \pmod{p} \\ &\equiv (p-1)! \pmod{p} \\ &\equiv -1 \pmod{p}. \end{aligned}$$

**Example.**  $x^2 \equiv -1 \pmod{17}$  has two solutions since  $17 \equiv 1 \pmod{4}$ . Indeed, these are  $x \equiv \pm\left(\frac{17-1}{2}\right)! \equiv \pm 8! \equiv \pm 13 \pmod{17}$ .

**Two results by Gauss.** The results above for calculating  $\left(\frac{a}{p}\right)$  are useful for small primes. For bigger primes, better tools are presented below in this chapter, including two results by K. F. Gauss, namely Gauss' Lemma and the Law of Quadratic Reciprocity.

**Gauss' Lemma.** Let  $p$  be an odd prime and  $a \in \mathbb{Z}$  with  $\gcd(a, p) = 1$ . Define the set  $S = \{a, 2a, \dots, \frac{p-1}{2}a\}$  with elements reduced modulo  $p$ , and let  $k = |\{s \in S : s > \frac{p-1}{2}\}|$ . Then  $\left(\frac{a}{p}\right) = (-1)^k$ .

**Proof.** The lemma is easy to verify if  $a = \pm 1$  so suppose that  $a \neq \pm 1$ . Since  $\gcd(a, p) = 1$ , the elements  $S = \{a_1, \dots, a_{\frac{p-1}{2}}\}$  are distinct. Let  $S' = \{s \in S : s > \frac{p-1}{2}\}$ . We may assume that  $S' = \{a_1, a_2, \dots, a_k\}$ . Then

$$p - a_1, p - a_2, \dots, p - a_k, a_{k+1}, \dots, a_{\frac{p-1}{2}}$$

are all less than or equal to  $\frac{p-1}{2}$ . Suppose that  $p - a_i = a_j$  for some  $i, j$ . Then  $p = a_i + a_j$  and thus  $a \mid p$ , or  $\gcd(a, p) = a \neq \pm 1$ , contradicting  $\gcd(a, p) = 1$ . Thus, the above numbers are distinct and must therefore be the numbers  $1, 2, \dots, \frac{p-1}{2}$  (in some order). It follows that

$$\begin{aligned} \left(\frac{p-1}{2}\right)! &\equiv (p - a_1)(p - a_2) \cdots (p - a_k) a_{k+1} \cdots a_{\frac{p-1}{2}} \pmod{p} \\ &\equiv (-a_1)(-a_2) \cdots (-a_k) a_{k+1} \cdots a_{\frac{p-1}{2}} \pmod{p} \\ &\equiv (-1)^k a(2a) \cdots \left(\frac{p-1}{2}a\right) \pmod{p} \\ &\equiv (-1)^k \left(\frac{p-1}{2}\right)! a^{\frac{p-1}{2}} \pmod{p}, \end{aligned}$$

so  $a^{\frac{p-1}{2}} \equiv (-1)^k \pmod{p}$ .

**Example.** Calculate  $\left(\frac{3}{11}\right)$  and  $\left(\frac{8}{29}\right)$ .

- For  $p = 11$  and  $a = 3$ ,  $\{a, 2a, \dots, \frac{p-1}{2}a\} = \{3, 6, 9, 12, 15\}$ , so

$$S = \{3, 6, 9, 1, 4\} \quad \text{and} \quad k = |\{s \in S : s > 5\}| = 2.$$

By Gauss' Lemma,  $\left(\frac{3}{11}\right) = (-1)^2 = 1$ .

- To minimize calculations, first note that  $\left(\frac{8}{29}\right) = \left(\frac{2^2}{23}\right)\left(\frac{2}{23}\right) = \left(\frac{2}{23}\right)$ . For  $p = 29$  and  $a = 2$ ,

$$S = \{2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28\},$$

so  $k = |\{s \in S : s > 14\}| = 8$ . By Gauss' Lemma,  $\left(\frac{8}{23}\right) = (-1)^8 = 1$ .

By itself, Gauss' Lemma is useful but not exceedingly efficient. However, it is a powerful theoretical tool for obtaining other good results. For instance, we can use it to find a nice and simple formula for  $\left(\frac{2}{p}\right)$ .

**Theorem.** If  $p$  is an odd prime, then  $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$ .

**Proof.** For  $a = 2$ ,  $S = \{2, 4, \dots, p-1\}$ , so  $k = |\{n : \frac{p-1}{2} < 2n \leq p-1\}|$ . The value of  $k$  depends on whether  $\frac{p-1}{2}$  is odd or even. If  $\frac{p-1}{2}$  is odd, then  $k = \frac{1}{2}((p-1) - (\frac{p-1}{2} - 1)) = \frac{p+1}{4}$ , and so by Gauss' Lemma,

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p+1}{4}} = \left((-1)^{\frac{p-1}{2}}\right)^{\frac{p+1}{4}} = (-1)^{\frac{p-1}{2} \cdot \frac{p+1}{4}} = (-1)^{\frac{p^2-1}{8}}.$$

If  $\frac{p-1}{2}$  is even, then  $k = \frac{1}{2}((p-1) - \frac{p-1}{2}) = \frac{p-1}{4}$ , and by Gauss' Lemma,

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p-1}{4}} = \left((-1)^{\frac{p-1}{2}}\right)^{\frac{p-1}{4}} = (-1)^{\frac{p-1}{2} \cdot \frac{p-1}{4}} = (-1)^{\frac{p^2-1}{8}}.$$

**Corollary.** For odd primes  $p$ ,  $\left(\frac{2}{p}\right) = \begin{cases} 1 & \text{if } p \equiv \pm 1 \pmod{8}; \\ -1 & \text{if } p \equiv \pm 3 \pmod{8}. \end{cases}$

**Example.** Since  $1001 \equiv 1 \pmod{8}$ ,  $\left(\frac{2}{1001}\right) = 1$ .

**Example.** Let  $q$  be an odd integer for which  $p = 12q + 1$  is prime.

- Then  $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = (-1)^{\frac{144q^2+24q}{8}} = (-1)^{q(18q+3)} = -1$ .
- Hence,  $2^{6q} = 2^{\frac{p-1}{2}} \equiv \left(\frac{2}{p}\right) \equiv -1 \pmod{p}$ .
- **Exercise:** Use this to find a prime factor  $p$  of  $2^{78} + 1$ .

**Quadratic Reciprocity.** Gauss regarded the Law of Quadratic Reciprocity (given on the next page) as one of his greatest achievements, and many proofs of it have been given, eight of which were found by him. For the proof given here, we need the following technical lemma.\*

**Lemma.** For  $a \in \mathbb{N}$  and primes  $p, q$  with  $p \equiv \pm q \pmod{4a}$ ,  $\left(\frac{a}{p}\right) = \left(\frac{a}{q}\right)$ .

**Proof.** Define  $k_p$  and  $k_q$  for  $p$  and  $q$ , respectively, as in Gauss' Lemma. Set  $b = \lfloor \frac{a}{2} \rfloor$ . Since  $0 < n < \frac{p}{2}$  if and only if  $0 < na < \frac{ap}{2}$ , it follows that

$$\begin{aligned} k_p &= \left| \left\{1, \dots, \frac{p-1}{2}\right\} \cap \bigcup_{j=1}^b \left\{n \in \mathbb{Z} : \frac{(2j-1)p}{2} < na < jp\right\} \right| \\ &= \left| \bigcup_{j=1}^b \left\{n \in \mathbb{Z} : \frac{(2j-1)p}{2a} < n < \frac{jp}{a}\right\} \right| \\ &= \sum_{j=1}^b \left| \left\{n \in \mathbb{Z} : \frac{(2j-1)p}{2a} < n < \frac{jp}{a}\right\} \right|. \end{aligned}$$

Similarly,

$$k_q = \sum_{j=1}^b \left| \left\{n \in \mathbb{Z} : \frac{(2j-1)q}{2a} < n < \frac{jq}{a}\right\} \right|.$$

---

\* 224 proofs of Gauss' Law of Quadratic Reciprocity are listed here: <http://www.rzuser.uni-heidelberg.de/~hb3/rchrono.html>

Suppose that  $p \equiv q \pmod{4a}$  and write  $p = q + 4am$  with  $m \in \mathbb{Z}$ . Then

$$\begin{aligned} k_p &= \sum_{j=1}^b \left| \left\{n \in \mathbb{Z} : \frac{(2j-1)q}{2a} + (4j-2)m < n < \frac{jq}{a} + 4jm\right\} \right| \\ &= \sum_{j=1}^b \left| \left\{n \in \mathbb{Z} : \frac{(2j-1)q}{2a} < n < \frac{jq}{a} + 2m\right\} \right| \\ &= \sum_{j=1}^b \left( \left| \left\{n \in \mathbb{Z} : \frac{(2j-1)q}{2a} < n < \frac{jq}{a}\right\} \right| + 2m \right) \\ &= k_q + 2bm. \end{aligned}$$

Hence by Gauss' Lemma,

$$\left(\frac{a}{p}\right) = (-1)^{k_p} = (-1)^{k_q+2bm} = (-1)^{k_q} = \left(\frac{a}{q}\right).$$

Suppose that  $p \equiv -q \pmod{4a}$  and write  $p = 4am - q$  for  $m \in \mathbb{Z}$ . Then

$$\begin{aligned} k_p &= \sum_{j=1}^b \left| \left\{n \in \mathbb{Z} : -\frac{(2j-1)q}{2a} + (4j-2)m < n < -\frac{jq}{a} + 4jm\right\} \right| \\ &= \sum_{j=1}^b \left| \left\{n \in \mathbb{Z} : -\frac{(2j-1)q}{2a} < n < -\frac{jq}{a} + 2m\right\} \right| \\ &= \sum_{j=1}^b \left| \left\{n \in \mathbb{Z} : \frac{jq}{a} < n < \frac{(2j-1)q}{2a} + 2m\right\} \right|. \end{aligned}$$

If  $m \geq 0$ , then

$$k_p + k_q = \sum_{j=1}^b \left| \left\{n \in \mathbb{Z} : \frac{(2j-1)q}{2a} < n < \frac{(2j-1)q}{2a} + 2m\right\} \right| = 2bm;$$

similarly,  $k_p + k_q = -2bm$  if  $m \leq 0$ . Hence,

$$\left(\frac{a}{p}\right) = (-1)^{k_p} = (-1)^{k_q \pm 2bm} = (-1)^{k_q} = \left(\frac{a}{q}\right).$$

**The Law of Quadratic Reciprocity.** If  $p$  and  $q$  are odd primes, then

$$\begin{aligned} \left(\frac{p}{q}\right) &= \left(\frac{q}{p}\right) & \text{if } p \equiv 1 \pmod{4} \text{ or } q \equiv 1 \pmod{4}; \\ \left(\frac{p}{q}\right) &= -\left(\frac{q}{p}\right) & \text{if } p \equiv 3 \pmod{4} \text{ and } q \equiv 3 \pmod{4}. \end{aligned}$$

Equivalently when  $p \neq q$ ,

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

**Proof.** Since  $p, q$  are odd,  $p \equiv \pm q \pmod{4}$ . Without loss of generality, assume that  $q > p$ . Then  $q = 4a \pm p$  with  $a \in \mathbb{N}$ , and

$$\left(\frac{q}{p}\right) = \left(\frac{4a \pm p}{p}\right) = \left(\frac{4a}{p}\right) = \left(\frac{a}{p}\right).$$

Furthermore,  $p \equiv \pm q \pmod{4a}$ , so by the last lemma,  $\left(\frac{a}{q}\right) = \left(\frac{a}{p}\right) = \left(\frac{q}{p}\right)$ . If  $q = 4a - p$ , then either  $p \equiv 1 \pmod{4}$  or  $q \equiv 1 \pmod{4}$ ; also,

$$\left(\frac{p}{q}\right) = \left(\frac{4a - q}{q}\right) = \left(\frac{4a}{q}\right) = \left(\frac{a}{q}\right) = \left(\frac{q}{p}\right).$$

Now suppose that  $q = 4a + p$ . Then

$$\left(\frac{p}{q}\right) = \left(\frac{q - 4a}{q}\right) = \left(\frac{-4a}{q}\right) = \left(\frac{-a}{q}\right) = \left(\frac{-1}{q}\right)\left(\frac{a}{q}\right) = \left(\frac{-1}{q}\right)\left(\frac{q}{p}\right).$$

If  $p \equiv q \equiv 1 \pmod{4}$ , then  $\left(\frac{-1}{q}\right) = 1$ , and thus  $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$ ; otherwise,  $p \equiv q \equiv 3 \pmod{4}$ , in which case  $\left(\frac{-1}{q}\right) = -1$ , and so  $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$ .

**Comment.** The results of this chapter form a very efficient set of tools to calculate  $\left(\frac{a}{p}\right)$  when  $p$  is prime. By reducing  $a$  modulo  $p$ , by the identity  $\left(\frac{bc}{p}\right) = \left(\frac{b}{p}\right)\left(\frac{c}{p}\right)$ , and by the Law of Quadratic Reciprocity, we may reduce  $\left(\frac{a}{p}\right)$  to a product of Legendre symbols of the form  $\left(\frac{a'}{p'}\right)$  where  $a' = 2, \pm 1$  or  $a'$  sufficiently small, and use the formulas for  $\left(\frac{a'}{p'}\right)$  or Euler's Criterion in these cases. Further reductions are had by noting that if  $p_i^{\alpha_i}$  are the prime power factors of  $a$ , then  $\left(\frac{p_i^2}{p}\right) = 1$ , and so

$$\left(\frac{a}{p}\right) = \prod_{i: 2 \nmid \alpha_i} \left(\frac{p_i}{p}\right).$$

**Example.** Calculate  $\left(\frac{360}{617}\right)$  and  $\left(\frac{323}{337}\right)$ :

- $\left(\frac{360}{617}\right) = \left(\frac{2}{617}\right)\left(\frac{5}{617}\right) = \left(\frac{617}{5}\right) = \left(\frac{2}{5}\right) = -1$ .
- $\left(\frac{323}{337}\right) = \left(\frac{-14}{337}\right) = \left(\frac{-1}{337}\right)\left(\frac{2}{337}\right)\left(\frac{7}{337}\right) = \left(\frac{7}{337}\right) = \left(\frac{337}{7}\right) = \left(\frac{1}{7}\right) = 1$ .

**Problem.** Does the equation  $x^2 + 6x - 5 \equiv 0 \pmod{127}$  have a solution?

**Solution.** Re-write the equation as  $(x + 3)^2 \equiv 14 \pmod{127}$ . Since

$$\left(\frac{14}{127}\right) = \left(\frac{14}{127}\right) = \left(\frac{2}{127}\right)\left(\frac{7}{127}\right) = \left(\frac{7}{127}\right) = -\left(\frac{127}{7}\right) = -\left(\frac{1}{7}\right) = -1,$$

the equation has no solution.

**Problem.** Is the diophantine equation  $x^2 - 43y^2 = 73$  solvable?

**Solution.** If a solution to this equation exists, then the congruence equation  $x^2 \equiv 73 \pmod{43}$  also has a solution. However,

$$\left(\frac{73}{43}\right) = \left(\frac{30}{43}\right) = \left(\frac{2}{43}\right)\left(\frac{3}{43}\right)\left(\frac{5}{43}\right) = \left(\frac{43}{3}\right)\left(\frac{43}{5}\right) = \left(\frac{1}{3}\right)\left(\frac{3}{5}\right) = -1,$$

so this is not true, and  $x^2 - 43y^2 = 73$  is not solvable.

To conclude this chapter, let us list  $\left(\frac{a}{p}\right)$  for small integers  $a$ :

**Theorem.** If  $p$  is an odd prime, then

- $\left(\frac{0}{p}\right) = 0$ ;  $\left(\frac{1}{p}\right) = 1$ ;
- $\left(\frac{-1}{p}\right) = 1$  if and only if  $p \equiv 1 \pmod{4}$ ;
- $\left(\frac{2}{p}\right) = 1$  if and only if  $p \equiv \pm 1 \pmod{8}$ ;
- $\left(\frac{3}{p}\right) = 1$  if and only if  $p \equiv \pm 1 \pmod{12}$ .

**Proof.** The theorem's first three parts have been shown, so consider  $\left(\frac{3}{p}\right)$ . By Euler's Criterion and the Law of Quadratic Reciprocity,

$$\left(\frac{3}{p}\right) \equiv p(-1)^{\frac{p-1}{2}} \pmod{3}.$$

It is now easy to see that of the six cases  $p \equiv 1, 3, 5, 7, 9, 11 \pmod{12}$ , the identity  $\left(\frac{3}{p}\right) = 1$  holds precisely when  $p \equiv 1, 11 \pmod{12}$ .

---

---

The University of New South Wales  
School of Mathematics and Statistics

MATH3521

ALGEBRAIC TECHNIQUES IN NUMBER THEORY

---

---

6. THE GAUSSIAN INTEGERS.

This chapter looks at the ring of Gaussian integers, objects that in many respects generalise the usual integers.\* By using certain properties of these Gaussian integers, we will be able to characterise those numbers such as  $26 = 5^2 + 1^2$ , but not as 6 for instance, that can be written as the sum of two integer squares.

**Definition.** The **Gaussian integers** are the complex numbers

$$\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\},$$

where  $i$  is a square root of  $-1$ .

As noted in Chapter 1, the Gaussian integers  $\mathbb{Z}[i]$  form a ring, namely a subring of the field of complex numbers  $\mathbb{C}$ . Indeed, as we shall show,  $\mathbb{Z}[i]$  is nearly a field itself; more precisely, it is a principal ideal domain that allows a generalised Euclidean algorithm.†

The Gaussian integers may be thought of as the points  $(a, b)$  of an infinite planar grid. As such vectors, each element  $\alpha = a + bi \in \mathbb{Z}[i]$  may be assigned a length or, squaring the length, a norm. More generally, we define the following norm for all complex numbers  $\alpha = a + ib$ :

**Definition.‡** The **norm** of  $\alpha = a + ib \in \mathbb{C}$  is  $N(\alpha) = \alpha\bar{\alpha} = a^2 + b^2$ .

---

\* Not surprisingly, Gaussian integers were introduced by K. F. Gauss, to simplify proofs of his Law of Quadratic Reciprocity from Chapter 5.

† Such integral domains are called **Euclidean domains**.

‡ This definition of a norm  $N(\cdot)$  is mostly used in number theory; in most other mathematical areas, the norm is defined as  $\sqrt{a^2 + b^2}$ .

The norm  $N(\cdot)$  has several fairly obvious but useful properties:

**Theorem.** Let  $\alpha, \beta \in \mathbb{Z}[i]$  and  $c \in \mathbb{Z}$  be given. Then

- $N(\alpha) \geq 0$  and  $N(\alpha) \in \mathbb{Z}$ ;
- $N(\alpha + \beta) \leq N(\alpha) + N(\beta)$ ;
- $N(\alpha\beta) = N(\alpha)N(\beta)$ ;
- $N(\alpha) \leq N(\alpha\beta)$  for  $\beta \neq 0$ ;
- $N(c\alpha) = c^2N(\alpha)$ ,  $N(c) = c^2$ , and  $N(N(\alpha)) = (N(\alpha))^2$ ;
- $N(\alpha) = 1$  if and only if  $\alpha \in \{\pm 1, \pm i\}$ ;
- $N(\alpha) = 0$  if and only if  $\alpha = 0$ .

**Proof.** Exercise.

Using that  $\mathbb{Z}[i]$  is a subring of the field  $\mathbb{C}$  and contains the identity 1, or by using the theorem above, we see that

**Corollary.**  $\mathbb{Z}[i]$  is an integral domain.

**Theorem.** The following are equivalent for  $\alpha \in \mathbb{Z}[i]$ :

- $\alpha$  is a unit;
- $N(\alpha) = 1$ ;
- $\alpha \in \{\pm 1, \pm i\}$ .

**Proof.** If  $\alpha$  is a unit, then  $\alpha\beta = 1$  for some element  $\beta \in \mathbb{Z}[i]$ ; and  $N(\alpha)N(\beta) = N(1) = 1$ , so  $N(\alpha) = 1$ . If  $N(\alpha) = 1$ , then  $\alpha \in \{\pm 1, \pm i\}$ . Finally, each of the elements  $\alpha \in \{\pm 1, \pm i\}$  is a unit in  $\mathbb{Z}[i]$ .

The division algorithm for integers (page 18) works due to the total ordering of the integers; that is, of any two integers, one is biggest. Such orders exist for the Gaussian integers but none is very natural. However, the partial order defined by the norm permits the following division algorithm for Gaussian integers.

**Theorem.** *Division algorithm for Gaussian integers.* For all  $\alpha, \beta \neq 0$ , elements  $q, r \in \mathbb{Z}[i]$  exist for which  $\alpha = q\beta + r$  and  $0 \leq N(r) < N(\beta)$ .

**Proof.** Write  $\frac{\alpha}{\beta} = x + yi$  where  $x, y \in \mathbb{Q}$ , choose integers  $s, t \in \mathbb{Z}$  so that  $|x - s| \leq \frac{1}{2}$  and  $|y - t| \leq \frac{1}{2}$ , and set  $q = s + ti$  and  $r = \alpha - q\beta$ . Then  $\alpha = q\beta + r$  and, since  $N(\frac{\alpha}{\beta} - q) = ((x - s)^2 + (y - t)^2) \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$ ,

$$N(r) = N(\alpha - q\beta) = N((\frac{\alpha}{\beta} - q)\beta) = N(\frac{\alpha}{\beta} - q)N(\beta) < N(\beta).$$

**Example.** Consider  $\alpha = 4 + 4i$  and  $\beta = 2 - i$ .

By the division algorithm, there are elements  $q, r \in \mathbb{Z}[i]$  that satisfy  $\alpha = q\beta + r$  and  $0 \leq N(r) < N(\beta) = 5$ . To find such elements, use the method given in the proof of the theorem. First, calculate the fraction  $\frac{\alpha}{\beta}$ :

$$\frac{\alpha}{\beta} = \frac{4 + 4i}{2 - i} = \frac{4 + 4i}{2 - i} \cdot \frac{2 + i}{2 + i} = \frac{4 + 12i}{5} = \frac{4}{5} + \frac{12}{5}i.$$

Next, let  $s$  and  $t$  be nearest integers to  $\frac{4}{5}$  and  $\frac{12}{5}$  of  $\frac{\alpha}{\beta}$ , respectively:  $s = 1$  and  $t = 2$ . Set  $q = 1 + 2i$  and let  $r$  be the remainder

$$r = \alpha - q\beta = (4 + 4i) - (1 + 2i)(2 - i) = (4 + 4i) - (4 + 3i) = i.$$

Then  $4 + 4i = \alpha = q\beta + r = (1 + 2i)(2 - i) + i$  and  $N(i) = 1 < 5 = N(2 - i)$ .

**Comment.** Unlike the division algorithm for integers, the division algorithm for Gaussian integers does not give unique results. For instance in the above example, we could have chosen  $(q, r)$  to be any of the pairs  $(2i, 2)$ ,  $(1 + 2i, i)$ , and  $(1 + 3i, -1 - i)$ . In general, we must choose the element  $q = s + ti \in \mathbb{Z}[i]$  within norm distance 1 from  $\frac{\alpha}{\beta}$ , so there are always two, three, or four valid pairs  $(q, r)$  to choose from.

Greatest common divisors may be defined for Gaussian integers:

**Definition.** A **greatest common divisor** of Gaussian integers  $\alpha$  and  $\beta$  is an element  $\gamma \in \mathbb{Z}[i]$  that is a factor of both  $\alpha$  and  $\beta$  and is a multiple of every such common factor. Let  $\mathbf{gcd}(\alpha, \beta)$  be the set of all greatest common divisors of  $\alpha$  and  $\beta$ .

For integers  $a, b \in \mathbb{Z}$ , the usual greatest common divisor  $\gcd(a, b)$  and its negative  $-\gcd(a, b)$  are both greatest common divisors of  $a$  and  $b$  in the sense given in the definition above; for convenience,  $\gcd(a, b)$  is chosen to be the positive of these two. For Gaussian integers  $\alpha$  and  $\beta$ , there is no natural way to choose a unique representative for the greatest common divisors  $\mathbf{gcd}(\alpha, \beta)$ . However, the Gaussian greatest common divisors always exist and are unique up to multiplication by units. This follows from the Euclidean algorithm modified for Gaussian integers:

**Theorem.** *The Euclidean algorithm for Gaussian integers.* For  $\alpha, \beta \neq 0$ , write  $r_0 = \beta$  and use the division algorithm to find  $q_1, r_1 \in \mathbb{Z}[i]$  for which

$$\alpha = q_1\beta + r_1 \quad \text{with} \quad 0 \leq N(r_1) < N(\beta).$$

If  $r_1 \neq 0$ , then write similarly

$$\beta = q_2r_1 + r_2 \quad \text{with} \quad 0 \leq N(r_2) < N(r_1).$$

Continuing like this while  $r_k \neq 0$ , we find  $q_{k+1}, r_{k+1} \in \mathbb{Z}[i]$  so that

$$r_{k-1} = q_{k+1}r_k + r_{k+1}, \quad \text{with} \quad 0 \leq N(r_{k+1}) < N(r_k).$$

Then  $r_{n+1} = 0$  for some  $n \geq 0$ , and  $\mathbf{gcd}(\alpha, \beta) = \{\pm r_n, \pm r_n i\}$ .

**Proof.** Each of the numbers  $N(r_0) > N(r_1) > N(r_2) > \dots$  is a positive integer, so  $r_{n+1} = 0$  for some  $n \geq 0$ , and the process will then terminate. If  $\gamma \in \mathbf{gcd}(r_n, 0)$ , then  $\gamma = mr_n$  and  $r_n = m'\gamma$  for elements  $m, m' \in \mathbb{Z}[i]$ . Then  $\gamma = (mm')\gamma$ , so  $mm'$  and thus  $m$  is a unit, and  $\gamma \in \{\pm r_n, \pm r_n i\}$ . Conversely,  $\{\pm r_n, \pm r_n i\} \subseteq \mathbf{gcd}(r_n, 0)$ , so  $\mathbf{gcd}(r_n, 0) = \{\pm r_n, \pm r_n i\}$ . Thus, since  $\mathbf{gcd}(\alpha, \beta) = \mathbf{gcd}(\beta, r_1) = \mathbf{gcd}(r_1, r_2) = \dots = \mathbf{gcd}(r_n, r_{n+1})$ ,

$$\mathbf{gcd}(\alpha, \beta) = \mathbf{gcd}(r_n, 0) = \{\pm r_n, \pm r_n i\}.$$

**Example.** Find  $\mathbf{gcd}(4 + 4i, 2 - i)$  by the Euclidean algorithm:

$$\begin{aligned} 4 + 4i &= (1 + 2i) \times (2 - i) + i \\ 2 - i &= (-1 - 2i) \times i + 0. \end{aligned}$$

Then  $i$  is a greatest common divisor of  $4 + 4i$  and  $2 - i$ , so

$$\mathbf{gcd}(4 + 4i, 2 - i) = \{\pm 1, \pm i\}.$$

**Comment.** The Euclidean algorithm above is a bit remarkable since it implies that the repeated use of the division algorithm must lead to one of the four greatest common divisors of  $\alpha$  and  $\beta$ , regardless of which pairs  $(q_i, r_i)$  are chosen along the way.

As in Chapter 1, an ideal  $I \subseteq \mathbb{Z}[i]$  is principal if it is generated by a single element. Using the Euclidean algorithm, we can show that the Gaussian integers form a principal ideal domain:

**Theorem.** Every ideal in the Gaussian integers is principal.

**Proof.** The ideal  $\{0\}$  is principal so consider an ideal  $I \neq \{0\}$  in  $\mathbb{Z}[i]$ . Choose an element  $\gamma \in I$  with smallest positive norm  $N(\gamma) > 0$ , and consider any  $\alpha \in I$ . By the division algorithm, elements  $q, r \in \mathbb{Z}[i]$  exist for which  $\alpha = q\gamma + r$  with  $N(r) < N(\gamma)$ . Then  $r = \alpha - q\gamma \in I$ , so since  $N(\alpha)$  is the smallest positive norm, we have  $N(r) = 0$  and so  $r = 0$ . Hence,  $\alpha = q\gamma \in \langle \gamma \rangle$ ; therefore,  $I \subseteq \langle \gamma \rangle$ . Since  $\alpha\gamma \in I$  for all  $\alpha \in \mathbb{Z}[i]$ , we also see that  $\langle \gamma \rangle \subseteq I$  and thus that  $I = \langle \gamma \rangle$ .

**Theorem.** For any elements  $\alpha, \beta \in \mathbb{Z}[i]$ , suppose that  $\gamma \in \gcd(\alpha, \beta)$ . Then  $\langle \alpha, \beta \rangle = \langle \gamma \rangle$  and so  $\gamma = m\alpha + n\beta$  for some  $m, n \in \mathbb{Z}[i]$ .

**Proof.** The ideal  $\langle \alpha, \beta \rangle$  is principal and is thus  $\langle \delta \rangle$  for some  $\delta \in \mathbb{Z}[i]$ . Then  $\delta \mid \alpha$  and  $\delta \mid \beta$ , so  $\delta \mid \gamma$ . Hence,  $\gamma \in \langle \delta \rangle = \langle \alpha, \beta \rangle$ , so  $\langle \gamma \rangle \subseteq \langle \alpha, \beta \rangle$ , and  $\gamma = m\alpha + n\beta$  for some elements  $m, n \in \mathbb{Z}[i]$ . Since  $\gamma \mid \alpha$  and  $\gamma \mid \beta$ , we also have that  $\langle \alpha, \beta \rangle \subseteq \langle \gamma \rangle$ , so  $\langle \alpha, \beta \rangle = \langle \gamma \rangle$ .

**Example.** Find a generator for the ideal  $\langle 2 - 11i, 4 - 7i \rangle$ .

By the theorem above, any greatest common divisor  $\gamma$  of  $2 - 11i$  and  $4 - 7i$  will do, so let us use the Euclidian algorithm to find such a divisor:

$$\begin{aligned} 2 - 11i &= 1 \times (4 - 7i) + (-2 - 4i); \\ 4 - 7i &= (1 + 2i) \times (-2 - 4i) + (-2 + i); \\ -2 - 4i &= 2i \times (-2 + i) + 0. \end{aligned}$$

Therefore,  $-2 + i$  is a greatest common divisor of  $2 - 11i$  and  $4 - 7i$ , and

$$\langle 2 - 11i, 4 - 7i \rangle = \langle -2 + i \rangle.$$

By reversing the steps of the Euclidean algorithm above, we can also find elements  $m, n \in \mathbb{Z}[i]$  so that  $-2 + i = m(2 - 11i) + n(4 - 7i)$ :

$$\begin{aligned} -2 + i &= 4 - 7i - (1 + 2i) \times (-2 - 4i) \\ &= 4 - 7i - (1 + 2i) \times ((2 - 11i) - (4 - 7i)) \\ &= (-1 - 2i) \times (2 - 11i) + (2 + 2i) \times (4 - 7i). \end{aligned}$$

**Gaussian primes and irreducibles.** As in Chapter 1, a non-zero, non-unit Gaussian integer  $\pi \in \mathbb{Z}[i]$  is **irreducible** if and only if

$$\alpha \text{ or } \beta \text{ is a unit} \quad \text{whenever} \quad \pi = \alpha\beta \text{ with } \alpha, \beta \in \mathbb{Z}[i]. \quad (\text{I})$$

Similarly,  $\pi \in \mathbb{Z}[i]$  is **prime** if and only if

$$\pi \mid \alpha \text{ or } \pi \mid \beta \quad \text{whenever} \quad \pi \mid \alpha\beta \text{ with } \alpha, \beta \in \mathbb{Z}[i]. \quad (\text{P})$$

**Definition.** The primes of  $\mathbb{Z}[i]$  are called **Gaussian primes**.

**Example.**  $7$  and  $1 + i$  are Gaussian primes but  $2 = (1 + i)(1 - i)$  is not. We will prove these claims later below.

By Chapter 1, the prime integers are exactly the irreducible integers. This is also true for Gaussian primes. To prove this, we can use the following general result about integral domains mentioned in Chapter 1:

**Lemma.** Every prime element of an integral domain is irreducible.

**Proof.** For a prime element  $p$  of an integral domain  $D$  with identity  $1$ , suppose that  $p = ab$  with  $a, b \in D$ . Then  $p \mid ab$ , so either  $p \mid a$  or  $p \mid b$ . If  $p \mid a$ , then  $a = pu$  for some  $u \in D$ , so  $p = ab = pub$ , and  $p(1 - ub) = 0$ . Now,  $p \neq 0$  and  $D$  has no zero-divisor, so  $ub = 1$ , and  $b$  is a unit. Similarly,  $a$  is a unit if  $p \mid b$ .

**Theorem.** A Gaussian integer is prime if and only if it is irreducible.

**Proof.** By the lemma above, we only need to show that each irreducible element  $\pi \in \mathbb{Z}[i]$  is prime, so suppose that  $\pi \mid \alpha\beta$  for some  $\alpha, \beta \in \mathbb{Z}[i]$ . Suppose that  $\pi \nmid \alpha$ . Since  $\pi$  is irreducible,  $\alpha$  and  $\pi$  only have units as common factors, so  $1 \in \gcd(\alpha, \pi)$ , and  $1 = x\alpha + y\pi$  for some  $x, y \in \mathbb{Z}[i]$ . Therefore,  $\beta = x\alpha\beta + y\pi\beta$ , so  $\pi \mid \beta$ .

The above theorem implies that Gaussian integers factor uniquely,\* just like normal integers but unlike  $\mathbb{Z}[\sqrt{-10}]$ , say:

**Theorem.** Each Gaussian integer factors uniquely into Gaussian primes up to order and multiplication by units.

**Example.**  $5$  factors uniquely into Gaussian primes as  $5 = (2 + i)(2 - i)$ , but also as  $5 = (2 - i)(2 + i)$  or  $5 = (2i - 1)(-2i + 1)$ , for instance.

\* Such integral domains are called **unique factorisation domains**.

**Theorem.** An integer prime  $p$  is a Gaussian prime iff  $p \equiv 3 \pmod{4}$ .

**Proof.** Suppose that  $p$  is a Gaussian prime. Now,  $pb \neq \pm 1$  if  $b \in \mathbb{Z}$ , so  $x \pm i \neq pa \pm pbi = p(a \pm bi)$  for all  $a, b, x \in \mathbb{Z}$ ; therefore,  $p \nmid x \pm i$ . Since  $p$  is a Gaussian prime, we see that  $p \nmid (x - i)(x + i) = x^2 + 1$ , so  $x^2 \equiv -1 \pmod{p}$  has no solution, and  $p \equiv 3 \pmod{4}$ . Conversely, suppose that  $p \equiv 3 \pmod{4}$ . If  $p = \beta\gamma$ , then  $p^2 = N(p) = N(\beta)N(\gamma)$ . Since  $p$  is prime, either  $N(\beta) = N(\gamma) = p$  or  $\{N(\beta), N(\gamma)\} = \{1, p^2\}$ . Write  $\beta = c + di$  for some  $c, d \in \mathbb{Z}$ . Since  $x^2 \equiv 0, 1 \pmod{4}$  for all  $x \in \mathbb{Z}$ , we see that  $N(\beta) = c^2 + d^2 \equiv 0, 1, 2 \not\equiv 3 \pmod{4}$ . Therefore,  $N(\beta) \neq p$ , so either  $N(\beta) = 1$  or  $N(\gamma) = 1$ , and  $p$  is a Gaussian prime.

**Theorem.** A Gaussian integer  $\alpha \in \mathbb{Z}[i]$  is prime if and only if either

- $N(\alpha)$  is a prime integer;
- or •  $\alpha = \epsilon p$  for a unit  $\epsilon \in \mathbb{Z}[i]$  and a prime  $p \in \mathbb{Z}$  with  $p \equiv 3 \pmod{4}$ .

**Proof.** Let  $\alpha = a + bi \neq 0$  be a Gaussian integer. If  $a = 0$  or  $b = 0$ ; then  $\alpha = \epsilon p$  for a unit  $\epsilon$  and some  $p \in \mathbb{Z}$ . Since  $\alpha$  is a Gaussian integer if and only if  $p$  is, we may assume that  $\alpha = p$ . If  $p$  is a Gaussian prime, then  $p$  is also an integer prime, so by the theorem above,  $\alpha = p$  is prime if and only if  $p$  is a prime integer and  $p \equiv 3 \pmod{4}$ .

Consider now the case in which  $a, b \neq 0$ . Suppose that  $\alpha$  is prime and write  $N(\alpha) = \alpha\bar{\alpha} = cd$  for  $c, d \in \mathbb{Z}$ . Then  $\alpha \mid cd$ , so  $\alpha \mid c$  or  $\alpha \mid d$ . Suppose the former; then  $c = \alpha\beta$  with  $\beta \in \mathbb{Z}[i]$ , and  $\alpha\bar{\alpha} = \alpha\beta d$ , so  $\bar{\alpha} = \beta d$ . Thus,  $d \mid \bar{\alpha} = a - bi$ , so  $d \mid a$  and  $d \mid b$ , and  $d \mid a + bi = \alpha$ . Since  $\alpha$  is prime,  $d = \pm 1$ . Similarly, if  $\alpha \mid d$ , then  $c = \pm 1$ , so  $N(\alpha)$  is a prime integer. Conversely, if  $N(\alpha)$  is a prime integer and  $\alpha = \beta\gamma$  with  $\beta, \gamma \in \mathbb{Z}[i]$ , then since  $N(\alpha) = N(\beta)N(\gamma)$ , either  $N(\beta) = 1$  or  $N(\gamma) = 1$ , so either  $\beta$  or  $\gamma$  is a unit, and  $\alpha$  is a Gaussian prime.

**Example.**  $(1 + i)$  is a Gaussian prime since  $N(1 + i) = 2$  is a prime, and  $7$  is a Gaussian prime since  $7$  is an integer prime and  $7 \equiv 3 \pmod{4}$ . However,  $2 = (1 + i)(1 - i)$  is not irreducible and is thus not a prime.

**Example.** Factorise  $15$  and  $24 - 3i$ :

- $5 = 3 \cdot 5 = 3(2 + i)(2 - i)$ ; by the above theorem, all terms are prime.
- Since  $N(24 - 3i) = 585 = 3^2 \cdot 5 \cdot 13$ , we see that  $24 - 3i = \epsilon\beta\gamma$  where  $\epsilon$  is a unit,  $\beta$  is  $2 - i$  or  $2 + i$ , and  $\gamma$  is  $3 - 2i$  or  $3 + 2i$ . By testing each of these possibilities, we find that  $24 - 3i = 3(2 + i)(3 - 2i)$ .

**Sums of two squares.** An integer is the sum of two integer squares if and only if it is the norm of some Gaussian integer. Therefore, the identity  $N(\alpha\beta) = N(\alpha)N(\beta)$  gives us the following useful result.

**Theorem.** If  $m$  and  $n$  are both the sum of two squares, then so is  $mn$ .

**Example.** The integers  $5$  and  $13$  are both sums of two squares:

$$5 = 2^2 + 1^2 = N(2 + i) \quad \text{and} \quad 13 = 3^2 + 2^2 = N(3 + 2i).$$

Therefore,  $65$  is also the sum of two squares:

$$65 = N(2 + i)N(3 + 2i) = N((2 + i)(3 + 2i)) = N(4 + 7i) = 4^2 + 7^2.$$

Since  $5 = N(2 - i)$ ,  $N(2 - i)N(3 + 2i) = N(8 + i)$ , and  $65 = 8^2 + 1^2$ . These are the only ways in which to write  $65$  as a sum of two squares.

**Theorem.** For an odd prime  $p$ , the following statements are equivalent:

- $p$  is the sum of two integer squares;
- $p$  is not a Gaussian prime;
- $p \equiv 1 \pmod{4}$ .

**Proof.** By the theorem at the top of page 141, we only need to show the equivalence of the first two statements. If  $p$  is the sum of two squares, say  $p = a^2 + b^2$ , then  $p = (a - ib)(a + ib)$ , so  $p$  is not a Gaussian prime. Conversely, if  $p$  is not a Gaussian prime, then  $p = \alpha\beta$  for some non-unit elements  $\alpha = u + vi, \beta = x + yi \in \mathbb{Z}[i]$ , where  $x, y, u, v$  are non-zero. Then  $N(\alpha)N(\beta) = N(p)$ , so  $p^2 = (x^2 + y^2)(u^2 + v^2)$ , and  $p = x^2 + y^2$ .

**Example.** Since  $29 \equiv 1 \pmod{4}$  and  $19 \equiv 3 \pmod{4}$ , it follows from the above theorem that  $29$  is the sum of two squares but  $19$  is not.

**Example.** For each odd prime  $p$ , the above theorem implies that the congruence equation  $x^2 \equiv -1 \pmod{p}$  has a solution if and only if  $p$  is the sum of two squares.

The following theorem completely characterises the numbers  $n$  that can be expressed as the sum of two integer squares.

**Theorem.** A positive integer  $n$  is the sum of two squares if and only if  $p \equiv 1 \pmod{4}$  for all odd primes  $p$  that divide  $n$  an odd number of times.

**Proof.** Suppose that  $n$  is the sum of two integer squares, say  $n = s^2 + t^2$ , and let  $p$  be any odd prime that divides  $n$  an odd number of times. Let  $b = \gcd(s, t)$  and write  $n = b^2(x^2 + y^2)$  where  $x = \frac{s}{b}$  and  $y = \frac{t}{b}$ . Note that  $x^2 + y^2 \equiv 0 \pmod{p}$  since  $p$  would otherwise divide  $n$  an even number of times. If  $p \mid y$ , then  $p \mid x$ , but then  $p \mid \gcd(x, y) = 1$ , a contradiction, so  $y \not\equiv 0 \pmod{p}$ . Therefore,  $(xy^{-1})^2 \equiv -1 \pmod{p}$ , so  $\left(\frac{-1}{p}\right) = 1$ , and  $p \equiv 1 \pmod{4}$ .

Conversely, if  $p \equiv 1 \pmod{4}$  for all odd primes  $p$  that divide  $n$  an odd number of times, then by the last theorem, each such prime  $p$  is the sum of two squares. All other prime factors of  $n$  are either 2 or divide  $n$  an even number of times and are thus also sums of two squares. Therefore, the product of these, namely  $n$ , is also the sum of two squares.

**Example.** Consider 792 and 29250.

- Since  $792 = 2^3 \cdot 3^2 \cdot 11$  and  $11 \equiv 3 \pmod{4}$ ,  
792 is not the sum of two squares.
- Since  $29250 = 2 \cdot 3^2 \cdot 5^3 \cdot 13$  and  $5 \equiv 13 \equiv 1 \pmod{4}$ ,  
29250 is the sum of two squares; for instance,

$$\begin{aligned} 29250 &= 2 \cdot 3^2 \cdot 5^3 \cdot 13 \\ &= N(1+i)N(3)N(5)N(1+2i)N(2+3i) \\ &= N((1+i) \cdot 3 \cdot 5 \cdot (1+2i)(2+3i)) \\ &= N(-165 + 45i) \\ &= 165^2 + 45^2, \end{aligned}$$

and

$$\begin{aligned} 29250 &= 2 \cdot 3^2 \cdot 5^3 \cdot 13 \\ &= N(1+i)N(3)N(5)N(1-2i)N(2+3i) \\ &= N((1+i) \cdot 3 \cdot 5 \cdot (1-2i)(2+3i)) \\ &= N(135 + 105i) \\ &= 135^2 + 105^2. \end{aligned}$$

**More on sums of squares.** The last theorem characterised all sums of two squares. The next two theorems characterise (without proof) all sums of three and four squares, respectively.

**Theorem.** A positive integer  $n$  is the sum of three integer squares if and only if  $n \neq 4^k(8m+7)$  for all non-negative integers  $k$  and  $m$ .

By the above theorem, most numbers are sums of three squares. In 1770, J. L. Lagrange showed that all numbers are sums of four squares:

**Theorem.** Every positive integer is the sum of four integer squares.

We have considered sums of squares but one could also look at sums of cubes, for instance, or sums of even higher powers of integers. Also in 1770, E. Waring proposed what we now call **Waring's Problem**, namely the problem of determining the least number  $g(n)$  for which each positive number is a sum of  $g(n)$   $n$ th powers of some integers; for instance,  $g(2) = 4$  by the theorems above. Although many good mathematicians have worked on the problem, or just special cases of it, not much progress has been made. Values of  $g(n)$  for small  $n$  have been found, and  $g(n)$  is conjectured to be the integer closest to  $2^n + \left(\frac{3}{2}\right)^n - 2$ .

---

---

The University of New South Wales  
School of Mathematics and Statistics

**MATH3521**

**ALGEBRAIC TECHNIQUES IN NUMBER THEORY**

---

---

**7. ALGEBRAIC NUMBER FIELDS.**

In the last chapter, we determined basic properties of the Gaussian integers, and these properties were used to characterise sums of integers. In the first half of this chapter, we will perform a similar analysis on the sets of polynomials over a field, partly in order to study this important ring, and partly in order to provide tools for a subsequent study of algebraic extension fields. This latter topic is interesting in its own right but mainly serves as the fundament for Chapter 8.

Throughout this chapter, let  $\mathbb{F}$  be a field.

**Definition.** A **polynomial**  $f$  over  $\mathbb{F}$  is a sum of the form ( $a_n \neq 0$ )

$$f = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

where  $a_i \in \mathbb{F}$  are the **coefficients** of  $f(x)$  and  $x$  is a variable.

If  $a_n = 1$ , then  $f$  is **monic**.

**Definition.** Let  $\mathbb{F}[x]$  denote the set of all polynomials  $f$  over  $\mathbb{F}$ .

**Example.** Each  $\mathbb{F}[x]$  contains the monic polynomial  $f = 1 + x + x^2 + x^3$ . In contrast,  $\mathbb{Q}[x]$  but not  $\mathbb{Z}_2[x]$  contains  $g = \frac{1}{2}x^2 - \frac{7}{13}x^3 \in \mathbb{Q}[x]$ .

Under the addition and multiplication of  $\mathbb{F}$ , the polynomials  $\mathbb{F}[x]$  form a ring with zero  $0 \in \mathbb{F}$  and identity  $1 \in \mathbb{F}$ . Since  $\mathbb{F}$  is commutative and has no zero-divisors, we can say somewhat more:

**Theorem.**  $\mathbb{F}[x]$  is an integral domain.

In the same way that we measured the integers  $\mathbb{Z}$  by their order  $<$  and cardinality  $|a|$ , and the Gaussian integers  $\alpha \in \mathbb{Z}[i]$  by norm  $N(\alpha)$ , we will measure the polynomials  $f \in \mathbb{F}[x]$  by their degree:

**Definition.** The **degree** of  $f = a_n x^n + \cdots + a_0$  ( $a_n \neq 0$ ) is  $\deg f = n$ ; the **degree** of  $f = 0$  is  $\deg 0 = -\infty$ .

The degree has the following obvious but useful properties:

**Theorem.** Let  $f, g \in \mathbb{F}[x]$  and non-zero  $c \in \mathbb{F}$  be given. Then

- $\deg f \geq 0$  and  $\deg f \in \mathbb{Z}$  unless  $f = 0$ ;
- $\deg (f + g) \leq \max\{\deg f, \deg g\}$ ;
- $\deg (fg) = \deg f + \deg g$ ;
- $\deg (cf) = \deg f$ .
- $\deg f \leq \deg (fg)$  unless  $g = 0$ .

**Proof.** Exercise.

**Theorem.** The following are equivalent for each non-zero  $f \in \mathbb{F}[x]$ :

- $f$  is a unit;
- $\deg f = 0$ ;
- $f = c$  for some non-zero  $c \in \mathbb{F}$ .

**Proof.** If  $f$  is a unit, then  $fg = 1$  for some unit  $g \in \mathbb{F}[x]$ ; therefore,  $0 \leq \deg f \leq \deg (fg) = \deg 1 = 0$ , and so  $\deg f = 0$ . If  $\deg f = 0$ , then  $f = c$  for some unit  $c \in \mathbb{F}$ . Finally, each unit  $c \in \mathbb{F}$  is a unit in  $\mathbb{F}[x]$ .

As for integers and Gaussian integers, a division algorithm exists for  $\mathbb{F}[x]$ :

**Theorem.** *Division Algorithm for  $\mathbb{F}[x]$ .* For all polynomials  $f, g \in \mathbb{F}[x]$  with  $f \neq 0$  and  $\deg g > 0$ , unique polynomials  $q, r \in \mathbb{F}[x]$  exist so that  $f = qg + r$  and  $\deg r < \deg g$ .

**Proof.** We will apply induction on  $\deg f$  to find such polynomials  $q, r$ . If  $\deg f < \deg g$ , then  $q = 0$  and  $r = f$  satisfy the theorem's conditions, so assume that this is true whenever  $\deg f < n$  for an integer  $n \geq \deg g$ . Suppose that  $f = a_n x^n + \cdots + a_0$  and  $g = b_k x^k + \cdots + b_0$  with  $a_n, b_k \neq 0$ . Since  $\deg (f - a_n b_k^{-1} x^{n-k} g) < \deg f = n$ , the induction assumption implies that  $q_0, r \in \mathbb{F}[x]$  exist such that  $\deg r < \deg g$  and

$$f - a_n b_k^{-1} x^{n-k} g = q_0 g + r.$$

Set  $q = q_0 + a_n b_k^{-1} x^{n-k}$  and note that  $f = qg + r$ ; now use induction.

To show that  $q, r$  are unique, suppose that  $q', r' \in \mathbb{F}[x]$  also satisfy conditions  $f = q'g + r'$  and  $\deg r' < \deg g$ . Then  $0 = (q - q')g + r - r'$ , so

$$\deg (q - q')g = \deg (r' - r) \leq \max\{\deg r', \deg r\} < \deg g,$$

and  $r' - r = q - q' = 0$ ; that is,  $r' = r$  and  $q = q'$ .

Note that the Division Algorithm for  $\mathbb{F}[x]$  gives unique  $q$  and  $r$ , just like the Division Algorithm for integers  $\mathbb{Z}$  but unlike that for the Gaussian integers  $\mathbb{Z}[i]$ , for instance.

**Corollary.** *The Remainder Theorem.* For each  $f \in \mathbb{F}[x]$  and each  $c \in \mathbb{F}$ , a unique polynomial  $q \in \mathbb{F}[x]$  satisfies  $f = q(x - c) + f(c)$ .

**Proof.** By the theorem above, unique polynomials  $q, r \in \mathbb{F}[x]$  exist such that  $f = q(x - c) + r$  and  $\deg r < \deg (x - c) = 1$ . Then  $r$  is a constant, namely  $r = r(c) = f(c) - q(c)(c - c) = f(c)$ .

The Remainder Theorem implies the following useful result:

**Corollary.** If  $f \in \mathbb{F}[x]$  and  $c \in \mathbb{F}$ , then  $f(c) = 0$  if and only if  $(x - c) \mid f$ .

**Definition.** An element  $\alpha \in \mathbb{F}$  is a **root** of  $f \in \mathbb{F}[x]$  if  $f(\alpha) = 0$ .

**Example.** Consider  $f = x^3 + x^2 + x + 1 = (x^2 + 1)(x + 1) \in \mathbb{F}[x]$ .

- $f(-1) = 0$  and  $f(0) = 1$ , so  $-1$  is a root of  $f$  but  $0$  is not;
- if  $\mathbb{F} = \mathbb{Q}$ , then  $f$  has no other roots; however,
- if  $\mathbb{F} = \mathbb{C}$ , then  $f$  has the three roots:  $-1, -i$ , and  $i$ .

From the above corollary, we get the following fundamental result:

**Corollary.** Each polynomial  $f \in \mathbb{F}[x]$  has at most  $\deg f$  roots.

As Chapters 1 and 6 showed for integers and Gaussian integers, the greatest common divisor properties of a ring determine much of the structure of that ring, in this case the ring  $\mathbb{F}[x]$ :

**Definition.**  $d \in \mathbb{F}[x]$  is a **greatest common divisor** of  $f, g \in \mathbb{F}[x]$  if it divides both  $f$  and  $g$  and is a multiple of all other such common factors. Let  $\mathbf{gcd}(f, g)$  be the set of all greatest common divisors of  $f$  and  $g$ .

For the integers, the notation “ $\gcd(a, b)$ ” denotes the positive of the two greatest common divisors of  $a$  and  $b$ . The Euclidean Algorithm for  $\mathbb{F}[x]$  states that the set  $\mathbf{gcd}(f, g)$  consist of all unit multiples of some greatest common divisor of polynomials  $f$  and  $g$ , so we could also replace the set  $\mathbf{gcd}(f, g)$  by a representative of this set. In particular, it would be natural to choose the monic polynomial in this set, that is, the greatest common divisor with leading coefficient  $a_n$  equal to 1. However, as for the Gaussian integers in Chapter 6, it is simpler to view the whole set  $\mathbf{gcd}(f, g)$  rather than a representative thereof.

**Example.** For  $f = x^3$  and  $g = x^4 - x^2$  over  $\mathbb{F}$ , the polynomial  $x^2$  is a greatest common divisor, as is any unit multiple thereof, and  $f, g$  have no other greatest common divisors, so  $\mathbf{gcd}(f, g) = \{cx^2 : c \in \mathbb{F}, c \neq 0\}$ .

As in Chapters 1 and 6, the Division Algorithm may be applied iteratively to pairs of elements to find their greatest common divisors; this is expressed as the Euclidean Algorithm:

**Theorem.** *The Euclidean Algorithm for  $\mathbb{F}[x]$ .* For  $f, g \in \mathbb{F}[x]$ ,  $g \neq 0$ , set  $r_0 = g$  and use the division algorithm to find  $q_1, r_1 \in \mathbb{F}[x]$  for which

$$f = q_1g + r_1 \quad \text{with} \quad \deg r_1 < \deg g.$$

If  $r_1 \neq 0$ , then write similarly

$$g = q_2r_1 + r_2 \quad \text{with} \quad \deg r_2 < \deg r_1.$$

Continuing like this while  $r_k \neq 0$ , find  $q_{k+1}, r_{k+1} \in \mathbb{F}[x]$  so that

$$r_{k-1} = q_{k+1}r_k + r_{k+1} \quad \text{with} \quad \deg r_{k+1} < \deg r_k.$$

Then  $r_{n+1} = 0$  for some  $n \geq 0$ , and  $\mathbf{gcd}(f, g) = \{cr_n : c \in \mathbb{F}, c \neq 0\}$ .

**Proof.** Modify slightly the corresponding proof on page 138.

As for  $\mathbb{Z}$  and  $\mathbb{Z}[i]$ , greatest common divisors of two elements may be written explicitly and uniquely as a linear combination of those elements:

**Theorem.** If  $f, g \in \mathbb{F}[x]$  and  $d \in \gcd(f, g)$ , then  $\langle f, g \rangle = \langle d \rangle$ . Also, unique  $m, n \in \mathbb{F}[x]$  exist so that  $d = mf + ng$ , and these may be found by reversing the steps of the Euclidean algorithm.

**Proof.** Modify very slightly the corresponding proof on page 139.

**Example.** Consider  $f, g \in \mathbb{Q}[x]$  where  $f = 3x^3 + 2x^2 + 1$  and  $g = x^2 - 1$ . Use the Euclidean algorithm to find some  $d \in \gcd(f, g)$ :

$$\begin{aligned} 3x^3 + 2x^2 + 1 &= (3x + 2) \times (x^2 - 1) + (3x + 3); \\ x^2 - 1 &= \frac{1}{3}x \times (3x + 3) + (-x - 1); \\ 3x + 3 &= -3 \times (-x - 1) + 0. \end{aligned}$$

Therefore,  $-(x + 1)$  is a greatest common divisor of  $f$  and  $g$ , so

$$\gcd(f, g) = \{c(x + 1) : c \in \mathbb{Q}, c \neq 0\}.$$

By the above theorem,  $x + 1$  is a generator for the ideal  $\langle f, g \rangle$ :

$$\langle x^3 + 2x^2 + 1, x^2 - 1 \rangle = \langle x + 1 \rangle,$$

and, by reversing the steps of the Euclidean algorithm above, we can also find polynomials  $m, n \in \mathbb{Q}[x]$  so that  $-x - 1 = mf + ng$ :

$$\begin{aligned} -x - 1 &= (x^2 - 1) - \frac{1}{3}x \times (3x - 3) \\ &= (x^2 - 1) - \frac{1}{3}x \times ((3x^3 + 2x^2 + 1) - (3x + 2) \times (x^2 - 1)) \\ &= -\frac{1}{3}x \times (3x^3 + 2x^2 + 1) + (x^2 + \frac{2}{3}x + 1) \times (x^2 - 1); \end{aligned}$$

that is,  $-x - 1 = mf + ng$  for  $m = -\frac{1}{3}x$  and  $n = x^2 + \frac{2}{3}x + 1$ .

**Corollary.** The common roots of  $f, g \in \mathbb{F}[x]$  are roots of all  $d \in \gcd(f, g)$ .

**Proof.** By the above theorem,  $d = mf + ng$  for polynomials  $m, n \in \mathbb{F}[x]$ . Thus if  $\alpha$  is a root of  $f$  and  $g$ , then

$$d(\alpha) = m(\alpha)f(\alpha) + n(\alpha)g(\alpha) = m(\alpha) \cdot 0 + n(\alpha) \cdot 0 = 0,$$

so  $\alpha$  is also a root of  $d$ .

From the above theorems, we see that each ideal of  $\mathbb{F}[x]$  is principal.

**Theorem.**  $\mathbb{F}[x]$  is a principal ideal domain.

**Proof.** Modify slightly the corresponding proof on page 139.

**Prime and irreducible polynomials.** As defined in Chapters 1 and 6, a non-zero, non-unit polynomial  $p \in \mathbb{F}[x]$  is **irreducible** if and only if

$$f \text{ or } g \text{ is a unit} \quad \text{whenever} \quad p = fg \text{ with } f, g \in \mathbb{F}[x]. \quad (\text{I})$$

Similarly,  $p \in \mathbb{F}[x]$  is **prime** if and only if

$$p \mid f \text{ or } p \mid g \quad \text{whenever} \quad p \mid fg \text{ with } f, g \in \mathbb{F}[x]. \quad (\text{P})$$

**Example.** Consider  $f = x^2 + 1 \in \mathbb{F}[x]$ . If  $\mathbb{F} = \mathbb{R}$ , then  $f$  is irreducible; whereas if  $\mathbb{F} = \mathbb{C}$ , then  $f = (1 + i)(1 - i)$ , and  $f$  is not irreducible.

In Chapters 1 and 6, we saw that integers and Gaussian integers each are prime if and only if they are irreducible. This is also true for polynomials  $f \in \mathbb{F}[x]$ , and the proof of this is much the same as the corresponding previous proofs. The lemma on page 140 implies that each prime polynomial in  $\mathbb{F}[x]$  is irreducible; from the theorem at the top of page 149, it follows that the converse is also true:

**Theorem.** Each element of  $\mathbb{F}[x]$  is prime if and only if it is irreducible.

**Proof.** Modify slightly the corresponding proof on page 140.

This result is useful but also important; among other things, it implies that  $\mathbb{F}[x]$  is a unique factorisation domain:

**Theorem.** Each  $f \in \mathbb{F}[x]$  factors uniquely into prime polynomials, up to order and multiplication by units.

As with integers, it is often computationally difficult to factorize polynomials. For polynomials of degrees 2 or 3, the following lemma provides an easy way of determining irreducibility.

**Lemma.** If  $f \in \mathbb{F}[x]$  has degree 2 or 3, then  $f$  is irreducible if and only if  $f$  does not have a root.

Although this simple lemma is often useful for small polynomials, it does not apply to polynomials with larger degrees:

**Example.** The polynomial  $f = x^4 + 4 \in \mathbb{R}[x]$  factors into primes as  $f = (x^2 + 2x + 2)(x^2 - 2x + 2)$  and thus neither has roots nor is irreducible.

The next result gives simple sufficient conditions for irreducibility of rational polynomials  $f \in \mathbb{Q}[x]$  with arbitrary degrees.

**Theorem.** *Eisenstein's Criterion.* Let  $f = a_n x^n + \cdots + a_0 \in \mathbb{Q}[x]$ . If there is a prime  $p \in \mathbb{Z}$  that does not divide  $a_n$  but divides every other coefficient of  $f$ , and  $p^2$  does not divide  $a_0$ , then  $f$  is irreducible.

**Proof.** Suppose that  $f = gh$  for some  $g, h \in \mathbb{Q}[x]$ , say

$$g = b_k x^k + \cdots + b_0 \quad \text{and} \quad h = c_l x^l + \cdots + c_0;$$

then  $a_t = \sum_{R_t} b_i c_j$  for all  $t = 0, \dots, n$ , where  $R_t = \{(i, j) : i + j = t\}$ .

Now  $p \mid a_0 = b_0 c_0$  and  $p^2 \nmid a_0$ , so  $p$  divides  $b_0$  or  $c_0$  but not both; suppose that  $p \mid b_0$  and  $p \nmid c_0$ . Now  $p \nmid b_k$  since  $p \nmid a_n = b_k c_l$ , so let  $m$  be a number so that  $p \nmid b_m$  but  $p \mid b_i$  for all  $i < m$ . Then  $p \mid b_0 c_{m-1} + \cdots + b_{m-1} c_0 = a_{m-1}$  and  $p \nmid b_0 c_m + \cdots + b_{m-1} c_1 + b_m c_0 = a_m$ , so  $m = n$ , and thus  $k = n$ . Therefore,  $l = n - k = 0$ , so the polynomial  $c_l x^l + \cdots + c_0$  is just the constant  $c_0$ , a unit, so  $f$  is irreducible.

**Example.** Consider the polynomial  $f = x^7 + 2x^3 + 4x + 6 \in \mathbb{Q}[x]$ . Since  $p = 2$  divides 2, 4, and 6 but not 1, and  $p^2 = 4$  does not divide 6, Eisenstein's Criterion implies that  $f$  is irreducible, and thus is prime.

The next example illustrates how to apply Eisenstein's Criterion indirectly in some of the cases in which it cannot be applied directly.

**Example.** Consider the polynomial  $f = 1 + x + x^2 + x^3 + x^4 \in \mathbb{Q}[x]$ . Since all coefficients of  $f$  equal 1, we cannot apply Eisenstein's Criterion directly to show that  $f$  is irreducible. However, Eisenstein's Criterion with  $p = 5$  implies that  $g = f(x + 1) = x^4 + 5x^3 + 10x^2 + 5x + 5$  is irreducible; therefore,  $f = g(x - 1)$  is also irreducible.

Eisenstein's Criterion provides sufficient conditions for irreducibility. For the field of rational numbers\*, the next lemma can provide additional sufficient conditions for irreducibility, especially for small polynomials.

**Lemma.** If  $f = a_n x^n + \cdots + a_0 \in \mathbb{Z}[x]$  has a root  $\alpha = \frac{p}{q}$  where  $p, q \in \mathbb{Z}$ , then  $p \mid a_0$  and  $q \mid a_n$ . In particular if  $a_n = 1$ , then  $\alpha$  is an integer.

**Example.** Consider the rational polynomial  $f = x^3 + 2x + 7 \in \mathbb{Q}[x]$ . Since  $f(7), f(-7) \neq 0$ , the above lemma implies that  $f$  has no roots. Since  $\deg f = 3$ , we conclude that  $f$  is irreducible.

\* and, more generally, a quotient field over an integral domain.

**Algebraic Extension Fields.** We saw in previous chapters several examples of subrings of rings; for instance, the Gaussian integers  $\mathbb{Z}[i]$  of Chapter 6 are principal ideal domains that are contained in the field of complex numbers  $\mathbb{C}$  and in turn contain the principal ideal domain  $\mathbb{Z}$ . We will now have a closer look at the relationship between two fields when one is a subfield of the other.

**Definition.** A field  $\mathbb{K}$  is an **extension (field)** of  $\mathbb{F}$  if  $\mathbb{F}$  is a subfield  $\mathbb{K}$ .

**Example.** The set  $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$  forms an extension field of  $\mathbb{Q}$  since it contains  $\mathbb{Q}$  and is itself a field. The last part is true because  $\mathbb{Q}(\sqrt{2})$  is an integral domain in which every non-zero element  $a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$  has an inverse:

$$(a + b\sqrt{2})^{-1} = \frac{1}{a + b\sqrt{2}} = \frac{1}{a + b\sqrt{2}} \frac{a - b\sqrt{2}}{a - b\sqrt{2}} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2} \sqrt{2} \in \mathbb{Q}(\sqrt{2}).$$

An extension field  $\mathbb{K}$  of a field  $\mathbb{F}$  is a *vector space* over that field  $\mathbb{F}$ . Therefore, we may compare the sizes of  $\mathbb{F}$  and  $\mathbb{K}$  by dimension:

**Definition.** Suppose that a field  $\mathbb{K}$  is an extension field over a field  $\mathbb{F}$ . Then let  $[\mathbb{K} : \mathbb{F}]$  denote the vector space dimension  $\dim_{\mathbb{F}} \mathbb{K}$  of  $\mathbb{K}$  over  $\mathbb{F}$ .

Note that  $[\mathbb{F} : \mathbb{F}] = 1$  since  $\mathbb{F}$  is spanned by the basis  $\{1\} \subset \mathbb{F}$ .

**Example.**  $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$  is an extension field of  $\mathbb{Q}$ . Since  $\mathbb{Q}(\sqrt{2}) = \text{span}_{\mathbb{Q}}\{1, \sqrt{2}\}$  and  $\sqrt{2} \notin \mathbb{Q}$ , it therefore follows that  $\{1, \sqrt{2}\}$  forms a basis for  $\mathbb{Q}(\sqrt{2})$  over  $\mathbb{Q}$ , and  $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$ .

**Theorem.** If  $\mathbb{F} \subseteq \mathbb{G} \subseteq \mathbb{K}$  for fields  $\mathbb{F}, \mathbb{G}, \mathbb{K}$ , then  $[\mathbb{K} : \mathbb{F}] = [\mathbb{K} : \mathbb{G}][\mathbb{G} : \mathbb{F}]$ .

**Proof.** Suppose that  $\mathbb{F} \subseteq \mathbb{G} \subseteq \mathbb{K}$ , let  $U$  be a basis for  $\mathbb{G}$  over  $\mathbb{F}$ , let  $V$  be a basis for  $\mathbb{K}$  over  $\mathbb{G}$ , and set  $W := \{uv : u \in U, v \in V\}$ . Since  $V$  is linearly independent over  $\mathbb{G}$ , the elements of  $W$  are distinct, so  $|W| = |U||V| = [\mathbb{K} : \mathbb{G}][\mathbb{G} : \mathbb{F}]$ . We want to show that  $[\mathbb{K} : \mathbb{F}] = |W|$ .

This is true if  $W$  is a basis of  $\mathbb{K}$  over  $\mathbb{F}$ , so suppose that  $k \in \mathbb{K}$ . Since  $V$  spans  $\mathbb{K}$ , we see that  $k = \sum_{v \in V} a_v v$  for some  $a_v \in \mathbb{G}$ , which in turn each may be expressed as  $a_v = \sum_{u \in U} b_{a_v} u$  for some  $b_{a_v} \in \mathbb{F}$ . Therefore,

$$k = \sum_{v \in V} a_v v = \sum_{v \in V} \left( \sum_{u \in U} b_{a_v} u \right) v = \sum_{uv \in W} b_{a_v} (uv),$$

so  $k \in \text{span}_{\mathbb{F}}W$ . In other words,  $W$  spans  $\mathbb{K}$ .

Now suppose that  $\sum_{uv \in W} b_{uv}(uv) = 0$  for some  $b_{uv} \in \mathbb{F}$ . Then

$$0 = \sum_{uv \in W} b_{uv}(uv) = \sum_{v \in V} \left( \sum_{u \in U} b_{uv}u \right) v,$$

so  $0 = \sum_{u \in U} b_{uv}u$  for all  $v \in V$  since  $V$  is linearly independent over  $\mathbb{G}$ . But  $U$  is linear independent over  $\mathbb{F}$ , so  $b_{uv} = 0$  for all  $u \in U$  and  $v \in V$ . Therefore,  $W$  is a basis for  $\mathbb{K}$  over  $\mathbb{F}$ , and  $[\mathbb{K} : \mathbb{F}] = |W| = [\mathbb{K} : \mathbb{G}][\mathbb{G} : \mathbb{F}]$ .

**Example.** The set  $\mathbb{Q}(\sqrt{2}, i) = \{a + b\sqrt{2} + ci + di\sqrt{2} : a, b, c, d \in \mathbb{Q}\}$  may be shown to be a field and is thus an extension field of  $\mathbb{Q}(\sqrt{2})$ . The set  $\{1, \sqrt{2}, i, i\sqrt{2}\}$  is a basis for  $\mathbb{Q}(\sqrt{2}, i)$  over  $\mathbb{Q}$ , so  $[\mathbb{Q}(\sqrt{2}, i) : \mathbb{Q}] = 4$ . As we saw in the example above,  $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$ , so the above theorem implies that  $[\mathbb{Q}(\sqrt{2}, i) : \mathbb{Q}(\sqrt{2})] = [\mathbb{Q}(\sqrt{2}, i) : \mathbb{Q}] / [\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 4/2 = 2$ . Indeed, we see that  $\{1, i\}$  is a basis for  $[\mathbb{Q}(\sqrt{2}, i) : \mathbb{Q}(\sqrt{2})]$  over  $\mathbb{Q}(\sqrt{2})$ .

If  $\mathbb{F}_1$  and  $\mathbb{F}_2$  are fields containing a common zero and identity, then the intersection  $\mathbb{F}_1 \cap \mathbb{F}_2$  is also a field. This allows the following definition:

**Definition.** Let  $\mathbb{F}$  be a subfield of a field  $\mathbb{K}$  and suppose that  $\alpha \in \mathbb{K}$ . Then let  $\mathbb{F}(\alpha)$  denote the smallest subfield of  $\mathbb{K}$  containing  $\mathbb{F}$  and  $\alpha$ ; similarly, let  $\mathbb{F}[\alpha]$  denote the smallest subring of  $\mathbb{K}$  containing  $\mathbb{F}$  and  $\alpha$ .

**Example.**  $\mathbb{Q}[\sqrt[3]{2}] = \{x + y\sqrt[3]{2} + z\sqrt[3]{4} : x, y, z \in \mathbb{Q}\}$  since  $(\sqrt[3]{2})^2 = \sqrt[3]{4}$ . It may be shown that  $\mathbb{Q}[\sqrt[3]{2}]$  is in fact a field, so  $\mathbb{Q}[\sqrt[3]{2}] = \mathbb{Q}(\sqrt[3]{2})$ .

We note some simple but useful facts about  $\mathbb{F}[\alpha]$  and  $\mathbb{F}(\alpha)$ :

**Theorem.** If  $\mathbb{F}$  is a subfield of a field  $\mathbb{K}$  and  $\alpha \in \mathbb{K}$ , then

- $\mathbb{F}[\alpha]$  is an integral domain;
- $\mathbb{F}[\alpha] \subseteq \mathbb{F}(\alpha) \subseteq \mathbb{K}$ ;
- $\mathbb{F}[\alpha] = \mathbb{F}(\alpha) = \mathbb{F}$  whenever  $\alpha \in \mathbb{F}$ .

**Proof.** Exercise.

**Example.** The ring  $\mathbb{F}[\alpha]$  and field  $\mathbb{F}(\alpha)$  sometimes coincide; for instance,  $\mathbb{C} = \mathbb{R}[i] = \mathbb{R}(i)$ ,  $\mathbb{Q}[\sqrt{2}] = \mathbb{Q}(\sqrt{2})$ , and  $\mathbb{Q}[\sqrt[3]{2}] = \mathbb{Q}(\sqrt[3]{2})$ . However, they usually differ; for instance,  $\mathbb{Q}[\pi] \neq \mathbb{Q}(\pi)$ .

**Theorem.** If  $\mathbb{F}$  is a subfield of a field  $\mathbb{K}$  and  $\alpha \in \mathbb{K}$ , then

$$\mathbb{F}[\alpha] = \{f(\alpha) : f \in \mathbb{F}[x]\}$$

and  $\mathbb{F}(\alpha) = \left\{ \frac{f(\alpha)}{g(\alpha)} : f, g \in \mathbb{F}[x], g \neq 0 \right\}.$

**Proof.** Set  $P = \{f(\alpha) : f \in \mathbb{F}[x]\}$  and note that  $P$  is the set of all elements obtained by adding and multiplying elements of  $\mathbb{F} \cup \{\alpha\}$ . Since rings are closed under addition and multiplication,  $P$  must be contained in all subrings of  $\mathbb{K}$  that contain  $\mathbb{F}$  and  $\alpha$ . But  $P$  is also such a ring, since  $\mathbb{F}[x]$  is a ring and  $\mathbb{F} \cup \{\alpha\} \subseteq P \subseteq \mathbb{K}$ . Therefore,  $\mathbb{F}[\alpha] = P$ .

Similarly, note that  $R = \left\{ \frac{f(\alpha)}{g(\alpha)} : f, g \in \mathbb{F}[x], g(\alpha) \neq 0 \right\}$  is the set of all elements obtained by adding, multiplying, and non-zero dividing elements of  $\mathbb{F} \cup \{\alpha\}$ . Since fields are closed under addition, multiplication, and non-zero division,  $R$  must be contained in all subfields of  $\mathbb{K}$  that contain both  $\mathbb{F}$  and  $\alpha$ . However,  $\mathbb{F} \cup \{\alpha\} \subseteq R \subseteq \mathbb{K}$  and  $\left(\frac{f(\alpha)}{g(\alpha)}\right)^{-1} = \frac{g(\alpha)}{f(\alpha)}$  for all non-zero  $\frac{f(\alpha)}{g(\alpha)} \in R$ , so  $R$  is itself such a field. Therefore,  $\mathbb{F}(\alpha) = R$ .

To study  $\mathbb{F}[\alpha] = \mathbb{F}(\alpha)$  further, we need some definitions:

**Definition.** Let  $\mathbb{F}$  be a subfield of a field  $\mathbb{K}$  and suppose that  $\alpha \in \mathbb{K}$ . Then  $\alpha$  is **algebraic** over  $\mathbb{F}$  if  $\alpha$  is the root of a polynomial  $f \in \mathbb{F}[x]$ ; otherwise,  $\alpha$  is **transcendental** over  $\mathbb{F}$ . A complex number is *algebraic* if it is algebraic over the rational numbers, and *transcendental* otherwise. If  $\alpha$  is algebraic over  $\mathbb{F}$ , then  $\mathbb{F}(\alpha)$  is a **simple algebraic extension** of  $\mathbb{F}$ . If  $\mathbb{G} \subseteq \mathbb{K}$  is obtained by a sequence of simple algebraic extensions of  $\mathbb{F}$ , then  $\mathbb{G}$  is an **algebraic extension field** of  $\mathbb{F}$ .

Note that  $\alpha \in \mathbb{F}$  is algebraic over  $\mathbb{F}$  since it is the root of  $x - \alpha \in \mathbb{F}[x]$ .

**Example.**  $i \in \mathbb{C}$  is algebraic over  $\mathbb{R}$  since it is a root of  $f = x^2 + 1 \in \mathbb{R}[x]$ . Similarly, the real numbers  $\sqrt{2}$  and  $\sqrt[3]{4}$  are algebraic (over  $\mathbb{Q}$ ) since they are roots of rational polynomials, such as  $(x^2 - 2)(x^3 - 4)$ , for instance.

Interestingly, the transcendental numbers have been proven to far outweigh the algebraic number but only few transcendental numbers have actually been found, including  $e$  and  $\pi$ , perhaps due to the difficulty in proving the non-existence of polynomials with these numbers as roots.

If  $\alpha \in \mathbb{K}$  is algebraic over  $\mathbb{F} \subseteq \mathbb{K}$ , then  $\alpha$  is a root of at least one polynomial  $f \in \mathbb{F}[x]$ . If  $\alpha$  is also a root of another polynomial  $g \in \mathbb{F}[x]$ , then, by the corollary on page 149,  $\alpha$  is a root of each greatest common divisor of  $f$  and  $g$ . This validates the following definition.

**Definition.** A unique monic, irreducible polynomial  $f \in \mathbb{F}[x]$  has  $\alpha$  as a root. This is the **minimal** (or **irreducible**) **polynomial** of  $\alpha$  over  $\mathbb{F}$ , and  $\deg f$  is the **degree** of  $\alpha$  over  $\mathbb{F}$ .

Note that the minimal polynomial of  $\alpha$  is *not* irreducible in  $\mathbb{K}$ .

**Example.** The minimal polynomial of  $i \in \mathbb{C}$  over  $\mathbb{R}$  is  $f = x^2 + 1 \in \mathbb{Q}[x]$ . Similarly, the minimal polynomials of the real numbers  $\sqrt{2}$  and  $\sqrt[3]{4}$  over  $\mathbb{Q}$  are the rational polynomials  $x^2 - 2$  and  $x^3 - 4$ , respectively. The degrees of  $\sqrt{2}$  and  $\sqrt[3]{4}$  over  $\mathbb{Q}$  are thus 2 and 3, respectively.

**Example.** Since  $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ , the real number  $\alpha = \cos 20^\circ$  is a root of  $8x^3 - 6x - 1 \in \mathbb{Q}[x]$ , so  $\alpha$  is algebraic (over  $\mathbb{Q}$ ).

**Example.** Consider the real number  $\alpha = \sqrt{2} + \sqrt{3}$ .

To find the minimal polynomial and degree of  $\alpha$  over  $\mathbb{Q}$ , first note that  $\alpha^2 = 5 + 2\sqrt{6}$  and thus that  $(\alpha^2 - 5)^2 = 24$ . Therefore,  $\alpha$  is a root of the polynomial  $f = x^4 - 10x^2 + 1 \in \mathbb{Q}[x]$ . Since  $y^2 - 10y + 1 \in \mathbb{R}[y]$  has roots  $5 \pm 2\sqrt{6}$ ,  $f$  has no rational root and is therefore irreducible in  $\mathbb{Q}$ , so it is the minimal polynomial of  $\alpha$  over  $\mathbb{Q}$ , and  $\alpha$  has degree 4 over  $\mathbb{Q}$ .

**Example.** Find the minimal polynomial and degree of  $\cos \frac{2\pi}{5}$  over  $\mathbb{Q}$ .

Set  $\alpha = \cos \frac{2\pi}{5}$  and  $\zeta = e^{2\pi i/5}$ , and note that  $\zeta^5 = 1$  and that  $2\alpha = \zeta + \frac{1}{\zeta}$ . If  $s = \zeta^4 + \zeta^3 + \zeta^2 + \zeta + 1$ , then  $s(1 - \zeta) = s - s\zeta = 1 - \zeta^5 = 1 - 1 = 0$ , so  $\zeta^4 + \zeta^3 + \zeta^2 + \zeta + 1 = s = 0$ , or dividing by  $\zeta^2$ ,  $(\zeta + \frac{1}{\zeta})^2 + (\zeta + \frac{1}{\zeta}) - 1 = 0$ . Thus,  $\alpha$  is a root of  $4x^2 + 2x - 1$  which has no rational roots and is thus irreducible in  $\mathbb{Q}$ . It follows that  $x^2 + \frac{1}{2}x - \frac{1}{4}$  is the minimal polynomial, and 2 the degree, of  $\alpha = \cos \frac{2\pi}{5}$  over  $\mathbb{Q}$ .

**Example.** Consider the real number  $\alpha = \sqrt{1 + \sqrt[3]{2}}$  over  $\mathbb{Q}$ .

To find the minimal polynomial and degree of  $\alpha$ , note that  $(\alpha^2 - 1)^3 = 2$ . Therefore,  $\alpha$  is a root of  $f = (x^2 - 1)^3 - 2 = x^6 - 3x^4 + 3x^2 - 3 \in \mathbb{Q}[x]$ . Applying Eisenstein's Criterion with  $p = 3$  shows that  $f$  is irreducible. Thus,  $f$  is the minimal polynomial, and 6 the degree, of  $\alpha$  over  $\mathbb{Q}$ .

**Theorem.** Let  $\mathbb{K}$  be an extension field of a field  $\mathbb{F}$  and consider  $\alpha \in \mathbb{K}$ . If  $\alpha$  is algebraic of degree  $n$  over  $\mathbb{F}$ , then

- $\mathbb{F}[\alpha] = \mathbb{F}(\alpha)$ ;
- $[\mathbb{F}(\alpha) : \mathbb{F}] = n$ ;
- $\{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$  is a basis for  $\mathbb{F}(\alpha)$  over  $\mathbb{F}$ ;
- each element of  $\mathbb{F}(\alpha)$  may be written uniquely in the form

$$a_0 + a_1\alpha + \dots + a_{n-1}\alpha^{n-1} \quad \text{where } a_i \in \mathbb{F}.$$

**Proof.** We first prove that  $S = \{a_0 + a_1\alpha + \dots + a_{n-1}\alpha^{n-1} : a_i \in \mathbb{F}\}$  forms a field. The set  $S$  is closed under addition; to see that it is closed under multiplication, we only need check that  $\alpha^r\alpha^s \in S$  for all  $r, s$ ; that is, that  $\alpha^k \in S$  for all  $k \geq 0$ . First, note that  $\alpha^k \in S$  for all  $k = 0, \dots, n-1$ , so assume that  $\alpha^k \in S$  for some  $k$ . By definition,  $\alpha^k = a_0 + a_1\alpha + \dots + a_{n-1}\alpha^{n-1}$  for some elements  $a_i \in \mathbb{F}$ , and therefore  $\alpha^{k+1} = a_0\alpha + a_1\alpha^2 + \dots + a_{n-1}\alpha^n$ . Let  $f = c_0 + c_1x + \dots + c_{n-1}x^{n-1} + c_nx^n$  be the minimal polynomial of  $\alpha$  over  $\mathbb{F}$ . Since  $f(\alpha) = 0$ , it follows that  $\alpha^n = -\frac{1}{c_n}(c_0 + c_1\alpha + \dots + c_{n-1}\alpha^{n-1})$ . Therefore,

$$\alpha^{k+1} = a_0\alpha + a_1\alpha^2 + \dots + a_{n-1}\left(-\frac{1}{c_n}(c_0 + c_1\alpha + \dots + c_{n-1}\alpha^{n-1})\right),$$

so  $\alpha^{k+1} \in S$ . By induction,  $\alpha^k \in S$  for all  $k \geq 0$ , so  $S$  is closed under multiplication. It is now easy to see that  $S$  forms a commutative ring with identity 1. To show that  $S$  forms a field, we need to show that each non-zero element  $s \in S$  has an inverse in  $S$ . By definition of  $S$ , there is a polynomial  $g \in \mathbb{F}[x]$  with  $\deg g < n$  and  $g(\alpha) = s$ . Since  $f$  is irreducible in  $\mathbb{F}[x]$ ,  $1 \in \gcd(f, g)$ , and so by the first theorem on page 149, there are polynomials  $m, n \in \mathbb{F}[x]$  so that  $1 = mf + ng$ . Then

$$n(\alpha)s = n(\alpha)g(\alpha) = 1 - m(\alpha)f(\alpha) = 1 - m(\alpha) \cdot 0 = 1,$$

so  $n(\alpha) = s^{-1}$  in  $\mathbb{F}[x]$ . Then  $n(\alpha) \in S$ , since  $n(\alpha)$  is a polynomial in  $\alpha$  and  $S$  is closed under addition and multiplication, so  $S$  is a field.

By the theorem on page 154,  $\mathbb{F} \cup \{\alpha\} \subseteq S \subseteq \mathbb{F}[\alpha] \subseteq \mathbb{F}(\alpha)$ , so the minimality of  $\mathbb{F}(\alpha)$  implies that  $S = \mathbb{F}[\alpha] = \mathbb{F}(\alpha)$ .

The remaining three statements are equivalent since  $\mathbb{F}(\alpha)$  is spanned by  $U = \{1, \alpha, \dots, \alpha^{n-1}\}$  over  $\mathbb{F}$ . We only need to show that  $U$  is linearly independent over  $\mathbb{F}$ . Suppose that  $a_0 + a_1\alpha + \dots + a_{n-1}\alpha^{n-1} = 0$  for some  $a_i \in \mathbb{F}$ . Then  $g(\alpha) = 0$  for  $g = a_0 + a_1x + \dots + a_{n-1}x^{n-1} \in \mathbb{F}[x]$ . Since  $\deg g < n$ , this only happens if  $g = 0$ . Thus,  $a_0 = a_1 = \dots = a_{n-1} = 0$ , and  $U$  is linearly independent over  $\mathbb{F}$ .

**Example.** We have seen that  $\sqrt{2} + \sqrt{3}$  is algebraic of degree 4 over  $\mathbb{Q}$ . By the above theorem,  $\mathbb{Q}[\sqrt{2} + \sqrt{3}]$  is a field;  $[\mathbb{Q}(\sqrt{2} + \sqrt{3}) : \mathbb{Q}] = 4$ ; and

$$\begin{aligned} \mathbb{Q}(\sqrt{2} + \sqrt{3}) &= \text{span}_{\mathbb{Q}}\{1, \sqrt{2} + \sqrt{3}, (\sqrt{2} + \sqrt{3})^2, (\sqrt{2} + \sqrt{3})^3\} \\ &= \text{span}_{\mathbb{Q}}\{1, \sqrt{2} + \sqrt{3}, 5 + 2\sqrt{6}, 11\sqrt{2} + 9\sqrt{3}\} \\ &= \text{span}_{\mathbb{Q}}\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}. \end{aligned}$$

We can now characterise when  $\mathbb{F}[\alpha] = \mathbb{F}(\alpha)$ .

**Corollary.** Let  $\mathbb{K}$  be an extension field of a field  $\mathbb{K}$  and consider  $\alpha \in \mathbb{K}$ . Then  $\alpha$  is algebraic over  $\mathbb{F}$  if and only if  $\mathbb{F}[\alpha] = \mathbb{F}(\alpha)$ .

**Proof.** By the theorem above,  $\mathbb{F}[\alpha] = \mathbb{F}(\alpha)$  when  $\alpha$  is algebraic over  $\mathbb{F}$ , so suppose that  $\mathbb{F}[\alpha] = \mathbb{F}(\alpha)$ . Then  $\mathbb{F}[\alpha]$  is a field, and so  $\alpha\beta = 1$  for some element  $\beta = a_0 + a_1\alpha + \dots + a_{n-1}\alpha^{n-1} \in \mathbb{F}[\alpha]$ . Now define  $g = x(a_0 + a_1x + \dots + a_{n-1}x^{n-1}) - 1 \in \mathbb{F}[x]$ . Then  $g(\alpha) = \alpha\beta - 1 = 0$ , so  $\alpha$  is a root of  $g \in \mathbb{F}[x]$  and is thus algebraic over  $\mathbb{F}$ .

**Example.**  $\pi$  is transcendental, so  $\mathbb{Q}[\pi] \neq \mathbb{Q}(\pi)$ , and  $\mathbb{Q}[\pi]$  is not a field.

**Complex Algebraic Extension Fields.** To conclude the chapter, we show that any sequence of simple algebraic extensions of a complex field may be obtained from a single algebraic number.\* More precisely, let  $\mathbb{F}(\alpha, \beta)$  denote the field extension  $(\mathbb{F}(\alpha))(\beta) = (\mathbb{F}(\beta))(\alpha)$ , that is, the smallest field containing  $\alpha$  and  $\beta$ .

**Theorem.** If  $\mathbb{F} \subseteq \mathbb{C}$  is a complex field and  $\alpha, \beta \in \mathbb{C}$  are algebraic over  $\mathbb{F}$ , then  $\mathbb{F}(\alpha, \beta) = \mathbb{F}(\theta)$  for some  $\theta \in \mathbb{C}$ .

**Proof.** Let  $f$  and  $g$  be the minimal polynomials of  $\alpha$  and  $\beta$  over  $\mathbb{F}$ , respectively, and let  $\alpha, \alpha_1, \dots, \alpha_n$  and  $\beta, \beta_1, \dots, \beta_m$  be the respective roots of  $f$  and  $g$ . Let  $t$  be a rational number for which  $\alpha + t\beta \neq \alpha_i + t\beta_j$  for all  $i, j$ . Set  $\theta = \alpha + t\beta \in \mathbb{F}(\alpha, \beta)$  and  $h = f(\theta - t\alpha)$ . Then  $h \in (\mathbb{F}(\theta))[x]$  and  $h(\beta) = 0$  but, by choice of  $t$ ,  $h(\beta_j) = f(\alpha + t\beta - t\beta_j) \neq 0$  for all  $j$ .

If  $k$  is the minimal polynomial of  $\beta$  over  $\mathbb{F}(\theta)$ , then  $k \mid g$  and  $k \mid h$ . Since  $g$  and  $h$  both have the root  $\beta$  but have no other root in common, it follows that  $k = x - \beta$  by choice of  $t$ . But  $k \in (\mathbb{F}(\theta))[x]$ , so  $\beta \in \mathbb{F}(\theta)$ .

Similarly,  $\alpha = \theta - t\beta \in \mathbb{F}(\theta)$ , so  $\mathbb{F}(\alpha, \beta) \subseteq \mathbb{F}(\theta) \subseteq \mathbb{F}(\alpha, \beta)$ .

**Example.** By the above theorem,  $\mathbb{Q}(\sqrt{2}, \sqrt{3})$  is a simple extension field  $\mathbb{Q}(\theta)$  for some  $\theta \in \mathbb{C}$ . To find such  $\theta$ , we apply the method of the proof. The minimal polynomials of  $\alpha = \sqrt{2}$  and  $\beta = \sqrt{3}$  over  $\mathbb{Q}$  are  $g = x^2 - 2$  and  $g = x^2 - 3$ , respectively, and their roots are  $(\alpha, \alpha_1) = (\sqrt{2}, -\sqrt{2})$  and  $(\beta, \beta_1) = (\sqrt{3}, -\sqrt{3})$ . Set  $t$  so that  $\alpha + t\beta \neq \alpha_1 + t\beta_1$ , say  $t = 1$ , and set  $\theta = \alpha + t\beta = \sqrt{2} + \sqrt{3}$ . By the theorem's proof,  $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3})$ .

---

\* This is a special case of the **Primitive Element Theorem**.

---



---

The University of New South Wales  
School of Mathematics and Statistics

**MATH3521**

**ALGEBRAIC TECHNIQUES IN NUMBER THEORY**

---



---

**8. CONSTRUCTIBILITY.**

The mathematics of the ancient Greeks was primarily geometric. The then-known numbers, namely the positive integers and rational numbers, and, later, also some irrational numbers, were mostly treated in terms of geometric lengths, areas, volumes, and ratios between any of these. Thus, for instance, the square root of a number  $x$  could be seen as the side length of a square with area  $x$ . In addition, the ancient Greeks were the first to consider geometric angles as numbers.

Much of ancient Greek geometry was about constructing geometric objects. Many construction tools were developed, including mechanical ones\*, but the standard set of tools allowed for geometric construction was an unmarked ruler and a compass. The rules of construction for these tools was to start with some points, draw lines through pairs of points or draw circles centering in and intersecting points, and obtain new points from the intersections of the lines and circles drawn.

Some simple examples of constructions are given below, including an example that shows how to bisect (halve) a given angle. The problem of trisecting a given angle was posed in the 5th century BC by the ancient Greeks but was not solved until 1837, by P. L. Wantzel, who also solved two other classical construction problems.

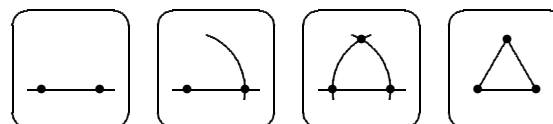
In this chapter, we will use the algebraic number fields of Chapter 7 to solve these three classical construction problems.

---

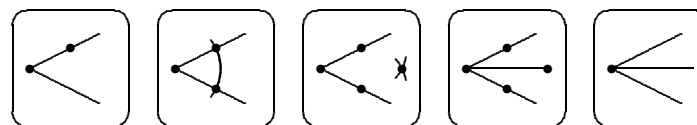
\* The ancient Greeks built several advanced mechanical computers: see [www.antikythera-mechanism.gr](http://www.antikythera-mechanism.gr).

Let us first look at some examples of constructions.

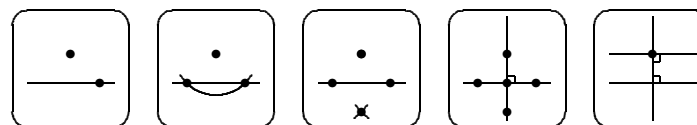
**Example.** We can construct a triangle as follows:



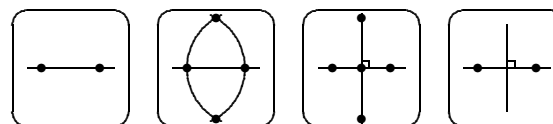
**Example.** Any angle can be bisected as follows:



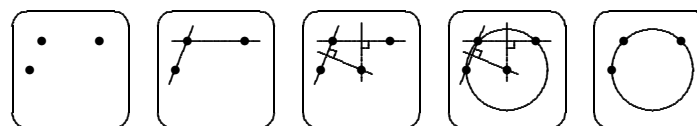
**Example.** We can construct the line orthogonal (i.e., at right angles) to a given line through a given point. By doing this twice, we can also draw the line parallel to the given line through the given point:



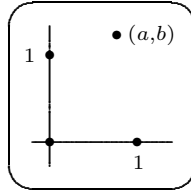
**Example.** A line interval can be bisected by an orthogonal line:



**Example.** By such bisecting, we can construct the circle intersecting three given points:



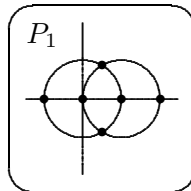
In order to examine such constructions more systematically, we will suppose that we initially have a set  $P_0$  of two points, labeled 0 and 1, say. Draw a line through these points, along with the orthogonal line that passes through the point 0. These lines form the two axes of a coordinate system in which any points subsequently constructed may be mapped; the points 0 and 1 lie in the positions  $(0, 0)$  and  $(1, 0)$ , respectively, and are identified with these coordinates.



Now we will perform a finite number of construction steps, described as follows. At step  $i$ , let  $P_i$  denote the set of points constructed so far. For each pair of points in  $P_i$ , draw an infinite line through the two points and also draw a circle that has one point as center and that intersects the other point. Then let  $P_{i+1}$  denote the set of points  $P_i$  together with a new point, chosen from the intersection points of lines and/or circles.

**Definition.** Every point in each constructed set  $P_i$  is **constructible**, as are all lines, circles, and angles constructed as above from  $P_i$ . Also, a real number  $a \in \mathbb{R}$  is said to be constructible if  $(a, 0)$  is constructible.

**Example.** Note that every integer is constructible and that  $(a, b)$  is constructible if and only if  $(b, a)$  is constructible; or, if and only if  $a$  and  $b$  both are constructible. By drawing circles with center  $(0, 0)$ , we also see that  $a$  is constructible if and only  $-a$  is. By the figure on the right, we see that  $P_1$  contains the set  $P_0$  and one of the points



$$(-1, 0) \quad (2, 0) \quad \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right) \quad \left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right),$$

so  $0, \pm 1, \pm \frac{1}{2}, \pm \frac{\sqrt{3}}{2}$ , and  $\pm 2$  are all constructible real numbers.

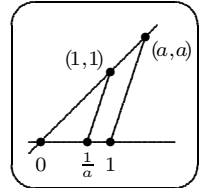
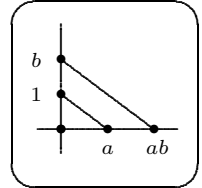
**Definition.** At each step  $i$ , let  $\mathbb{K}_i$  denote the smallest field containing the constructible numbers  $\{a : (a, b) \in P_i\}$ , and set

$$\mathbb{K}_\infty = \mathbb{K}_0 \cup \mathbb{K}_1 \cup \mathbb{K}_2 \cup \dots$$

**Example.**  $\mathbb{Q}$  is the smallest field containing the real set  $P_0 = \{0, 1\}$ , so  $\mathbb{K}_0 = \mathbb{Q}$ . Here usual addition and multiplication is assumed; otherwise,  $P_0$  may itself be a field, say  $\mathbb{Z}_2$ ! If  $\mathbb{K}_1$  contains  $\frac{\sqrt{3}}{2}$ , then  $\mathbb{K}_1 = \mathbb{Q}(\sqrt{3})$ ; otherwise  $\mathbb{K}_1 = \mathbb{K}_0$ . Note that  $\mathbb{K}_0 \subseteq \mathbb{K}_1 \subseteq \dots \subseteq \mathbb{K}_i \subseteq \dots \subseteq \mathbb{K}_\infty \subseteq \mathbb{R}$ .

**Theorem.** The field  $\mathbb{K}_\infty$  consists of all constructible numbers.

**Proof.** It suffices to show that the constructible numbers form a field. If  $a, b \in \mathbb{R}$  are constructible, then  $a + b$  and  $-a$  are constructible. To show that  $ab$  is also constructible, draw a line through points  $(0, 1)$  and  $(a, 0)$ , and draw a second line through the point  $(0, b)$  parallel to the first line; see the figure on the right. The second line intersects the  $x$ -axis in the point  $(ab, 0)$ , so  $ab$  is constructible. To show that  $\frac{1}{a}$  is constructible, suppose that  $a > 0$ . Draw a line through points  $(1, 0)$  and  $(a, a)$ , draw a second line through the point  $(1, 1)$  parallel to the first line, and draw a line through points  $(0, 0)$  and  $(a, a)$ ; see the figure on the right. The second line meets the  $x$ -axis in a point, say  $(x, 0)$ . By triangle similarity, we can see that  $\frac{a}{1} = \frac{1}{x}$ , and so  $\frac{1}{a} = x$  is constructible. Therefore, the constructible numbers form a field.

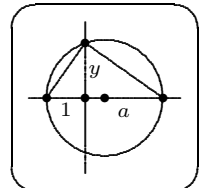


**Comment.** By above theorem, each element  $a \in \mathbb{K}_\infty$  is constructible and is thus a number in some set  $P_i$ ; that is,  $a$  may be constructed in a finite number of steps by compass and ruler operations.

In addition to being able to add, subtract, multiply, and divide constructible numbers by compass and ruler constructions, we can also construct the square roots of each non-negative constructible number:

**Theorem.** If  $a \geq 0$  is constructible, then  $\sqrt{a}$  is also constructible.

**Proof.** The number  $b = \frac{a-1}{2}$  is constructible, so draw the circle with centre  $(b, 0)$  that intersects points  $(-1, 0)$  and  $(a, 0)$ . The circle intersects the  $y$ -axis in a point, say  $(0, y)$ . Draw a line between the two points  $(-1, 0)$  and  $(0, y)$ , and draw a line between the points  $(0, y)$  and  $(a, 0)$ . With the axes, these lines define two adjacent right-angled triangles; see the figure. By triangle similarity,  $\frac{a}{y} = \frac{y}{1}$ , so  $a = y^2$ . Hence,  $\sqrt{a} = y$  is constructible.



**Corollary.** The field  $\mathbb{K}_\infty$  is closed with respect to taking square roots.

On adding an intersection point to our collection of points, the corresponding field remains the same or is extended by a single element.

**Theorem.**  $\mathbb{K}_{i+1} = \mathbb{K}_i(\sqrt{u})$  for some  $u \in \mathbb{K}_i$ .

**Proof.** The point  $p = (a, b)$  added to  $P_i$  to form  $P_{i+1}$  lies in one of the intersections of lines and/or circles determined by the points  $P_i$ . Now, lines determined by the points  $P_i$  are given by linear equations with coefficients in  $\mathbb{K}_i$ ; similarly, circles determined by  $P_i$  are given by quadratic equations with coefficients in  $\mathbb{K}_i$ . Thus, if  $p$  is the intersection of two lines, then  $a$  is a rational form in the coefficients of these lines, so  $a \in \mathbb{K}$ , and  $\mathbb{K}_{i+1} = \mathbb{K}_i = \mathbb{K}_i(\sqrt{1})$ . If  $p$  is the intersection of a circle and a line or another circle, then  $a = \frac{s+\sqrt{u}}{t}$  for some  $s, t, u \in \mathbb{K}_i$ . If  $u$  is square in  $\mathbb{K}_i$ , then  $a \in \mathbb{K}_i$ , and  $\mathbb{K}_{i+1} = \mathbb{K}_i = \mathbb{K}_i(\sqrt{1})$ ; otherwise,  $a \in \mathbb{K}_i(\sqrt{u})$ , and  $\mathbb{K}_{i+1} = \mathbb{K}_i(\sqrt{u})$ .

**Corollary.**  $[\mathbb{K}_i : \mathbb{K}_{i-1}] \leq 2$ .

**Proof.** By the theorem,  $\mathbb{K}_{i+1} = \mathbb{K}_i(\sqrt{u})$  for some  $u \in \mathbb{K}_i$ . If  $\sqrt{u} \in \mathbb{K}_i$ , then  $\mathbb{K}_{i+1} = \mathbb{K}_i$ , so  $[\mathbb{K}_{i+1} : \mathbb{K}_i] = 1$ ; otherwise,  $u$  is a root of  $f = x^2 - u$ , a monic and irreducible polynomial over  $\mathbb{K}_i$ , so  $[\mathbb{K}_{i+1} : \mathbb{K}_i] = \deg f = 2$ .

**Example.** If  $\frac{\sqrt{3}}{2} \in \mathbb{K}_1$ , then  $[\mathbb{K}_1 : \mathbb{K}_0] = [\mathbb{K}_0(\sqrt{3}) : \mathbb{K}_0] = 2$ ; otherwise,  $[\mathbb{K}_1 : \mathbb{K}_0] = [\mathbb{K}_0 : \mathbb{K}_0] = 1$ .

**Corollary.**  $[\mathbb{K}_i : \mathbb{Q}] = 2^m$  for some  $m \in \mathbb{N}_0$ .

**Proof.** Since  $\mathbb{K}_0 = \mathbb{Q}$ ,  $[\mathbb{K}_i : \mathbb{Q}] = [\mathbb{K}_i : \mathbb{K}_{i-1}][\mathbb{K}_{i-1} : \mathbb{K}_{i-2}] \cdots [\mathbb{K}_1 : \mathbb{K}_0]$ , so the result follows from the previous corollary.

The following result describes the constructible numbers.

**Theorem.** Each constructible number  $c \in \mathbb{K}_\infty$  is algebraic of degree  $2^k$  over  $\mathbb{Q}$  for some  $k \in \mathbb{N}_0$ .

**Proof.** The number  $c$  is in some field  $\mathbb{K}_i$  for some  $i$ , and  $\mathbb{Q}(c) \subseteq \mathbb{K}_i$ . Then  $[\mathbb{K}_i : \mathbb{Q}] = [\mathbb{K}_i : \mathbb{Q}(c)][\mathbb{Q}(c) : \mathbb{Q}]$ , so  $[\mathbb{Q}(c) : \mathbb{Q}]$  is an even power.

**Examples.** The numbers  $\sqrt[3]{2}$  and  $\cos 20^\circ$  have the minimal polynomials  $x^3 - 2 \in \mathbb{Q}[x]$  and  $x^3 - \frac{3}{4}x - \frac{1}{8} \in \mathbb{Q}[x]$ , respectively, so they are both algebraic of degree 3 over  $\mathbb{Q}$ . Therefore, neither number is constructible.

By the above theorem, no transcendental real number, like  $e$  or  $\pi$ , can be constructed with ruler and compass. Therefore, almost no real numbers are constructible. However, the rational numbers are dense in  $\mathbb{R}$  and are constructible, so we can use ruler and compass constructions to approximate any real number with arbitrary precision.

**The Classical Construction Problems:** We now have tools that allow us to solve four of the construction problems that the ancient Greeks spent a lot of effort in trying to solve; three of these problems were not solved until more than two millennia later, in 1837, by P. L. Wantzel.

**Squaring the Circle:** Given a circle, is it always possible to construct a square with the same area as the circle? The answer to this problem is clearly no, since the circle with radius 1 has area  $\pi$ , so the square to be constructed would have side length  $\sqrt{\pi}$ , which is not algebraic over  $\mathbb{Q}$  (note that we have not actually proved this last fact):

**Theorem.** One cannot always construct by compass and ruler a square with the same area as a given circle.

**Duplicating the Cube:** Is it possible to construct a cube with double the volume of a given cube? As above, the answer is clearly no, since the cube with volume 2 doubles the volume as the cube with side lengths 1 but has itself side lengths  $\sqrt[3]{2}$  which is not constructible:

**Theorem.** It is not always possible by compass and ruler to construct a cube with double the volume of a given cube.

**Trisecting the Angle:** Is it possible to trisect an arbitrary angle? Once again, the answer is clearly no, since the angle  $60^\circ$  trisected is  $20^\circ$  but  $\cos 20^\circ$  is not constructible:

**Theorem.** One cannot trisect all angles by compass and ruler.

**Regular Polygons:** is it possible to construct all regular polygons by compass and ruler? It is clear that a regular  $n$ -gon can be constructed if and only if  $\cos \frac{2\pi}{n} = \cos \frac{360^\circ}{n}$  is constructible.

**Example.** By an example on page 155,  $\cos \frac{2\pi}{5}$  is algebraic of degree 2 over  $\mathbb{Q}$  and is thus constructible, so the pentagon is constructible.

In general, the answer to this problem is no; for instance, the regular heptagon is not constructible, since  $\cos \frac{2\pi}{7}$  is algebraic of degree 3 over  $\mathbb{Q}$ . Gauss determined which regular polygons are constructible and which are not; we give his result here without proof:

**Theorem.** The regular  $n$ -gon is constructible if and only if  $n$  is not divided by any odd prime square and each odd prime factor of  $n$  is of the form  $2^{2^m} + 1$ .

**Example.** The regular 60-gon is constructible since  $60 = 2^2 \cdot 3 \cdot 5$  and since  $3 = 2^{2^0} + 1$  and  $5 = 2^{2^1} + 1$  do not divide  $n$  more than once each.

**Corollary.** For any prime  $p$ , the regular  $p$ -gon is constructible if and only if  $p = 2^{2^n} + 1$  for some integer  $n \in \mathbb{N}_0$ .

Primes of the form  $2^{2^m} + 1$  are known as **Fermat primes**, the first of which are 3, 5, 17, 257, and 65537. No other Fermat primes have been found, and some conjecture that these are the only Fermat primes.\*

As a 19-year old and before his result above, Gauss proved that a regular 17-gon could be constructed, and J. Erchinger found an actual construction of this polygon a few years later. The regular 257-gon was constructed in 1832 by F. J. Richelot, and J. G. Hermes spent 10 years on a 200-page construction of the regular 65537-gon, although doubts have recently been raised about the validity of his construction.

---

\* See [www.research.att.com/~njas/sequences/A019434](http://www.research.att.com/~njas/sequences/A019434).