Contents

Part I Algebra Homeworks

Part II Math 103B Lecture Notes

Part I

Algebra Homeworks

0

Math 103B Homework Problems

The problems are from *Contemporary Abstract Algebra* by Gallian, 6th Edition [2] unless otherwise noted. We will label the problems via Chapter.Number, so that 4.12 will refer to exercise 12 in Chapter 4.

0.1 Homework #1 (Due Thursday, April 2)

Hand in Problems:

- Chapter 12: #2, 6, 7, 18, 22, 26, 42, 44*, 46 (*Hint: consider $(-a)^n$).

Extra problems for practice (do not hand in):

- Chapter 12: # 14, 19, 29, 31, 37, 39, 41, 47

0.2 Homework #2 (Due Thursday, April 9)

Hand in Problems:

- Chapter 13: #4, 6, 13, 14, 16, 22, 44, 54
  Chapter 14: #2, 4, 10, 12

Extra problems for practice (do not hand in):

- Chapter 13, #5, 15, 21, 25, 29, 33, 39, 45, 53.
  Chapter 14, #1, 5, 7, **11,** 17, 19.

Hints:
13.54: Use Lagrange's theorem from last quarter.

0.3 Homework #3 (Due Thursday, April 16)

Hand in Problems:

- Chapter 14, #18, 40
  Chapter 15, #6, 12, 14, 16, 28, 30

Exercise 0.1. Let R be an integral domain with characteristic zero and $\varphi, \psi : \mathbb{Q} \to R$ be two ring homomorphisms such that $\varphi(1) = \psi(1)$. Show $\varphi(a) = \psi(a)$ for all $a \in \mathbb{Q}$. Thus φ is uniquely determined by its value on $1 \in \mathbb{Q}$.

Hints:
14.18: In other words, find all possible choices of a pair of ideals I and J of $\mathbb{Z}$ which satisfy, $\langle 35 \rangle \subsetneq J \subsetneq I$.
15.28: Consider the relationship mod 3.
Extra problems for practice (do not hand in):

- Chapter 14, #25, 27, 59
  Chapter 15, #11, 13

0.4 Homework #4 (Due Thursday, April 23)

Hand in Problems:

- Chapter 14, #24, 34, 37, 54
- Chapter 15, #20, 26, 36, 40.
- Chapter 16, #2, 4, 6, 14

Hints:

15.26: The 1$^{\text{st}}$ - isomorphism theorem for rings may help.
15.36. Use Exercise 0.1 to see that any homomorphism $\varphi : \mathbb{Q} \to \mathbb{Q}$ is determined uniquely by the value, $a = \varphi(1)$. Now use the multiplicativity property of φ to determine the allowed values for a.
15.40. Think about the first isomorphism theorem and make use of Exercise 25 of Chapter 14.
Extra problems for practice (do not hand in):

- Chapter 14, #29, 33
- Chapter 15, #39, 45, 51.

0.5 Homework #5 (Due Thursday, April 30)

Hand in Problems:

Exercise 0.2 (This problem is to be handed in!). Let R be a commutative ring with identity. Then R is a field iff R has no non-trivial proper ideals. (Recall that $I \subset R$ is the trivial ideal if $I = \{0\}$ and is a proper ideal if $I \subsetneq R$.)

- Chapter 14, #28, 32, 36, 52
- Chapter 15, #58, 60

Hints:
14.36: Let $\mathcal{S} := \{a + bi : a \in \mathbb{Z}_4 \text{ and } b \in \mathbb{Z}_2\}$. To count the number of element in $\mathbb{Z}[i] / \langle 2 + 2i \rangle$ you might show

$$\mathcal{S} \ni (a + bi) \xrightarrow{\psi} [a + bi] \in \mathbb{Z}[i] / \langle 2 + 2i \rangle$$

is a bijection.
14.52: One way is to prove that $\mathbb{Z}[i] / \langle 1 - i \rangle$ is isomorphic to $\mathbb{Z}_2$.
Extra problems for practice (do not hand in):

- Chapter 14, #31
- Chapter 15, #27, 29, 42, 62

0.6 Homework #6 (Due Thursday, May 7)

Hand in Problems:

- Chapter 16, #12, 18, 20, 24, 30, 36, 38, 48.
  Hint:

  16.20. Think about the roots of the polynomial $h = f - g$.
  Extra problems for practice (do not hand in):
- Chapter 16, #1, 11, 13, 15, 19, 41

0.7 Homework #7 (Due Thursday, May 14)

Hand in Problems:

- Chapter 17, #2, 4, 6, 8, 12, 14, 25 (17.4 is a special case of 17.25!)

Extra problems for practice (do not hand in):

- Chapter 17, #1, 3, 5, 7, 21

0.8 Homework #8 (Due Thursday, May 21)

Hand in all problems below:

- Chapter 17, #10a, c, e, l0 b & d, 32

Exercise 0.3. Prove Proposition 20.4.

Exercise 0.4. Let $\psi : R \to T$ be an **onto** ring homomorphism of commutative rings, R and T and $I := \ker(\psi)$.

1. Explain why $\bar{\psi} : R[x] \to T[x]$ is onto.
2. Show $\ker\left(\bar{\psi}\right) = I[x]$.
3. Use the first isomorphism theorem to conclude $R[x] / I[x]$ is isomorphic $T[x]$.

Exercise 0.5. Let R be a commutative ring and $I \subset R$ be an ideal. Use the results of Exercise 0.4 to show $R[x] / I[x]$ is isomorphic to $(R/I)[x]$.

Exercise 0.6. Use Exercise 0.5 to give another proof Exercise 16.38 on page 300 of the book. This proof should be very short and similar in spirit to the proof of Gauss' Lemma on p. 305 of the book.

Exercise 0.7. Prove Proposition 18.7.

Exercise 0.8. Prove Proposition 23.3

Exercise 0.9. Show $x^8 - 20/9 \in \mathbb{Q}[x]$ is irreducible over $\mathbb{Q}$. Conclude that $\sqrt[8]{20/9} \notin \mathbb{Q}$.

Hints:
17.10b & d. Try the mod p irreducibility test, using some small prime p. Follow the method of examples 7 and 8 in the book.
17.32. One (but not the only) possibility is to show $\mathbb{Z}[x] / \left\langle x^2 + 1 \right\rangle$ is isomorphic to a ring we have studied frequently. Then use Theorems 14.3 and 14.4 of the book.

0.9 Homework #9 (Due Thursday, May 28)

Hand in Problems:

- Chapter 18, #4, 12, 28, 30.
- Chapter 19, #8, 22, 26.

Exercise 0.10. Let $f(x) := 6x^3 - 14x^2 - 2x + 2 \in \mathbb{Z}[x]$. Factor $f(x)$ into irreducible polynomials; 1) over $\mathbb{R}$, 2) over $\mathbb{Q}$, and 3) over $\mathbb{Z}$. (The answers are different in each case.) **Hint:** first find the rational roots of $f(x)$.

Extra problems for practice (do not hand in):

- Chapter 18, #1, 5, 27
- Chapter19, #7, 13, 14.

0.10 Homework #10 (Due Thursday, June 4)

Hand in Problems:

- Chapter 20, #1, 3, 5
- Chapter 21, #10, 14

- Chapter 23, #14

Exercise 0.11. Let $g(x) = x^3 + 2x + 1 \in \mathbb{Q}[x]$. Let $\alpha \in \mathbb{C}$ be a root of g over the complex numbers (do not try to find α explicitly; you don't need to in order to do this problem.)

1. Show $g(x)$ is irreducible.
2. What is the minimal polynomial of α over $\mathbb{Q}$.
3. Explain how you know that
$$K = \mathbb{Q}[\alpha] = \mathbb{Q} + \mathbb{Q}\alpha + \mathbb{Q}\alpha^2 = \{a + b\alpha + c\alpha^2 : a, b, c \in \mathbb{Q}\}$$
is a subfield of $\mathbb{C}$.
4. Find explicit $a, b, c \in \mathbb{Q}$ such that $\alpha^{-1} = a + b\alpha + c\alpha^2$. (Note that some such expression must exist since K is a field.)
5. Find the degree, $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ of the field extension $\mathbb{Q} \subset \mathbb{Q}(\alpha)$. Recall that this is defined to be the dimension of $\mathbb{Q}(\alpha)$ as a vector space over $\mathbb{Q}$.

Hints:

20.1: The hint in the back is useless. One way to “describe the elements” is to find a basis for the given field as a $\mathbb{Q}$-vector space. (Hint: it has dimension 3.)

21.14: See Examples 28.17 and 28.18.

23.14: Use Exercise 17.25 to show the relevant polynomial is irreducible.

Part II

Math 103B Lecture Notes

1

Lecture 1

1.1 Definition of Rings and Examples

A ring will be a set of elements, R, with both an **addition** and **multiplication** operation satisfying a number of "natural" axioms.

Axiom 1.1 (Axioms for a ring) *Let R be a set with 2 binary operations called addition (written $a+b$) and multiplication (written ab). R is called a* ***ring*** *if for all $a, b, c \in R$ we have*

1. $(a+b)+c = a+(b+c)$
2. *There exists an element $0 \in R$ which is an identity for $+$.*
3. *There exists an element $-a \in R$ such that $a+(-a) = 0$.*
4. $a+b = b+a$.
5. $(ab)c = a(bc)$.
6. $a(b+c) = ab+ac$ *and* $(b+c)a = ba+bc$.

Items 1. – 4. are the axioms for an abelian group, $(R, +)$. Item 5. says multiplication is associative, and item 6. says that is both left and right distributive over addition. Thus we could have stated the definition of a ring more succinctly as follows.

Definition 1.2. *A* ***ring*** *R is a set with two binary operations "$+$" = addition and "$\cdot$"= multiplication, such that $(R, +)$ is an abelian group (with identity element we call 0), "$\cdot$" is an associative multiplication on R which is both left and right distributive over addition.*

Remark 1.3. The multiplication operation might not be commutative, i.e., $ab \neq ba$ for some $a, b \in R$. If we have $ab = ba$ for all $a, b \in R$, we say R is a **commutative ring**. Otherwise R is **noncommutative.**

Definition 1.4. *If there exists and element $1 \in R$ such that $a1 = 1a = a$ for all $a \in R$, then we call 1 the* ***identity element*** *of R [the book calls it the unity.]*

Most of the rings that we study in this course will have an identity element.

Lemma 1.5. *If R has an identity element 1, then 1 is unique. If an element $a \in R$ has a multiplicative inverse b, then b is unique, and we write $b = a^{-1}$.*

Proof. Use the same proof that we used for groups! I.e. $1 = 1 \cdot 1' = 1'$ and if b, b' are both inverses to a, then $b = b(ab') = (ba)b' = b'$. ■

Notation 1.6 (Subtraction) *In any ring R, for $a \in R$ we write the additive inverse of a as $(-a)$. So at $a + (-a) = (-a) + a = 0$ by definition. For any $a, b \in R$ we abbreviate $a + (-b)$ as $a - b$.*

Let us now give a number of examples of rings.

Example 1.7. Here are some examples of commutative rings that we are already familiar with.

1. $\mathbb{Z}$ = all integers with usual $+$ and $\cdot$.
2. $\mathbb{Q}$ = all $\frac{m}{n}$ such that $m, n \in \mathbb{Z}$ with $n \neq 0$, usual $+$ and $\cdot$. (We will generalize this later when we talk about "fields of fractions.")
3. $\mathbb{R}$ = reals, usual $+$ and $\cdot$.
4. $\mathbb{C}$ = all complex numbers, i.e. $\{a + ib : a, b \in \mathbb{R}\}$, usual $+$ and $\cdot$ operations. (We will explicitly verify this in Proposition 3.7 below.)

Example 1.8. $2\mathbb{Z} = \{\dots, -4, -2, 0, 2, 4, \dots\}$ is a ring without identity.

Example 1.9 (Integers modulo m). For $m \geq 2$, $\mathbb{Z}_m = \{0, 1, 2, \dots, m-1\}$ with

$$\begin{aligned} + &= \text{addition} \mod m \\ \cdot &= \text{multiplication} \mod n. \end{aligned}$$

Recall from last quarter that $(\mathbb{Z}_m, +)$ is an abelian group and we showed,

$$[(ab) \bmod m \cdot c] \bmod m = [abc] = [a(bc) \bmod m] \bmod m \quad \text{(associativity)}$$

and

$$\begin{aligned} [a \cdot (b+c) \bmod m] \bmod m &= [a \cdot (b+c)] \bmod m \\ &= [ab + ac] \bmod m = (ab) \bmod m + (ac) \bmod m \end{aligned}$$

which is the distributive property of multiplication $\bmod m$. Thus $\mathbb{Z}_m$ is a ring with identity, 1.

Example 1.10. $M_2(F) = 2\times 2$ matrices with entries from F, where $F = \mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$ with binary operations;

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} a+a' & b+b' \\ c+c' & d+d' \end{bmatrix} \text{ (addition)}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{bmatrix}. \text{ (multiplication)}$$

That is multiplication is the usual matrix product. You should have checked in your linear algebra course that $M_2(F)$ is a non-commutative ring with identity,

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

For example let us check that left distributive law in $M_2(\mathbb{Z})$;

$$\begin{aligned} &\begin{bmatrix} a & b \\ c & d \end{bmatrix} \left(\begin{bmatrix} e & f \\ g & h \end{bmatrix} + \begin{bmatrix} p & q \\ r & s \end{bmatrix} \right) \\ &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} p+e & f+q \\ g+r & h+s \end{bmatrix} \\ &= \begin{bmatrix} b(g+r)+a(p+e) & a(f+q)+b(h+s) \\ d(g+r)+c(p+e) & c(f+q)+d(h+s) \end{bmatrix} \\ &= \begin{bmatrix} bg+ap+br+ae & af+bh+aq+bs \\ dg+cp+dr+ce & cf+dh+cq+ds \end{bmatrix} \end{aligned}$$

while

$$\begin{aligned} &\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} p & q \\ r & s \end{bmatrix} \\ &= \begin{bmatrix} bg+ae & af+bh \\ dg+ce & cf+dh \end{bmatrix} + \begin{bmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{bmatrix} \\ &= \begin{bmatrix} bg+ap+br+ae & af+bh+aq+bs \\ dg+cp+dr+ce & cf+dh+cq+ds \end{bmatrix} \end{aligned}$$

which is the same result as the previous equation.

Example 1.11. We may realize $\mathbb{C}$ as a sub-ring of $M_2(\mathbb{R})$ as follows. Let

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in M_2(\mathbb{R}) \text{ and } \mathbf{i} := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

and then identify $z = a + ib$ with

$$aI + b\mathbf{i} := a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Since

$$\mathbf{i}^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = I$$

it is straight forward to check that

$$\begin{aligned} (aI+b\mathbf{i})(cI+d\mathbf{i}) &= (ac-bd)\,I + (bc+ad)\,\mathbf{i} \text{ and} \\ (aI+b\mathbf{i}) + (cI+d\mathbf{i}) &= (a+c)\,I + (b+d)\,\mathbf{i} \end{aligned}$$

which are the standard rules of complex arithmetic. The fact that $\mathbb{C}$ is a ring now easily follows from the fact that $M_2(\mathbb{R})$ is a ring.

In this last example, the reader may wonder how did we come up with the matrix $\mathbf{i} := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ to represent i. The answer is as follows. If we view $\mathbb{C}$ as $\mathbb{R}^2$ in disguise, then multiplication by i on $\mathbb{C}$ becomes,

$$(a,b) \sim a + ib \to i(a+ib) = -b + ai \sim (-b, a)$$

while

$$\mathbf{i} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix}.$$

Thus $\mathbf{i}$ is the 2×2 real matrix which implements multiplication by i on $\mathbb{C}$.

Theorem 1.12 (Matrix Rings). *Suppose that R is a ring and $n \in \mathbb{Z}_+$. Let $M_n(R)$ denote the $n\times n$ – matrices $A = (A_{ij})_{i,j=1}^n$ with entries from R. Then $M_n(R)$ is a ring using the addition and multiplication operations given by,*

$$\begin{aligned} (A+B)_{ij} &= A_{ij} + B_{ij} \text{ and} \\ (AB)_{ij} &= \sum_k A_{ik} B_{kj}. \end{aligned}$$

Moreover if $1 \in R$, then

$$I := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is the identity of $M_n(R)$.

Proof. I will only check associativity and left distributivity of multiplication here. The rest of the proof is similar if not easier. In doing this we will make use of the results about sums in the Appendix 1.2 at the end of this lecture.

Let A, B, and C be $n\times n$ – matrices with entries from R. Then

$$
\begin{aligned}
[A(BC)]_{ij} &= \sum_k A_{ik}(BC)_{kj} = \sum_k A_{ik}\left(\sum_l B_{kl}C_{lj}\right) \\
&= \sum_{k,l} A_{ik}B_{kl}C_{lj}
\end{aligned}
$$

while

$$
\begin{aligned}
[(AB)C]_{ij} &= \sum_l (AB)_{il}C_{lj} = \sum_l\left(\sum_k A_{ik}B_{kl}\right)C_{lj} \\
&= \sum_{k,l} A_{ik}B_{kl}C_{lj}.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
[A(B+C)]_{ij} &= \sum_k A_{ik}(B_{kj}+C_{kj}) = \sum_k (A_{ik}B_{kj}+A_{ik}C_{kj}) \\
&= \sum_k A_{ik}B_{kj} + \sum_k A_{ik}C_{kj} = [AB]_{ij} + [AC]_{ij}.
\end{aligned}
$$

■

Example 1.13. In $\mathbb{Z}_6$, 1 is an identity for multiplication, but 2 has no multiplicative inverse. While in $M_2(\mathbb{R})$, a matrix A has a multiplicative inverse if and only if $\det(A) \neq 0$.

Example 1.14 (Another ring without identity). Let

$$R = \left\{\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix} : a \in \mathbb{R}\right\}$$

with the usual addition and multiplication of matrices.

$$\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The identity element for multiplication "wants" to be $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, but this is not in R.

More generally if $(R,+)$ is any abelian group, we may make it into a ring in a trivial way by setting $ab = 0$ for all $a, b \in R$. This ring clearly has no multiplicative identity unless $R = \{0\}$ is the trivial group.

1.2 Appendix: Facts about finite sums

Throughout this section, suppose that $(R,+)$ is an abelian group, Λ is any set, and $\Lambda \ni \lambda \to r_\lambda \in R$ is a given function.

Theorem 1.15. *Let $\mathcal{F} := \{A \subset \Lambda : |A| < \infty\}$. Then there is a unique function, $S : \mathcal{F} \to R$ such that;*

1. *$S(\emptyset) = 0$,*
2. *$S(\{\lambda\}) = r_\lambda$ for all $\lambda \in \Lambda$.*
3. *$S(A \cup B) = S(A) + S(B)$ for all $A, B \in \mathcal{F}$ with $A \cap B = \emptyset$.*

Moreover, for any $A \in \mathcal{F}$, $S(A)$ only depends on $\{r_\lambda\}_{\lambda \in A}$.

Proof. Suppose that $n \geq 2$ and that $S(A)$ has been defined for all $A \in \mathcal{F}$ with $|A| < n$ in such a way that S satisfies items 1. – 3. provided that $|A \cup B| < n$. Then if $|A| = n$ and $\lambda \in A$, we must define,

$$S(A) = S(A \setminus \{\lambda\}) + S(\{\lambda\}) = S(A \setminus \{\lambda\}) + r_\lambda.$$

We should verify that this definition is independent of the choice of $\lambda \in A$. To see this is the case, suppose that $\lambda' \in A$ with $\lambda' \neq \lambda$, then by the induction hypothesis we know,

$$
\begin{aligned}
S(A \setminus \{\lambda\}) &= S([A \setminus \{\lambda,\lambda'\}] \cup \{\lambda'\}) \\
&= S(A \setminus \{\lambda,\lambda'\}) + S(\{\lambda'\}) = S(A \setminus \{\lambda,\lambda'\}) + r_{\lambda'}
\end{aligned}
$$

so that

$$
\begin{aligned}
S(A \setminus \{\lambda\}) + r_\lambda &= [S(A \setminus \{\lambda,\lambda'\}) + r_{\lambda'}] + r_\lambda \\
&= S(A \setminus \{\lambda,\lambda'\}) + (r_{\lambda'} + r_\lambda) \\
&= S(A \setminus \{\lambda,\lambda'\}) + (r_\lambda + r_{\lambda'}) \\
&= [S(A \setminus \{\lambda,\lambda'\}) + r_\lambda] + r_{\lambda'} \\
&= [S(A \setminus \{\lambda,\lambda'\}) + S(\{\lambda\})] + r_{\lambda'} \\
&= S(A \setminus \{\lambda'\}) + r_{\lambda'}
\end{aligned}
$$

as desired. Notice that the "moreover" statement follows inductively using this definition.

Now suppose that $A, B \in \mathcal{F}$ with $A \cap B = \emptyset$ and $|A \cup B| = n$. Without loss of generality we may assume that neither A or B is empty. Then for any $\lambda \in B$, we have used the inductive hypothesis, that

$$
\begin{aligned}
S(A \cup B) &= S(A \cup [B \setminus \{\lambda\}]) + r_\lambda = (S(A) + S(B \setminus \{\lambda\})) + r_\lambda \\
&= S(A) + (S(B \setminus \{\lambda\}) + r_\lambda) = S(A) + (S(B \setminus \{\lambda\}) + S(\{\lambda\})) \\
&= S(A) + S(B).
\end{aligned}
$$

Thus we have defined S inductively on the size of $A \in \mathcal{F}$ and we had no choice in how to define S showing S is unique. ∎

Notation 1.16 *Keeping the notation used in Theorem 1.15, we will denote $S(A)$ by $\sum_{\lambda \in A} r_\lambda$. If $A = \{1, 2, \dots, n\}$ we will often write,*

$$\sum_{\lambda \in A} r_\lambda = \sum_{i=1}^{n} r_i.$$

Corollary 1.17. *Suppose that $A = A_1 \cup \dots \cup A_n$ with $A_i \cap A_j = \emptyset$ for $i \neq j$ and $|A| < \infty$. Then*

$$S(A) = \sum_{i=1}^{n} S(A_i) \;\; i.e. \;\; \sum_{\lambda \in A} r_\lambda = \sum_{i=1}^{n} \left(\sum_{\lambda \in A_i} r_\lambda \right).$$

Proof. As usual the proof goes by induction on n. For $n = 2$, the assertion is one of the defining properties of $S(A) := \sum_{\lambda \in A} r_\lambda$. For $n \geq 2$, we have used the induction hypothesis and the definition of $\sum_{i=1}^{n} S(A_i)$ that

$$\begin{aligned} S(A_1 \cup \dots \cup A_n) &= S(A_1 \cup \dots \cup A_{n-1}) + S(A_n) \\ &= \sum_{i=1}^{n-1} S(A_i) + S(A_n) = \sum_{i=1}^{n} S(A_i). \end{aligned}$$

∎

Corollary 1.18 (Order does not matter). *Suppose that A is a finite subset of Λ and B is another set such that $|B| = n = |A|$ and $\sigma : B \to A$ is a bijective function. Then*

$$\sum_{b \in B} r_{\sigma(b)} = \sum_{a \in A} r_a.$$

In particular if $\sigma : A \to A$ is a bijection, then

$$\sum_{a \in A} r_{\sigma(a)} = \sum_{a \in A} r_a.$$

Proof. We again check this by induction on $n = |A|$. If $n = 1$, then $B = \{b\}$ and $A = \{a := \sigma(b)\}$, so that

$$\sum_{x \in B} r_{\sigma(x)} = r_{\sigma(b)} = \sum_{a \in A} r_a$$

as desired. Now suppose that $N \geq 1$ and the corollary holds whenever $n \leq N$. If $|B| = N + 1 = |A|$ and $\sigma : B \to A$ is a bijective function, then for any $b \in B$, we have with $B' := B' \setminus \{b\}$ that

$$\sum_{x \in B} r_{\sigma(x)} = \sum_{x \in B'} r_{\sigma(x)} + r_{\sigma(b)}.$$

Since $\sigma|_{B'} : B' \to A' := A \setminus \{\sigma(b)\}$ is a bijection, it follows by the induction hypothesis that $\sum_{x \in B'} r_{\sigma(x)} = \sum_{\lambda \in A'} r_\lambda$ and therefore,

$$\sum_{x \in B} r_{\sigma(x)} = \sum_{\lambda \in A'} r_\lambda + r_{\sigma(b)} = \sum_{\lambda \in A} r_\lambda.$$

∎

Lemma 1.19. *If $\{a_\lambda\}_{\lambda \in A}$ and $\{b_\lambda\}_{\lambda \in A}$ are two sequences in R, then*

$$\sum_{\lambda \in A} (a_\lambda + b_\lambda) = \sum_{\lambda \in A} a_\lambda + \sum_{\lambda \in A} b_\lambda.$$

Moreover, if we further assume that R is a ring, then for all $r \in R$ we have the right and left distributive laws;,

$$\begin{aligned} r \cdot \sum_{\lambda \in A} a_\lambda &= \sum_{\lambda \in A} r \cdot a_\lambda \;\; \textit{and} \\ \left(\sum_{\lambda \in A} a_\lambda \right) \cdot r &= \sum_{\lambda \in A} a_\lambda \cdot r. \end{aligned}$$

Proof. This follows by induction. Here is the key step. Suppose that $\alpha \in A$ and $A' := A \setminus \{\alpha\}$, then

$$\begin{aligned} \sum_{\lambda \in A} (a_\lambda + b_\lambda) &= \sum_{\lambda \in A'} (a_\lambda + b_\lambda) + (a_\alpha + b_\alpha) \\ &= \sum_{\lambda \in A'} a_\lambda + \sum_{\lambda \in A'} b_\lambda + (a_\alpha + b_\alpha) \quad \text{(by induction)} \\ &= \left(\sum_{\lambda \in A'} a_\lambda + a_\lambda + \right) \left(\sum_{\lambda \in A'} b_\lambda + b_\alpha \right) \quad \begin{pmatrix} \text{commutativity} \\ \text{and associativity} \end{pmatrix} \\ &= \sum_{\lambda \in A} a_\lambda + \sum_{\lambda \in A} b_\lambda. \end{aligned}$$

The multiplicative assertions follows by induction as well,

$$\begin{aligned} r \cdot \sum_{\lambda \in A} a_\lambda = r \cdot \left(\sum_{\lambda \in A'} a_\lambda + a_\alpha \right) &= r \cdot \left(\sum_{\lambda \in A'} a_\lambda \right) + r \cdot a_\alpha \\ &= \left(\sum_{\lambda \in A'} r \cdot a_\lambda \right) + r \cdot a_\alpha \\ &= \sum_{\lambda \in A} r \cdot a_\lambda. \end{aligned}$$

∎

2

Lecture 2

Recall that a ring is a set, R, with two binary operations "+" = addition and "$\cdot$"= multiplication, such that $(R,+)$ is an abelian group (with identity element we call 0), $(\cdot)$ is an associative multiplication on R which is left and right distributive over "+." Also recall that if there is a multiplicative identity, $1 \in R$ (so $1a = a1 = a$ for all a), we say R is a ring with identity (unity). Furthermore we write $a - b$ for $a + (-b)$. This shows the importance of distributivity. We now continue with giving more examples of rings.

Example 2.1. Let R denote the continuous functions, $f : \mathbb{R} \to \mathbb{R}$ such that $\lim_{x \to \pm\infty} f(x) = 0$. As usual, let $f + g$ and $f \cdot g$ be pointwise addition and multiplication of functions, i.e.

$$(f+g)(x) = f(x) + g(x) \text{ and } (f \cdot g)(x) = f(x)\, g(x) \text{ for all } x \in \mathbb{R}.$$

Then R is a ring without identity. (If we remove the restrictions on the functions at infinity, R would be a ring with identity, namely $\mathbf{1}(x) \equiv 1$.)

Example 2.2. For any collection of rings $R_1, R_2, \ldots, R_m$, define the direct sum to be

$$R = R_1 \oplus \cdots \oplus R_n = \{(r_1, r_2, \ldots, r_n) : r_i \in R_i \text{ all } i\}$$

the set of all m-tuples where the i^{th} coordinate comes from R_i. R is a ring if we define

$$(r_1, r_2, \ldots, r_m) + (s_1, s_2, \ldots, s_m) = (r_1 s_1, r_2 s_2, \ldots, r_m s_m),$$

and

$$(r_1, r_2, \ldots, r_m) + (s_1, s_2, \ldots, s_m) = (r_1 + s_1, r_2 + s_2, \ldots, r_m + s_m).$$

The identity element 0 is $(0, 0, \ldots, 0)$. (Easy to check)

2.1 Polynomial Ring Examples

Example 2.3 (Polynomial rings). Let $R = \mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{Z}$ and let $R[x]$ denote the polynomials in x with coefficients from R. We add and multiply polynomials in the usual way. For example if $f = 3x^2 - 2x + 5$ and $g = 5x^2 + 1$, then

$$\begin{aligned} f + g &= 8x^2 - 2x + 6 \text{ and} \\ fg &= (5x^3 + 1)(3x^2 - 2x + 5) \\ &= 5 - 2x + 3x^2 + 25x^3 - 10x^4 + 15x^5. \end{aligned}$$

One may check (see Theorem 2.4 below) that $R[x]$ with these operations is a commutative ring with identity, $\mathbf{1} = 1$.These rules have been chosen so that $(f+g)(\alpha) = f(\alpha) + g(\alpha)$ and $(f \cdot g)(\alpha) = f(\alpha)\, g(\alpha)$ for all $\alpha \in R$ where

$$f(\alpha) := \sum_{i=0}^{\infty} a_i \alpha^i.$$

Theorem 2.4. *Let R be a ring and $R[x]$ denote the collection of polynomials with the usual addition and multiplication rules of polynomials. Then $R[x]$ is again a ring. To be more precise,*

$$R[x] = \left\{ p = \sum_{i=0}^{\infty} p_i x^i : p_i \in R \text{ with } p_i = 0 \text{ a.a.} \right\},$$

where we say that $p_i = 0$ a.a. (read as almost always) provided that $|\{i : p_i \neq 0\}| < \infty$. If $q := \sum_{i=0}^{\infty} q_i x^i \in R[x]$, then we set,

$$p + q := \sum_{i=0}^{\infty} (p_i + q_i)\, x^i \text{ and} \tag{2.1}$$

$$p \cdot q := \sum_{i=0}^{\infty} \left(\sum_{k+l=i} p_k q_l \right) x^i = \sum_{i=0}^{\infty} \left(\sum_{k=0}^{i} p_k q_{i-k} \right) x^i. \tag{2.2}$$

Proof. The proof is similar to the matrix group examples. Let me only say a few words about the associativity property of multiplication here, since this is the most complicated property to check. Suppose that $r = \sum_{i=0}^{\infty} r_i x^i$, then

$$\begin{aligned} p\left(qr\right) &= \sum_{n=0}^{\infty}\left(\sum_{i+j=n} p_i\left(qr\right)_j\right)x^n \\ &= \sum_{n=0}^{\infty}\left(\sum_{i+j=n} p_i\left(\sum_{k+l=j} q_k r_l\right)\right)x^n \\ &= \sum_{n=0}^{\infty}\left(\sum_{i+k+l=n} p_i q_k r_l\right)x^n. \end{aligned}$$

As similar computation shows,

$$\left(pq\right)r = \sum_{n=0}^{\infty}\left(\sum_{i+k+l=n} p_i q_k r_l\right)x^n$$

and hence the multiplication rule in Eq. (2.2) is associative. ■

2.2 Subrings and Ideals I

We now define the concept of a subring in a way similar to the concept of subgroup.

Definition 2.5 (Subring). *Let R be a ring. If S is subset of R which is itself a ring under the same operations $+,\cdot$ of R restricted to the set S, then S is called a **subring** of R.*

Lemma 2.6 (Subring test). *$S \subset R$ is a subring if and only if S is a subgroup of $(R,+)$ and S is closed under multiplication. In more detail, S is a subring of R, iff for all $a,b \in S$, that*

$$a+b \in S,\ \ -a \in S,\ \text{ and } ab \in S.$$

Alternatively we may check that

$$a-b \in S,\ \text{ and } ab \in S \text{ for all } a,b \in S.$$

Put one last way, S is a subring of R if $(S,+)$ is a subgroup of $(R,+)$ which is closed under the multiplication operation, i.e. $S \cdot S \subset S$.

Proof. Either of the conditions, $a+b \in S$, $-a \in S$ or $a-b \in S$ for all $a,b \in S$ implies that $(S,+)$ is a subgroup of $(R,+)$. The condition that $(S,\cdot)$ is a closed shows that "$\cdot$" is well defined on S. This multiplication on S then inherits the associativity and distributivity laws from those on R. ■

Definition 2.7 (Ideals). *Let R be a ring. A (two sided) ideal, I, of R is a subring, $I \subset R$ such that $RI \subset R$ and $IR \subset R$. Alternatively put, $I \subset R$ is an ideal if $(I,+)$ is a subgroup of $(R,+)$ such that $RI \subset R$ and $IR \subset R$. (Notice that every ideal, I, of R is also a subring of R.)*

Example 2.8. Suppose that R is a ring with identity 1 and I is an ideal. If $1 \in I$, then $I = R$ since $R = R \cdot 1 \subset RI \subset I$.

Example 2.9. Given a ring R, R itself and $\{0\}$ are always ideals of R. $\{0\}$ is the trivial ideal. An ideal (subring) $I \subset R$ for which $I \neq R$ is called a proper ideal (subring).

Example 2.10. If R is a commutative ring and $b \in R$ is any element, then the **principle ideal generated by** b, denoted by $\langle b\rangle$ or Rb, is

$$I = Rb = \{rb : r \in R\}.$$

To see that I is an ideal observer that if $r,s \in R$, then rb and sb are generic elements of I and

$$rb - sb = (r-s)b \in Rb.$$

Therefore I is an additive subgroup of R. Moreover, $(rb)\,s = s\,(rb) = (sr)\,b \in I$ so that $RI = IR \subset I$.

Theorem 2.11. *Suppose that $R = \mathbb{Z}$ or $R = \mathbb{Z}_m$ for some $m \in \mathbb{Z}_+$. Then the subgroups of $(R,+)$ are the same as the subrings of R which are the same as the ideals of R. Moreover, every ideal of R is a principle ideal.*

Proof. If $R = \mathbb{Z}$, then $\langle m\rangle = m\mathbb{Z}$ inside of $\mathbb{Z}$ is the principle ideal generated by m. Since every subring, $S \subset \mathbb{Z}$ is also a subgroup and all subgroups of $\mathbb{Z}$ are of the form $m\mathbb{Z}$ for some $m \in \mathbb{Z}$, it flows that all subgroups of $(\mathbb{Z},+)$ are in fact also principle ideals.

Suppose now that $R = \mathbb{Z}_n$. Then again for any $m \in \mathbb{Z}_n$,

$$\langle m\rangle = \{km : k \in \mathbb{Z}\} = m\mathbb{Z}_n \tag{2.3}$$

is the principle ideal in $\mathbb{Z}_n$ generated by m. Conversely if $S \subset \mathbb{Z}_n$ is a sub-ring, then S is in particular a subgroup of $\mathbb{Z}_n$. From last quarter we know that this implies $S = \langle m\rangle = \langle \gcd(n,m)\rangle$ for some $m \in \mathbb{Z}_n$. Thus every subgroup of $(\mathbb{Z}_n,+)$ is a principle ideal as in Eq. (2.3). ■

Example 2.12. The set,

$$S = \left\{\begin{bmatrix} a & b \\ 0 & d\end{bmatrix} : a,b,d \in \mathbb{R}\right\},$$

is a subring of $M_2(\mathbb{R})$. To check this observe that;

$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} - \begin{bmatrix} a' & b' \\ 0 & d' \end{bmatrix} = \begin{bmatrix} a - a & b - b' \\ 0 & d - d' \end{bmatrix} \in S$$

and

$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} a' & b' \\ 0 & d' \end{bmatrix} = \begin{bmatrix} a'a & ab' + bd' \\ 0 & dd' \end{bmatrix} \in S.$$

S is not an ideal since,

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ a & b \end{bmatrix} \notin S \text{ if } a \neq 0.$$

Example 2.13. Consider $\mathbb{Z}_m$ and the subset $U(m)$ the set of units in $\mathbb{Z}_m$. Then $U(m)$ is never a subring of $\mathbb{Z}_m$, because $0 \notin U(m)$.

Example 2.14. The collection of matrices,

$$S = \left\{ \begin{bmatrix} 0 & a \\ b & c \end{bmatrix} : a, b, c \in \mathbb{R} \right\},$$

is not a subring of $M_2(\mathbb{R})$. It is an additive subgroup which is however not closed under matrix multiplication;

$$\begin{bmatrix} 0 & a \\ b & c \end{bmatrix} \begin{bmatrix} 0 & a' \\ b' & c' \end{bmatrix} = \begin{bmatrix} ab' & ac' \\ cb' & ba + cc' \end{bmatrix} \notin S$$

Definition 2.15. *Let R be a ring with identity. We say that $S \subset R$ is a* ***unital subring*** *of R if S is a sub-ring containing 1_R. (Most of the subrings we will consider later will be unital.)*

Example 2.16. Here are some examples of unital sub-rings.

1. S in Example 2.12 is a unital sub-ring of $M_2(\mathbb{R})$.
2. The polynomial functions on $\mathbb{R}$ is a unital sub-ring of the continuous functions on $\mathbb{R}$.
3. $\mathbb{Z}[x]$ is a unital sub-ring of $\mathbb{Q}[x]$ or $\mathbb{R}[x]$ or $\mathbb{C}[x]$.
4. The **Gaussian integres**, $\mathbb{Z}[i] := \{a + ib : a, b \in \mathbb{Z}\}$ is a unital subring of $\mathbb{C}$. (For some number theoretic applications of the Gaussian integers see [1, Sections 12.3, p. 364 – 371.].)

Example 2.17. Here are a few examples of non-unital sub-rings.

1. $n\mathbb{Z} \subset \mathbb{Z}$ is a non-unital subring of $\mathbb{Z}$ for all $n \neq 0$ since $n\mathbb{Z}$ does not even contain an identity element.
2. If $R = \mathbb{Z}_8$, then every non-trivial proper subring, $S = \langle m \rangle$, of R has no identity. The point is if $k \in \mathbb{Z}_8$ is going to be an identity for some sub-ring of $\mathbb{Z}_8$, then $k^2 = k$. It is now simple to check that $k^2 = k$ in $\mathbb{Z}_8$ iff $k = 0$ or 1 which are not contained in any proper non-trivial sub-ring of $\mathbb{Z}_8$. (See Remark 2.18 below.)
3. Let $R := \mathbb{Z}_6$ and $S = \langle 2 \rangle = \{0, 2, 4\}$ is a sub-ring of $\mathbb{Z}_6$. Moreover, one sees that $1_S = 4$ is the unit in S ($4^2 = 4$ and $4 \cdot 2 = 2$) which is not $1_R = 1$. Thus again, S is not a unital sub-ring of $\mathbb{Z}_6$.
4. The set,
$$S = \left\{ \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} : a \in \mathbb{R} \right\} \subset R = M_2(\mathbb{R}),$$
is a subring of $M_2(\mathbb{R})$ with
$$1_S = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1_R$$
and hence is not a unital subring of $M_2(\mathbb{R})$.
5. Let v be a non-zero column vector in $\mathbb{R}^2$ and define,
$$S := \{A \in M_2(\mathbb{R}) : Av = 0\}.$$
Then S is a non-unital subring of $M_2(\mathbb{R})$ which is not an ideal. (You should verify these assertions yourself!)

Remark 2.18. Let $n \in \mathbb{Z}_+$ and $S := \langle m \rangle$ be a sub-ring of $\mathbb{Z}_n$. It is natural to ask, when does S have an identity element. To answer this question, we begin by looking for $m \in \mathbb{Z}_n$ such that $m^2 = m$. Given such a m, we claim that m is an identity for $\langle m \rangle$ since

$$(km)\, m = km^2 = k_1 m \text{ for all } km \in \langle m \rangle.$$

The condition that $m^2 = m$ is equivalent to $m(m-1) = 0$, i.e. $n|m(m-1)$. Thus $\langle m \rangle = \langle \gcd(n, m) \rangle$ is a ring with identity iff $n|m(m-1)$.

Example 2.19. Let us take $m = 6$ in the above remark so that $m(m-1) = 30 = 3 \cdot 2 \cdot 5$. In this case 10, 15 and 30 all divide $m(m-1)$ and therefore 6 is the identity element in $\langle 6 \rangle$ thought of as a subring of either, $\mathbb{Z}_{10}$, or $\mathbb{Z}_{15}$, or $\mathbb{Z}_{30}$. More explicitly 6 is the identity in

$$\begin{aligned} \langle 6 \rangle &= \langle \gcd(6, 10) \rangle = \langle 2 \rangle = \{0, 2, 4, 6, 8\} \subset \mathbb{Z}_{10}, \\ \langle 6 \rangle &= \langle \gcd(6, 15) \rangle = \langle 3 \rangle = \{0, 3, 6, 9, 12\} \subset \mathbb{Z}_{15}, \text{ and} \\ \langle 6 \rangle &= \langle \gcd(6, 30) \rangle = \{0, 6, 12, 18, 24\} \subset \mathbb{Z}_{30}. \end{aligned}$$

Example 2.20. On the other hand there is no proper non-trivial subring of $\mathbb{Z}_8$ which contains an identity element. Indeed, if $m \in \mathbb{Z}_8$ and $8 = 2^3|m(m-1)$, then either $2^3|m$ if m is even or $2^3|(m-1)$ if m is odd. In either the only $m \in \mathbb{Z}_8$ with this property is $m = 0$ and $m = 1$. In the first case $\langle 0 \rangle = \{0\}$ is the trivial subring of $\mathbb{Z}_8$ and in the second case $\langle 1 \rangle = \mathbb{Z}_8$ is not proper.

3

Lecture 3

3.1 Some simple ring facts

The next lemma shows that the distributive laws force 0, 1, and the symbol "$-$" to behave in familiar ways.

Lemma 3.1 (Some basic properties of rings). *Let R be a ring. Then;*

1. $a0 = 0 = 0a$ *for all* $a \in R$.
2. $(-a)b = -(ab) = a(-b)$ *for all* $a, b \in R$
3. $(-a)(-b) = ab$ *for all* $a, b \in R$. *In particular, if R has identity* 1, *then*
$$\begin{aligned}(-1)(-1) &= 1 \text{ and} \\ (-1)a &= -a \text{ for all } a \in R.\end{aligned}$$
(This explains why minus times minus is a plus! It has to be true in any structure with additive inverses and distributivity.)
4. *If* $a, b, c \in R$, *then* $a(b-c) = ab - ac$ *and* $(b-c)a = ba - ca$.

Proof. For all $a, b \in R$;

1. $a0 + 0 = a0 = a(0+0) = a0 + a0$, and hence by cancellation in the abelian group, $(R, +)$, we conclude that , so $0 = a0$. Similarly one shows $0 = 0a$.
2. $(-a)b + ab = (-a+a)b = 0b = 0$, so $(-a)b = -(ab)$. Similarly $a(-b) = -ab$.
3. $(-a)(-b) = -(a(-b)) = -(-(ab)) = ab$, where in the last equality we have used the inverting an element in a group twice gives the element back.
4. This last item is simple since,
$$a(b-c) := a(b + (-c)) = ab + a(-c) = ab + (-ac) = ab - ac.$$
Similarly one shows that $(b-c)a = ba - ca$.

■

In proofs above the reader should not be fooled into thinking these things are obvious. The elements involved are not necessarily familiar things like real numbers. For example, in $M_2(\mathbb{R})$ item 2 states, $(-I)A = -(IA) = -A$, i.e.

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix} \checkmark$$

The following example should help to illustrate the significance of Lemma 3.1.

Example 3.2. Consider $R = \langle 2 \rangle = \{0, 2, 4, 6, 8\} \subset \mathbb{Z}_{10}$. From Example 2.19 we know that $1_R = 6$ which you can check directly as well. So $-1_R = -6 \bmod 10 = 4$. Taking $a = 2$ let us write out the meaning of the identity, $(-1_R) \cdot a = -a$;

$$(-1_R) \cdot a = 4 \cdot 2 = 8 = -a.$$

Let us also work out $(-2)(-4)$ and compare this with $2 \cdot 4 = 8$;

$$(-2)(-4) = 8 \cdot 6 = 48 \bmod 10 = 8.$$

Lastly consider,

$$\begin{aligned} 4 \cdot (8-2) &= 4 \cdot 6 = 24 \bmod 10 = 4 \text{ while} \\ 4 \cdot 8 - 4 \cdot 2 &= 2 - 8 = -6 \bmod 10 = 4. \end{aligned}$$

3.2 The $R[S]$ subrings I

Here we will construct some more examples of rings which are closely related to polynomial rings. In these examples, we will be given a commutative ring R (usually commutative) and a set S equipped with some sort of multiplication, we then are going to define $R[S]$ to be the collection of linear combinations of elements from the set, $\cup_{n=0}^{\infty} RS^n$. Here RS^n consists of formal symbols of the form $rs_1 \dots s_n$ with $r \in R$ and $s_i \in S$. The next proposition gives a typical example of what we have in mind.

A typical case will be where $S = \{s_1, \dots, s_n\}$ is a finite set then

Proposition 3.3. *If $R \subset \bar{R}$ is a sub-ring of a commutative ring $\bar{R}$ and $S = \{s_1, \dots, s_n\} \subset \bar{R}$. Let*

$$R[S] = R[s_1, \dots, s_n] = \left\{ \sum_k a_k s^k : a_k \in R \text{ with } a_k = 0 \text{ a.a.} \right\},$$

where $k = (k_1, \dots k_n) \in \mathbb{N}^n$ and $s^k = s_1^{k_1} \dots s_n^{k_n}$ with $a_0 s^0 := a_0 \in R$. Then $R[s_1, \dots, s_n]$ is a sub-ring of $\bar{R}$.

Proof. If $f = \sum_k a_k s^k$ and $g = \sum_k b_k s^k$, then

$$\begin{aligned} f + g &= \sum_k (a_k + b_k) s^k \in R[S], \\ -g &= \sum_k -b_k s^k \in R[S], \text{ and} \\ f \cdot g &= \sum_k a_k s^k \cdot \sum_l b_l s^l \\ &= \sum_{k,l} a_k b_l s^k s^l = \sum_{k,l} a_k b_l s^{k+l} \\ &= \sum_n \left(\sum_{k+l=n} a_k b_l \right) s^n \in R[S]. \end{aligned}$$

■

Example 3.4 (Gaussian Integers). Let $i := \sqrt{-1} \in \mathbb{C}$. Then $\mathbb{Z}[i] = \{x + yi : x, y \in \mathbb{Z}\}$. To see this notice that $i^2 = -1 \in \mathbb{Z}$, and therefore

$$\begin{aligned} \sum_{k=0}^{\infty} a_k (i)^k &= \sum_{l=0}^{\infty} \left[a_{4l} (i)^{4l} + a_{4l+1} (i)^{4l+1} + a_{4l+2} (i)^{4l+2} + a_{4l+3} (i)^{4l+3} \right] \\ &= \sum_{l=0}^{\infty} [a_{4l} + a_{4l+1} i - a_{4l+2} - a_{4l+3} i] \\ &= \sum_{l=0}^{\infty} [a_{4l} - a_{4l+2}] + \left(\sum_{l=0}^{\infty} [a_{4l+1} - a_{4l+3}] \right) i \\ &= x + yi \end{aligned}$$

where

$$x = \sum_{l=0}^{\infty} [a_{4l} - a_{4l+2}] \text{ and } y = \sum_{l=0}^{\infty} [a_{4l+1} - a_{4l+3}].$$

Example 3.5. Working as in the last example we see that

$$\mathbb{Z}\left[\sqrt{2}\right] = \left\{ a + b\sqrt{2} : a, b \in \mathbb{Z} \right\}$$

is a sub-ring of $\mathbb{R}$.

Example 3.6 (Gaussian Integers mod m*).* For any $m \geq 2$, let

$$\mathbb{Z}_m[i] = \{x + yi : x, y \in \mathbb{Z}_m\}$$

with the obvious addition rule and multiplication given by

$$(x + yi)(u + vi) = ux - vy + (uy + vx) i \text{ in } \mathbb{Z}_m.$$

The next proposition shows that this is a commutative ring with identity, 1.

Proposition 3.7. *Let R be a commutative ring with identity and let*

$$R[i] := \{a + bi : a, b \in R\} \cong \{(a, b) : a, b \in R\} = R^2.$$

Define addition and multiplication of $R[i]$ as one expects by,

$$(a + bi) + (c + di) = (a + c) + (b + d) i$$

and

$$(a + bi) \cdot (c + di) = (ac - bd) + (bc + ad) i.$$

Then $(R[i], +, \cdot)$ is a commutative ring with identity.

Proof. This can be checked by brute force. Rather than use brute force lets give a proof modeled on Example 1.11, i.e. we will observe that we may identify $R[i]$ with a unital subring of $M_2(R)$. To do this we take,

$$\mathbf{i} := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in M_2(R) \text{ and } 1 := I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in M_2(R).$$

Thus we take,

$$a + ib \longleftrightarrow aI + b\mathbf{i} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \in M_2(R).$$

Since

$$\begin{aligned} (aI + b\mathbf{i}) + (cI + d\mathbf{i}) &= \begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \\ &= \begin{bmatrix} a + c & -b - d \\ b + d & a + c \end{bmatrix} \\ &= (a + c) I + (b + d) \mathbf{i} \end{aligned}$$

and

$$\begin{aligned} (aI + b\mathbf{i})(cI + d\mathbf{i}) &= \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \\ &= \begin{bmatrix} ac - bd & -ad - bc \\ ad + bc & ac - bd \end{bmatrix} \\ &= (ac - bd) I + (bc + ad) \mathbf{i} \end{aligned}$$

we see that

$$S := \left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = aI + b\mathbf{i} : a, b \in R \right\}$$

is indeed a unital sub-ring of $M_2(R)$. Moreover, the multiplication rules on S and $R[i]$ agree under the identification; $a + ib \longleftrightarrow aI + b\mathbf{i}$. Therefore we may conclude that $(R[i], +, \cdot)$ satisfies the properties of a ring. ■

3.3 Appendix: $R\left[S\right]$ rings II

You may skip this section on first reading.

Definition 3.8. *Suppose that S is a set which is equipped with an associative binary operation, $\cdot$, which has a unique unit denoted by e. (We do not assume that $(S,\cdot)$ has inverses. Also suppose that R is a ring, then we let $R\left[S\right]$ consist of the formal sums, $\sum_{s\in S} a_s s$ where $\left\{a_s\right\}_{s\in S} \subset R$ is a sequence with finite support, i.e. $\left|\left\{s\in S : a_s \neq 0\right\}\right| < \infty$. We define two binary operations on $R\left[S\right]$ by,*

$$\sum_{s\in S} a_s s + \sum_{s\in S} b_s s := \sum_{s\in S} \left(a_s + b_s\right) s$$

and

$$\begin{aligned}\sum_{s\in S} a_s s \cdot \sum_{s\in S} b_s s &= \sum_{s\in S} a_s s \cdot \sum_{t\in S} b_t t \\ &= \sum_{s,t\in S} a_s b_t st = \sum_{u\in S}\left(\sum_{st=u} a_s b_t\right) u.\end{aligned}$$

So really we $R\left[S\right]$ are those sequences $a := \left\{a_s\right\}_{s\in S}$ with finite support with the operations,

$$\left(a+b\right)_s = a_s + b_s \text{ and } \left(a\cdot b\right)_s = \sum_{uv=s} a_u b_v \text{ for all } s \in S.$$

Theorem 3.9. *The set $R\left[S\right]$ equipped with the two binary operations $(+,\cdot)$ is a ring.*

Proof. Because $(R,+)$ is an abelian group it is easy to check that $\left(R\left[S\right],+\right)$ is an abelian group as well. Let us now check that $\cdot$ is associative on $R\left[S\right]$. To this end, let $a,b,c \in R\left[S\right]$, then

$$\begin{aligned}\left[a\left(bc\right)\right]_s &= \sum_{uv=s} a_u \left(bc\right)_v = \sum_{uv=s} a_u \left(\sum_{\alpha\beta=v} b_\alpha c_\beta\right) \\ &= \sum_{u\alpha\beta=s} a_u b_\alpha c_\beta\end{aligned}$$

while

$$\begin{aligned}\left[\left(ab\right)c\right]_s &= \sum_{\alpha\beta=s} \left(ab\right)_\alpha c_\beta = \sum_{\alpha\beta=s}\sum_{uv=\alpha} a_u b_v c_\beta \\ &= \sum_{uv\beta=s} a_u b_v c_\beta = \sum_{u\alpha\beta=s} a_u b_\alpha c_\beta = \left[a\left(bc\right)\right]_s\end{aligned}$$

as desired. Secondly,

$$\begin{aligned}\left[a\cdot\left(b+c\right)\right]_s &= \sum_{uv=s} a_u \left(b+c\right)_v = \sum_{uv=s} a_u \left(b_v + c_v\right) \\ &= \sum_{uv=s} a_u b_v + \sum_{uv=s} a_u c_v \\ &= \left[a\cdot b\right]_s + \left[a\cdot c\right]_s = \left[a\cdot b + a\cdot c\right]_s\end{aligned}$$

from which it follows that $a\cdot\left(b+c\right) = a\cdot b + a\cdot c$. Similarly one shows that $\left(b+c\right)\cdot a = b\cdot a + c\cdot a$.

Lastly if S has an identity, e, and $\mathbf{e}_s := 1_{s=e} \in R$, then

$$\left[a\cdot\mathbf{e}\right]_s = \sum_{uv=s} a_u \mathbf{e}_v = a_s$$

from which it follows that $\mathbf{e}$ is the identity in $R\left[S\right]$. ■

Example 3.10 (Polynomial rings). Let x be a formal symbol and let $S := \left\{x^k : k = 0,1,2\dots\right\}$ with $x^k x^l := x^{k+l}$ being the binary operation of S. Notice that x^0 is the identity in S under this multiplication rule. Then for any ring R, we have

$$R\left[S\right] = \left\{p\left(x\right) := \sum_{k=0}^{n} p_k x^k : p_k \in R \text{ and } n\in\mathbb{N}\right\}.$$

The multiplication rule is given by

$$p\left(x\right)q\left(x\right) = \sum_{k=0}^{\infty}\left(\sum_{j=0}^{k} p_j q_{k-j}\right) x^k$$

which is the usual formula for multiplication of polynomials. In this case it is customary to write $R\left[x\right]$ rather than $R\left[S\right]$.

This example has natural generalization to multiple indeterminants as follows.

Example 3.11. Suppose that $x = \left(x_1,\dots,x_d\right)$ are d indeterminants and $k = \left(k_1,\dots,k_d\right)$ are multi-indices. Then we let

$$S := \left\{x^k := x_1^{k_1}\dots x_d^{k_d} : k\in\mathbb{N}^d\right\}$$

with multiplication law given by

$$x^k x^{k'} := x^{k+k'}.$$

Then

$$R[S] = \left\{ p(x) := \sum_k p_k x^k : p_k \in R \text{ with } p_k = 0 \text{ a.a.} \right\}.$$

We again have the multiplication rule,

$$p(x)\, q(x) = \sum_k \left(\sum_{j \le k} p_j q_{k-j} \right) x^k.$$

As in the previous example, it is customary to write $R[x_1, \dots, x_d]$ for $R[S]$.

In the next example we wee that the multiplication operation on S need not be commutative.

Example 3.12 (Group Rings). In this example we take $S = G$ where G is a group which need not be commutative. Let R be a ring and set,

$$R[G] := \{a : G \to R| \ |\{g :\in G\} : a(g) \neq 0| < \infty\}.$$

We will identify $a \in R[G]$ with the formal sum,

$$a := \sum_{g \in G} a(g)\, g.$$

We define $(a+b)(g) := a(g) + b(g)$ and

$$\begin{aligned} a \cdot b &= \left(\sum_{g \in G} a(g)\, g \right) \left(\sum_{k \in G} b(k)\, k \right) = \sum_{g,k \in G} a(g)\, b(k)\, gk \\ &= \sum_{h \in G} \left(\sum_{gk = h} a(g)\, b(k) \right) h = \sum_{h \in G} \left(\sum_{g \in G} a(g)\, b\left(g^{-1}h\right) \right) h. \end{aligned}$$

So formally we define,

$$\begin{aligned} (a \cdot b)(h) &:= \sum_{g \in G} a(g)\, b\left(g^{-1}h\right) = \sum_{g \in G} a(hg)\, b\left(g^{-1}\right) = \sum_{g \in G} a\left(hg^{-1}\right) b(g) \\ &= \sum_{gk = h} a(g)\, b(k). \end{aligned}$$

We now claim that R is a ring which is non − commutative when G is non-abelian.

Let us check associativity and distributivity of $\cdot$. To this end,

$$\begin{aligned} [(a \cdot b) \cdot c](h) &= \sum_{gk = h} (a \cdot b)(g) \cdot c(k) \\ &= \sum_{gk = h} \left[\sum_{uv = g} a(u) \cdot b(v) \right] \cdot c(k) \\ &= \sum_{uvk = h} a(u) \cdot b(v) \cdot c(k) \end{aligned}$$

while on the other hand,

$$\begin{aligned} [a \cdot (b \cdot c)](h) &= \sum_{uy = h} a(u) \cdot (b \cdot c)(y) \\ &= \sum_{uy = h} a(u) \cdot \left(\sum_{vk = y} b(v) \cdot c(y) \right) \\ &= \sum_{uvk = h} a(u) \cdot (b(v) \cdot c(y)) \\ &= \sum_{uvk = h} a(u) \cdot b(v) \cdot c(k). \end{aligned}$$

For distributivity we find,

$$\begin{aligned} [(a+b) \cdot c](h) &= \sum_{gk = h} (a+b)(g) \cdot c(k) = \sum_{gk = h} (a(g) + b(g)) \cdot c(k) \\ &= \sum_{gk = h} (a(g) \cdot c(k) + b(g) \cdot c(k)) \\ &= \sum_{gk = h} a(g) \cdot c(k) + \sum_{gk = h} b(g) \cdot c(k) \\ &= [a \cdot c + b \cdot c](h) \end{aligned}$$

with a similar computation showing $c \cdot (a+b) = c \cdot a + c \cdot b$.

Lecture 4

4.1 Units

Definition 4.1. *Suppose R is a ring with identity. A **unit** of a ring is an element $a \in R$ such that there exists an element $b \in R$ with $ab = ba = 1$. We let $U(R) \subset R$ denote the units of R.*

Notice that in fact $a = b$ in this definition since,

$$a = a \cdot 1 = a(ub) = (au)b = 1 \cdot b = b.$$

Moreover this argument shows that a satisfying $au = 1 = ua$ is unique if it exists. For this reason we will write u^{-1} for a.

Proposition 4.2. *The set $U(R)$ equipped the multiplication law of R is a group.*

Proof. This is a straight forward verification – see the homework assignment. The main point is to observe that $u, v \in U(R)$, then $a := v^{-1}u^{-1}$ satisfies, $a(uv) = 1 = (uv)a$, showing $U(R)$ is closed under the multiplication operation of R. ■

Example 4.3. In $M_2(\mathbb{R})$, the units in this ring are exactly the elements in $GL(2, \mathbb{R})$, i.e.

$$U(M_2(\mathbb{R})) = GL(2, \mathbb{R}) = \{A \in M_2(\mathbb{R}) : \det A \neq 0\}.$$

If you look back at last quarters notes you will see that we have already proved the following theorem. I will repeat the proof here for completeness.

Theorem 4.4 ($U(\mathbb{Z}_m) = U(m)$). *For any $m \geq 2$,*

$$U(\mathbb{Z}_m) = U(m) = \{a \in \{1, 2, \dots, m-1\} : \gcd(a, m) = 1\}.$$

Proof. If $a \in U(\mathbb{Z}_m)$, there there exists $r \in \mathbb{Z}_m$ such that $1 = r \cdot a = ra \bmod m$. Equivalently put, $m|(ra-1)$, i.e. there exists t such that $ra - 1 = tm$. Since $1 = ra - tm$ it follows that $\gcd(a, m) = 1$, i.e. that $a \in U(m)$.

Conversely, if $a \in U(m) \iff \gcd(a, m) = 1$ which we know implies there exists $s, t \in \mathbb{Z}$ such that $sa + tm = 1$. Taking this equation $\bmod\, m$ and letting $b := s \bmod m \in \mathbb{Z}_m$, we learn that $b \cdot a = 1$ in $\mathbb{Z}_m$, i.e. $a \in U(\mathbb{Z}_m)$. ■

Example 4.5. In $\mathbb{R}$, the units are exactly the elements in $\mathbb{R}^\times := \mathbb{R} \setminus \{0\}$ that is $U(\mathbb{R}) = \mathbb{R}^\times$.

Example 4.6. Let R be the non-commutative ring of linear maps from $\mathbb{R}^\infty$ to $\mathbb{R}^\infty$ where

$$\mathbb{R}^\infty = \{(a_1, a_2, a_3, \dots) : a_i \in \mathbb{R} \text{ for all } i\},$$

which is a vector space over $\mathbb{R}$. Further let $A, B \in R$ be defined by

$$A(a_1, a_2, a_3, \dots) = (0, a_1, a_2, a_3, \dots) \text{ and}$$
$$B(a_1, a_2, a_3, \dots) = (a_2, a_3, a_4, \dots).$$

Then $BA = \mathbf{1}$ where

$$\mathbf{1}(a_1, a_2, a_3, \dots) = (a_1, a_2, a_3, \dots)$$

while

$$AB(a_1, a_2, a_3, \dots) = (0, a_2, a_3, \dots) \neq \mathbf{1}(a_1, a_2, a_3, \dots).$$

This shows that even though $BA = \mathbf{1}$ it is not necessarily true that $AB = 1$. Neither A nor B are units of $\mathbb{R}^\infty$.

4.2 (Zero) Divisors and Integral Domains

Definition 4.7 (Divisors). *Let R be a ring. We say that for elements $a, b \in R$ that a **divides** b if there exists an element c such that $ac = b$.*

Note that if $R = \mathbb{Z}$ then this is the usual notion of whether one integer evenly divides another, e.g., 2 divides 6 and 2 doesn't divide 5.

Definition 4.8 (Zero divisors). *A nonzero element $a \in R$ is called a **zero divisor** if there exists another nonzero element $b \in R$ such that $ab = 0$, i.e. a divides 0 in a nontrivial way. (The trivial way for $a|0$ is; $0 = a \cdot 0$ as this always holds.)*

Definition 4.9 (Integral domain). *A commutative ring R with no zero divisors is called an **integral domain** (or just a **domain**). Alternatively put, R should satisfy, $ab \neq 0$ for all $a, b \in R$ with $a \neq 0 \neq b$.*

Example 4.10. The most familiar rings to you, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ have no zero-divisors and hence are integral domains.. In these number systems, it is a familiar fact that $ab = 0$ implies either $a = 0$ or $b = 0$. Another integral domain is the polynomial ring $\mathbb{R}[x]$, see Proposition 4.13 below.

Example 4.11. The ring, $\mathbb{Z}_6$, is not an integral domain. For example, $2 \cdot 3 = 0$ with $2 \neq 0 \neq 3$, so both 2 and 3 are zero divisors.

Lemma 4.12. *The ring $\mathbb{Z}_m$ is an integral domain iff m is prime.*

Proof. If m is prime we know that $U(\mathbb{Z}_m) = U(m) = \mathbb{Z}_m \setminus \{0\}$. Therefore if $a, b \in \mathbb{Z}_m$ with $a \neq 0$ and $ab = 0$ then $b = a^{-1}ab = a^{-1}0 = 0$.

If $m = a \cdot b$ with $a, b \in \mathbb{Z}_m \setminus \{0\}$, then $ab = 0$ while both a and b are not equal to zero in $\mathbb{Z}_m$. ■

Proposition 4.13. *If R is an integral domain, then so is $R[x]$. Conversely if R is not an integral domain then neither is $R[x]$.*

Proof. If $f, g \in R[x]$ are two non-zero polynomials. Then $f = a_n x^n +$ l.o.ts. (lower order terms) and $g = b_m x^m +$ l.o.ts. with $a_n \neq 0 \neq b_m$ and therefore,

$$fg = a_n b_m x^{n+m} + \text{l.o.ts.} \neq 0 \text{ since } a_n b_m \neq 0.$$

The proof of the second assertion is left to the reader. ■

Example 4.14. All of the following rings are integral domains; $\mathbb{Z}[x], \mathbb{Q}[x], \mathbb{R}[x]$, and $\mathbb{C}[x]$. We also know that $\mathbb{Z}_m[x]$ is an integral domain iff m is prime.

Example 4.15. If R is the direct product of at least 2 rings, then R has zero divisors. For example if $R = \mathbb{Z} \oplus \mathbb{Z}$, then $(0, b)(a, 0) = (0, 0)$ for all $a, b \in \mathbb{Z}$.

Example 4.16. If R is an integral domain, then any unital subring $S \subset R$ is also an integral domain. In particular, for any $\theta \in \mathbb{C}$, then $\mathbb{Z}[\theta], \mathbb{Q}[\theta]$, and $\mathbb{R}[\theta]$ are all integral domains.

Remark 4.17. It is not true that if R is not an integral domain then every sub-ring, $S \subset R$ is also not an integral domain. For an example, take $R := \mathbb{Z} \oplus \mathbb{Z}$ and $S := \{(a, a) : a \in \mathbb{Z}\} \subset R$. (In the language of Section 5.1 below, $S = \{n \cdot (1, 1) : n \in \mathbb{Z}\}$ which is the sub-ring generated by $1 = (1, 1)$. Similar to this counter example, commutative ring with identity which is not an integral domain but has characteristic being either 0 or prime would give a counter example.)

Domains behave more nicely than arbitrary rings and for a lot of the quarter we will concentrate exclusively on domains. But in a lot of ring theory it is very important to consider rings that are not necessarily domains like matrix rings.

Theorem 4.18 (Cancellation). *If R is an integral domain and $ab = ac$ with $a \neq 0$, then $b = c$. Conversely if R is a commutative ring with identity satisfying this cancellation property then R has no zero divisors and hence is an integral domain.*

Proof. If $ab = ac$, then $a(b - c) = 0$. Hence if $a \neq 0$ and R is an integral domain, then $b - c = 0$, i.e. $b = c$.

Conversely, if R satisfies cancellation and $ab = 0$. If $a \neq 0$, then $ab = a \cdot 0$ and so by cancellation, $b = 0$. This shows that R has no zero divisors. ■

Example 4.19. The ring, $M_2(\mathbb{R})$ contains many zero divisors. For example

$$\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

So in $M_2(\mathbb{R})$ we can not conclude that $B = 0$ if $AB = 0$ with $A \neq 0$, i.e. cancellation does not hold.

4.3 Fields

If we add one more restriction to a domain we get a familiar class of objects called fields.

Definition 4.20 (Fields). *A ring R is a* ***field*** *if R is a commutative ring with identity and $U(R) = R \setminus \{0\}$, that is, every non-zero element of R is a unit, in other words has a multiplicative inverse.*

Lemma 4.21 (Fields are domains). *If R is a field then R is an integral domain.*

Proof. If R is a field and $xy = 0$ in R for some x, y with $x \neq 0$, then

$$0 = x^{-1}0 = x^{-1}xy = y.$$

■

Example 4.22. $\mathbb{Z}$ is an integral domain that is not a field. For example $2 \neq 0$ has no multiplicative inverse. The inverse to 2 should be $\frac{1}{2}$ which exists in $\mathbb{Q}$ but not in $\mathbb{Z}$. On the other hand, $\mathbb{Q}$ and $\mathbb{R}$ are fields as the non-zero elements have inverses back in $\mathbb{Q}$ and $\mathbb{R}$ respectively.

Example 4.23. We have already seen that $\mathbb{Z}_m$ is a field iff m is prime. This follows directly form the fact that $U(\mathbb{Z}_m) = U(m)$ and $U(m) = \mathbb{Z}_m \setminus \{0\}$ iff m is prime. Recall that we also seen that $\mathbb{Z}_m$ is an integral domain iff m is prime so it follows $\mathbb{Z}_m$ is a field iff it is an integral domain iff m is prime. When p is prime, we will often denote $\mathbb{Z}_p$ by $\mathbb{F}_p$ to indicate that we are viewing $\mathbb{Z}_p$ is a field.

5

Lecture 5

In fact, there is another way we could have seen that $\mathbb{Z}_p$ is a field, using the following useful lemma.

Lemma 5.1. *If R be an integral domain with finitely many elements, then R is a field.*

Proof. Let $a \in R$ with $a \neq 0$. We need to find a multiplicative inverse for a. Consider $a, a^2, a^3, \ldots$. Since R is finite, the elements on this list are not all distinct. Suppose then that $a^i = a^j$ for some $i > j \geq 1$. Then $a^j a^{i-j} = a^j \cdot 1$. By cancellation, since R is a domain, $a^{i-j} = 1$. Then a^{i-j-1} is the inverse for a. Note that $a^{i-j-1} \in R$ makes sense because $i - j - 1 \geq 0$. ■

For general rings, a^n only makes sense for $n \geq 1$. If $1 \in R$ and $a \in U(R)$, we may define $a^0 = 1$ and $a^{-n} = \left(a^{-1}\right)^n$ for $n \in \mathbb{Z}_+$. As for groups we then have $a^n a^m = a^{n+m}$ for all $m, n \in \mathbb{Z}$. makes sense for all $n \in \mathbb{Z}$, but in generally negative powers don't always make sense in a ring. Here is another very interesting example of a field, different from the other examples we've written down so far.

Example 5.2. Lets check that $\mathbb{C}$ is a field. Given $0 \neq a + bi \in \mathbb{C}$, $a, b \in \mathbb{R}$, $i = \sqrt{-1}$, we need to find $(a + ib)^{-1} \in \mathbb{C}$. Working formally; we expect,

$$\begin{aligned}(a+ib)^{-1} &= \frac{1}{a+bi} = \frac{1}{a+bi}\frac{a-bi}{a-bi}\frac{a-bi}{a^2+b^2} \\ &= \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2} i \in \mathbb{C},\end{aligned}$$

which makes sense if $N(a + ib) := a^2 + b^2 \neq 0$, i.e. $a + ib \neq 0$. A simple direct check show that this formula indeed gives an inverse to $a + ib$;

$$\begin{aligned}&(a+ib)\left[\frac{a}{a^2+b^2} - \frac{b}{a^2+b^2} i\right] \\ &= \frac{1}{a^2+b^2}(a+ib)(a-ib) = \frac{1}{a^2+b^2}\left(a^2+b^2\right) = 1.\end{aligned}$$

So if $a + ib \neq 0$ we have shown

$$(a+bi)^{-1} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2} i.$$

Example 5.3. I claim that $R := \mathbb{Z}_3[i] = \mathbb{Z}_3 + i\mathbb{Z}_3$ is a field where we use the multiplication rule,

$$(a+ib)(c+id) = (ac - bd) + i(bc + ad).$$

The main point to showing this is a field beyond showing R is a ring (see Proposition 3.7) is to show $(a + ib)^{-1}$ exists in R whenever $a + ib \neq 0$. Working formally for the moment we should have,

$$\frac{1}{a+ib} = \frac{a-ib}{a^2+b^2}.$$

This suggest that

$$(a+ib)^{-1} = \left(a^2+b^2\right)^{-1}(a-ib).$$

In order for the latter expression to make sense we need to know that $a^2 + b^2 \neq 0$ in $\mathbb{Z}_3$ if $(a, b) \neq 0$ which we can check by brute force;

a	0	0	0	1	1	1	2	2	2
b	0	1	2	0	1	2	0	1	2
$N(a+ib)$ $= a^2+b^2$	0	1	1	1	2	2	1	2	2

.

Alternatively we may show $\mathbb{Z}_3[i]$ is an integral domain and then use Lemma 5.1. Notice that

$$\begin{aligned}(a+ib)(c+id) = 0 \implies (a-ib)(a+ib)(c+id) = 0 \text{ i.e.} \\ \left(a^2+b^2\right)(c+id) = 0.\end{aligned}$$

So using the chart above, we see that $a^2 + b^2 = 0$ iff $a + ib = 0$ and therefore, if $a + ib \neq 0$ then $c + id = 0$.

5.1 Characteristic of a Ring

Notation 5.4 *Suppose that $a \in R$ where R is a ring. Then for $n \in \mathbb{Z}$ we define $n \cdot a \in R$ by, $0_{\mathbb{Z}} \cdot a = 0_R$ and*

$$n \cdot a = \begin{cases} \overbrace{a + \cdots + a}^{n \text{ times}} & \text{if } n \geq 1 \\ -\overbrace{(a + \cdots + a)}^{|n| \text{ times}} = |n| \cdot (-a) & \text{if } n \leq -1 \end{cases}.$$

So $3 \cdot a = a + a + a$ *while* $-2 \cdot a = -a - a$.

Lemma 5.5. *Suppose that* R *is a ring and* $a, b \in R$. *Then for all* $m, n \in \mathbb{Z}$ *we have*

$$(m \cdot a)\, b = m \cdot (ab)\,, \tag{5.1}$$
$$a\,(m \cdot b) = m \cdot (ab)\,. \tag{5.2}$$

We also have

$$-\,(m \cdot a) = (-m) \cdot a = m \cdot (-a) \;\; \textit{and} \tag{5.3}$$
$$m \cdot (n \cdot a) = mn \cdot a. \tag{5.4}$$

Proof. If $m = 0$ both sides of Eq. (5.1) are zero. If $m \in \mathbb{Z}_+$, then using the distributive associativity laws repeatedly gives;

$$\begin{aligned} (m \cdot a)\, b &= \overbrace{(a + \cdots + a)}^{m \text{ times}} b \\ &= \overbrace{(ab + \cdots + ab)}^{m \text{ times}} = m \cdot (ab)\,. \end{aligned}$$

If $m < 0$, then

$$(m \cdot a)\, b = (|m| \cdot (-a))\, b = |m| \cdot ((-a)\, b) = |m| \cdot (-ab) = m \cdot (ab)$$

which completes the proof of Eq. (5.1). The proof of Eq. (5.2) is similar and will be omitted.

If $m = 0$ Eq. (5.3) holds. If $m \geq 1$, then

$$-\,(m \cdot a) = -\overbrace{(a + \cdots + a)}^{m \text{ times}} = \overbrace{((-a) + \cdots + (-a))}^{m \text{ times}} = m \cdot (-a) = (-m) \cdot a.$$

If $m < 0$, then

$$-\,(m \cdot a) = -\,(|m| \cdot (-a)) = (-\,|m|) \cdot (-a) = m \cdot (-a)$$

and

$$-\,(m \cdot a) = -\,(|m| \cdot (-a)) = (|m| \cdot (-\,(-a))) = |m| \cdot a = (-m) \cdot a.$$

which proves Eq. (5.3).

Letting $x := \operatorname{sgn}(m) \operatorname{sgn}(n) a$, we have

$$\begin{aligned} m \cdot (n \cdot a) = |m| \cdot (|n| \cdot x) &= \overbrace{(|n| \cdot x + \cdots + |n| \cdot x)}^{|m| \text{ times}} \\ &= \overbrace{\overbrace{(x + \cdots + x)}^{|n| \text{ times}} + \cdots + \overbrace{(x + \cdots + x)}^{|n| \text{ times}}}^{|m| \text{ times}} \\ &= (|m|\, |n|) \cdot x = mn \cdot a. \end{aligned}$$

■

Corollary 5.6. *If* R *is a ring,* $a, b \in R$, *and* $m, n \in \mathbb{Z}$, *then*

$$(m \cdot a)\,(n \cdot b) = mn \cdot ab. \tag{5.5}$$

Proof. Using Lemma 5.5 gives;

$$(m \cdot a)\,(n \cdot b) = m \cdot (a\,(n \cdot b)) = m \cdot (n \cdot (ab)) = mn \cdot ab.$$

■

Corollary 5.7. *Suppose that* R *is a ring and* $a \in R$. *Then for all* $m, n \in \mathbb{Z}$,

$$(m \cdot a)\,(n \cdot a) = mn \cdot a^2.$$

In particular if $a = 1 \in R$ *we have,*

$$(m \cdot 1)\,(n \cdot 1) = mn \cdot 1.$$

Unlike the book, we will only bother to define the characteristic for rings which have an identity, $1 \in R$.

Definition 5.8 (Characteristic of a ring). *Let* R *be a ring with* $1 \in R$. *The characteristic,* $\operatorname{chr}(R)$, *of* R *is is the order of the element* 1 *in the additive group* $(R, +)$. *Thus* n *is the smallest number in* $\mathbb{Z}_+$ *such that* $n \cdot 1 = 0$. *If no such* $n \in \mathbb{Z}_+$ *exists, we say that characteristic of* R *is* 0 *by convention and write* $\operatorname{chr}(R) = 0$.

Lemma 5.9. *If* R *is a ring with identity and* $\operatorname{chr}(R) = n \geq 1$, *then* $n \cdot x = 0$ *for all* $x \in R$.

Proof. For any $x \in R$, $n \cdot x = n \cdot (1x) = (n \cdot 1)\, x = 0x = 0$. ■

Lemma 5.10. *Let* R *be a domain. If* $n = \operatorname{chr}(R) \geq 1$, *then* n *is a prime number.*

Proof. If n is not prime, say $n = pq$ with $1 < p < n$ and $1 < q < n$, then

$$(p \cdot 1_R)(q \cdot 1_R) = pq \cdot (1_R 1_R) = pq \cdot 1_R = n \cdot 1_R = 0.$$

As $p \cdot 1_R \neq 0$ and $q \cdot 1_R \neq 0$ and we may conclude that both $p \cdot 1_R$ and $q \cdot 1_R$ are zero divisors contradicting the assumption that R is an integral domain. ■

Example 5.11. The rings $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}\left[\sqrt{d}\right] := \mathbb{Q} + \mathbb{Q}\sqrt{d}$, $\mathbb{Z}[x]$, $\mathbb{Q}[x]$, $\mathbb{R}[x]$, and $\mathbb{Z}[x]$ all have characteristic 0.

For each $m \in \mathbb{Z}_+$, $\mathbb{Z}_m$ and $\mathbb{Z}_m[x]$ are rings with characteristic m.

Example 5.12. For each prime, p, $\mathbb{F}_p := \mathbb{Z}_p$ is a field with characteristic p. We also know that $\mathbb{Z}_3[i]$ is a field with characteristic 3. Later, we will see other examples of fields of characteristic p.

6

Lecture 6

6.1 Square root field extensions of $\mathbb{Q}$

Recall that $\sqrt{2}$ is irrational. Indeed suppose that $\sqrt{2} = m/n \in \mathbb{Q}$ and, with out loss of generality, assume that $\gcd(m,n) = 1$. Then $m^2 = 2n^2$ from which it follows that $2|m^2$ and so $2|m$ by Euclid's lemma. However, it now follows that $2^2|2n^2$ and so $2|n^2$ which again by Euclid's lemma implies $2|n$. However, we assumed that m and n were relatively prime and so we have a contradiction and hence $\sqrt{2}$ is indeed irrational. As a consequence of this fact, we know that $\left\{1, \sqrt{2}\right\}$ are linearly independent over $\mathbb{Q}$, i.e. if $a + b\sqrt{2} = 0$ then $a = 0 = b$.

Example 6.1. In this example we will show,

$$R = \mathbb{Q}\left[\sqrt{2}\right] = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\} \tag{6.1}$$

is a field. Using similar techniques to those in Example 3.4 we see that $\mathbb{Q}\left[\sqrt{2}\right]$ may be described as in Eq. (6.1) and hence is a subring of $\mathbb{Q}$ by Proposition 3.3. Alternatively one may check directly that the right side of Eq. (6.1) is a subring of $\mathbb{Q}$ since;

$$a + b\sqrt{2} - (c + d\sqrt{2}) = (a - c) + (b - d)\sqrt{2} \in R$$

and

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

So by either means we see that R is a ring and in fact an integral domain by Example 4.16. It does not have finitely many elements so we can't use Lemma 5.1 to show it is a field. However, we can find $\left(a + b\sqrt{2}\right)^{-1}$ directly as follows. If $\xi = \left(a + b\sqrt{2}\right)^{-1}$, then

$$1 = (a + b\sqrt{2})\xi$$

and therefore,

$$a - b\sqrt{2} = (a - b\sqrt{2})(a + b\sqrt{2})\xi = \left(a^2 - 2b^2\right)\xi$$

which implies,

$$\xi = \frac{a}{a^2 - 2b^2} + \frac{-b}{a^2 - 2b^2}\sqrt{2} \in \mathbb{Q}\left[\sqrt{2}\right].$$

Moreover, it is easy to check this ξ works provided $a^2 - 2b^2 \neq 0$. But if $a^2 - 2b^2 = 0$ with $b \neq 0$, then $\sqrt{2} = |a| / |b|$ showing $\sqrt{2}$ is irrational which we know to be false – see Proposition 6.2 below for details. Therefore, $\mathbb{Q}\left[\sqrt{2}\right]$ is a field.

Observe that $\mathbb{Q} \subsetneq R := \mathbb{Q}\left[\sqrt{2}\right] \subsetneq \mathbb{R}$. Why is this? One reason is that $R := \mathbb{Q}\left[\sqrt{2}\right]$ is countable and $\mathbb{R}$ is uncountable. Or it is not hard to show that an irrational number selected more or less at random is not in R. For example, you could show that $\sqrt{3} \notin R$. Indeed if $\sqrt{3} = a + b\sqrt{2}$ for some $a, b \in \mathbb{Q}$ then

$$3 = a^2 + 2ab\sqrt{2} + 2b^2$$

and hence $2ab\sqrt{2} = 3 - a^2 - 2b^2$. Since $\sqrt{2}$ is irrational, this can only happen if either $a = 0$ or $b = 0$. If $b = 0$ we will have $\sqrt{3} \in \mathbb{Q}$ which is false and if $a = 0$ we will have $3 = 2b^2$. Writing $b = \frac{k}{l}$, this with $\gcd(k,l) = 1$, we find $3l^2 = 2k^2$ and therefore $2|l$ by Gauss' lemma. Hence $2^2|2k^2$ which implies $2|k$ and therefore $\gcd(k,l) \geq 2 > 1$ which is a contradiction. Hence it follows that $\sqrt{3} \neq a + b\sqrt{2}$ for any $a, b \in \mathbb{Q}$.

The following proposition is a natural extension of Example 6.1.

Proposition 6.2. *For all $d \in \mathbb{Z} \setminus \{0\}$, $F := \mathbb{Q}\left[\sqrt{d}\right]$ is a field. (As we will see in the proof, we need only consider those d which are "square prime" free.*

Proof. As $F := \mathbb{Q}\left[\sqrt{d}\right] = \mathbb{Q} + \mathbb{Q}\sqrt{d}$ is a subring of $\mathbb{R}$ which is an integral domain, we know that F is again an integral domain. Let $d = \varepsilon p_1^{k_1} \dots p_n^{k_n}$ with $\varepsilon \in \{\pm 1\}$, $p_1, \dots, p_n$ being distinct primes, and $k_i \geq 1$. Further let $\delta = \varepsilon \prod_{i:k_i \text{ is odd}} p_i$, then $\sqrt{d} = m\sqrt{\delta}$ for some integer m and therefore it easily follows that $F = \mathbb{Q}\left[\sqrt{\delta}\right]$. So let us now write $\delta = \varepsilon p_1 \dots p_k$ with $\varepsilon \in \{\pm 1\}$, $p_1, \dots, p_k$ being distinct primes so that δ is **square prime free.**

Working as above we look for the inverse to $a + b\sqrt{\delta}$ when $(a,b) \neq 0$. Thus we will look for $u, v \in \mathbb{Q}$ such that

$$1 = \left(a + b\sqrt{\delta}\right)\left(u + v\sqrt{\delta}\right).$$

Multiplying this equation through by $a - b\sqrt{\delta}$ shows,

$$a-b\sqrt{\delta}=\left(a^2-b^2\delta\right)\left(u+v\sqrt{\delta}\right)$$

so that

$$u+v\sqrt{\delta}=\frac{a}{a^2-b^2\delta}-\frac{b}{a^2-b^2\delta}\sqrt{\delta}. \tag{6.2}$$

Thus we may define,

$$\left(a+b\sqrt{\delta}\right)^{-1}=\frac{a}{a^2-b^2\delta}-\frac{b}{a^2-b^2\delta}\sqrt{\delta}$$

provided $a^2-b^2\delta\neq 0$ when $(a,b)\neq(0,0)$.

Case 1. If $\delta<0$ then $a^2-b^2\delta=a^2+|\delta|\,b^2=0$ iff $a=0=b$.

Case 2. If $\delta\geq 2$ and suppose that $a,b\in\mathbb{Q}$ with $a^2=b^2\delta$. For sake of contradiction suppose that $b\neq 0$. By multiplying $a^2=b^2\delta$ though by the denominators of a^2 and b^2 we learn there are integers, $m,n\in\mathbb{Z}_+$ such that $m^2=n^2\delta$. By replacing m and n by $\frac{m}{\gcd(m,n)}$ and $\frac{n}{\gcd(m,n)}$, we may assume that m and n are relatively prime.

We now have $p_1|\left(n^2\delta\right)$ implies $p_1|m^2$ which by Euclid's lemma implies that $p_1|m$. Thus we learn that $p_1^2|m^2=n^2p_1,\dots,p_k$ and therefore that $p_1|n^2$. Another application of Euclid's lemma shows $p_1|n$. Thus we have shown that p_1 is a divisor of both m and n contradicting the fact that m and n were relatively prime. Thus we must conclude that $b=0=a$. Therefore $a^2-b^2\delta=0$ only if $a=0=b$. ■

Later on we will show the following;

Fact 6.3 *Suppose that $\theta\in\mathbb{C}$ is the root of some polynomial in $\mathbb{Q}[x]$, then $\mathbb{Q}[\theta]$ is a sub-field of $\mathbb{C}$.*

Recall that we already know $\mathbb{Q}[\theta]$ is an integral domain. To prove that $\mathbb{Q}[\theta]$ is a field we will have to show that for every nonzero $z\in\mathbb{Q}[\theta]$ that the inverse, $z^{-1}\in\mathbb{C}$, is actually back in $\mathbb{Q}[\theta]$.

6.2 Homomorphisms

Definition 6.4. *Let R and S be rings. A function $\varphi:R\to S$ is a **homomorphism** if*

$$\varphi(r_1r_2)=\varphi(r_1)\varphi(r_2)\text{ and}$$
$$\varphi(r_1+r_2)=\varphi(r_1)+\varphi(r_2)$$

*for all $r_1,r_2\in R$. That is, φ preserves addition and multiplication. If we further assume that φ is an invertible map (i.e. one to one and onto), then we say $\varphi:R\to S$ is an **isomorphism** and that R and S are **isomorphic.***

Example 6.5 (Conjugation isomorphism). Let $\varphi:\mathbb{C}\to\mathbb{C}$ be defined by $\varphi(z)=\bar{z}$ where for $z=x+iy$, $\bar{z}:=x-iy$ is the complex conjugate of z. Then it is routine to check that φ is a ring isomorphism. Notice that $z=\bar{z}$ iff $z\in\mathbb{R}$. There is analogous conjugation isomorphism on $\mathbb{Q}[i]$, $\mathbb{Z}[i]$, and $\mathbb{Z}_m[i]$ (for $m\in\mathbb{Z}_+$) with similar properties.

Here is another example in the same spirit of the last example.

Example 6.6 (Another conjugation isomorphism). Let $\varphi:\mathbb{Q}\left[\sqrt{2}\right]\to\mathbb{Q}\left[\sqrt{2}\right]$ be defined by

$$\varphi\left(a+b\sqrt{2}\right)=a-b\sqrt{2}\text{ for all }a,b\in\mathbb{Q}.$$

Then φ is a ring isomorphism. Again this is routine to check. For example,

$$\begin{aligned}\varphi\left(a+b\sqrt{2}\right)\varphi\left(u+v\sqrt{2}\right)&=\left(a-b\sqrt{2}\right)\left(u-v\sqrt{2}\right)\\&=au+2bv-(av+bu)\sqrt{2}\end{aligned}$$

while

$$\begin{aligned}\varphi\left(\left(a+b\sqrt{2}\right)\left(u+v\sqrt{2}\right)\right)&=\varphi\left(au+2bv+(av+bu)\sqrt{2}\right)\\&=au+2bv-(av+bu)\sqrt{2}.\end{aligned}$$

Notice that for $\xi\in\mathbb{Q}\left[\sqrt{2}\right]$, $\varphi(\xi)=\xi$ iff $\xi\in\mathbb{Q}$.

Example 6.7. The only ring homomorphisms, $\varphi:\mathbb{Z}\to\mathbb{Z}$ are $\varphi(a)=a$ and $\varphi(a)=0$ for all $a\in\mathbb{Z}$. Indeed, if $\varphi:\mathbb{Z}\to\mathbb{Z}$ is a ring homomorphism and $t:=\varphi(1)$, then $t^2=\varphi(1)\varphi(1)=\varphi(1\cdot 1)=\varphi(1)=t$. The only solutions to $t^2=t$ in $\mathbb{Z}$ are $t=0$ and $t=1$. In the first case $\varphi\equiv 0$ and in the second $\varphi=id$.

Lecture 7

Example 7.1. Suppose that $g \in M_2(\mathbb{R})$ is a unit, i.e. g^{-1} exists. Then $\varphi : M_2(\mathbb{R}) \to M_2(\mathbb{R})$ defined by,

$$\varphi(A) := gAg^{-1} \text{ for all } A \in M_2(\mathbb{R}),$$

is a ring isomorphism. For example,

$$\varphi(A)\varphi(B) = \left(gAg^{-1}\right)\left(gBg^{-1}\right) = gAg^{-1}gBg^{-1} = gABg^{-1} = \varphi(AB).$$

Observe that $\varphi^{-1}(A) = g^{-1}Ag$ and $\varphi(I) = I$.

Proposition 7.2 (Homomorphisms from $\mathbb{Z}$). *Suppose that R is a ring and $a \in R$ is an element such that $a^2 = a$. Then there exists a unique ring homomorphism, $\varphi : \mathbb{Z} \to R$ such that $\varphi(1) = a$. Moreover, $\varphi(k) = k \cdot a$ for all $k \in \mathbb{Z}$.*

Proof. Recall from last quarter that, $\varphi(n) := n \cdot a$ for all $n \in \mathbb{Z}$ is a group homomorphism. This is also a ring homomorphism since,

$$\varphi(m)\varphi(n) = (m \cdot a)(n \cdot a) = mn \cdot a^2 = mn \cdot a = \varphi(mn),$$

wherein we have used Corollary 5.6 for the second equality. ■

Corollary 7.3. *Suppose that R is a ring with $1_R \in R$. Then there is a unique homomorphism, $\varphi : \mathbb{Z} \to R$ such that $\varphi(1_{\mathbb{Z}}) = 1_R$.*

Proposition 7.4. *Suppose that $\varphi : R \to S$ is a ring homomorphism. Then;*

1. *$\varphi(0) = 0$,*
2. *$\varphi(-r) = -\varphi(r)$ for all $r \in R$,*
3. *$\varphi(r_1 - r_2) = \varphi(r_1) - \varphi(r_2)$ for all $r_1, r_2 \in R$.*
4. *If $1_R \in R$ and φ is surjective, then $\varphi(1_R)$ is an identity in S.*
5. *If $\varphi : R \to S$ is an isomorphism of rings, then $\varphi^{-1} : S \to R$ is also a isomorphism.*

Proof. Noting that $\varphi : (R, +) \to (S, +)$ is a group homomorphism, it follows that items 1. – 3. were covered last quarter when we studied groups. The proof of item 5. is similar to the analogous statements for groups and hence will be omitted. So let me prove item 4. here.

To each $s \in S$, there exists $a \in R$ such that $\varphi(a) = s$. Therefore,

$$\begin{gathered}\varphi(1_R)s = \varphi(1_R)\varphi(a) = \varphi(1_R a) = \varphi(a) = s \\ \text{and} \\ s\varphi(1_R) = \varphi(a)\varphi(1_R) = \varphi(a1_R) = \varphi(a) = s.\end{gathered}$$

Since these equations hold for all $s \in S$, it follows that $\varphi(1_R)$ is an (the) identity in S. ■

Definition 7.5. *As usual, if $\varphi : R \to S$ is a ring homomorphism we let*

$$\ker(\varphi) := \{r \in R : \varphi(r) = 0\} = \varphi^{-1}(\{0_S\}) \subset R.$$

Lemma 7.6. *If $\varphi : R \to S$ is a ring homomorphism, then $\ker(\varphi)$ is an ideal of R.*

Proof. We know from last quarter that $\ker(\varphi)$ is a subgroup of $(R, +)$. If $r \in R$ and $n \in \ker(\varphi)$, then

$$\begin{aligned}\varphi(rn) &= \varphi(r)\varphi(n) = \varphi(r)0 = 0 \text{ and} \\ \varphi(nr) &= \varphi(n)\varphi(r) = 0\varphi(r) = 0,\end{aligned}$$

which shows that rn and $nr \in \ker(\varphi)$ for all $r \in R$ and $n \in \ker(\varphi)$. ■

Example 7.7. Let us find all of the ring homomorphisms, $\varphi : \mathbb{Z} \to \mathbb{Z}_{10}$ and their kernels. To do this let $t := \varphi(1)$. Then $t^2 = \varphi(1)\varphi(1) = \varphi(1 \cdot 1) = \varphi(1) = t$. The only solutions to $t^2 = t$ in $\mathbb{Z}_{10}$ are $t = 0$, $t = 1$, $t = 5$ and $t = 6$.

1. If $t = 0$, then $\varphi \equiv 0$ and $\ker(\varphi) = \mathbb{Z}$.
2. If $t = 1$, then $\varphi(x) = x \bmod 10$ and $\ker\varphi = \langle 10 \rangle = \langle 0 \rangle = \{0\} \subset \mathbb{Z}$.
3. If $t = 5$, then $\varphi(x) = 5x \bmod 10$ and $x \in \ker\varphi$ iff $10|5x$ iff $2|x$ so that $\ker(\varphi) = \langle 2 \rangle = \{0, 2, 4, 8\}$.
4. If $t = 6$, then $\varphi(x) = 6x \bmod 10$ and $x \in \ker\varphi$ iff $10|6x$ iff $5|x$ so that $\ker(\varphi) = \langle 5 \rangle = \{0, 5\} \subset \mathbb{Z}$.

Proposition 7.8. *Suppose $n \in \mathbb{Z}_+$, R is a ring, and $a \in R$ is an element such that $a^2 = a$ and $n \cdot a = 0$. Then there is a unique homomorphism, $\varphi : \mathbb{Z}_n \to R$ such that $\varphi(1) = a$ and in fact $\varphi(k) = k \cdot a$ for all $k \in \mathbb{Z}_n$.*

Proof. This has a similar proof to the proof of Proposition 7.2. ■

Corollary 7.9. *Suppose that R is a ring, $1_R \in R$, and $\operatorname{chr}(R) = n \in \mathbb{Z}_+$. Then there is a unique homomorphism, $\varphi : \mathbb{Z}_n \to R$ such that $\varphi(1_{\mathbb{Z}_n}) = 1_R$ which is given by $\varphi(m) = m \cdot 1_R$ for all $m \in \mathbb{Z}_n$. Moreover, $\ker(\varphi) = \langle 0 \rangle = \{0\}$.*

Example 7.10. Suppose that $\varphi : \mathbb{Z}_{10} \to \mathbb{Z}_{10}$ is a ring homomorphism and $t := \varphi(1)$. Then $t^2 = \varphi(1)^2 = \varphi(1) = t$, and therefore $t^2 = t$. Moreover we must have $0 = \varphi(0) = \varphi(10 \cdot 1) = 10 \cdot t$ which is not restriction on t. As we have seen the only solutions to $t^2 = t$ in $\mathbb{Z}_{10}$ are $t = 0$, $t = 1$, $t = 5$ and $t = 6$. Thus φ must be one of the following; $\varphi \equiv 0$, $\varphi = id$, $\varphi(x) = 5x$, or $\varphi(x) = 6x$ for all $x \in \mathbb{Z}_{10}$. The only ring isomorphism is the identity in this case. If $\varphi(x) = 5x$

Example 7.11. Suppose that $\varphi : \mathbb{Z}_{12} \to \mathbb{Z}_{10}$ is a ring homomorphism and let $t := \varphi(1)$. Then as before, $t^2 = t$ and this forces $t = 0, 1, 5$, or 6. In this case we must also require $12 \cdot t = 0$, i.e. $10 | 12 \cdot t$, i.e. $5 | t$. Therefore we may now only take $t = 0$ or $t = 5$, i.e.

$$\varphi(x) = 0 \text{ for all } x \in \mathbb{Z}_{12} \text{ or}$$
$$\varphi(x) = 5x \bmod 10 \text{ for all } x \in \mathbb{Z}_{12}$$

are the only such homomorphisms.

Theorem 7.12 (Not covered in class). *If $\varphi : \mathbb{R} \to \mathbb{R}$ is a ring homomorphism, then φ is either the zero or the identity homomorphism.*

Proof. If $t = \varphi(1)$, then as above, $t^2 = t$, i.e. $t(t-1) = 0$. Since $\mathbb{R}$ is a field this implies that $t = 0$ or $t = 1$. If $t = 0$, then for all $a \in \mathbb{R}$,

$$\varphi(a) = \varphi(a \cdot 1) = \varphi(a)\varphi(1) = \varphi(a) \cdot 0 = 0,$$

i.e. φ is the zero homomorphism. So we may now assume that $t = 1$.

If $t = 1$,

$$\varphi(n) = \varphi(n \cdot 1) = n \cdot \varphi(1) = n \cdot 1 = n$$

for all $n \in \mathbb{Z}$. Therefore for $n \in \mathbb{N} \setminus \{0\}$ and $m \in \mathbb{Z}$,

$$m = \varphi(m) = \varphi\left(n \cdot \frac{m}{n}\right) = \varphi(n)\varphi\left(\frac{m}{n}\right) = n\varphi\left(\frac{m}{n}\right)$$

from which it follows that $\varphi(m/n) = m/n$. Thus we now know that $\varphi|_{\mathbb{Q}}$ is the identity.

Since $\ker(\varphi) \neq \mathbb{R}$, we must have $\ker(\varphi) = \{0\}$ so that φ is injective. In particular $\varphi(b) \neq 0$ for all $b \neq 0$. Moreover if $a > 0$ in $\mathbb{R}$ and $b := \sqrt{a}$, then

$$\varphi(a) = \varphi(b^2) = [\varphi(b)]^2 > 0.$$

So if $y, x \in \mathbb{R}$ with $y > x$, then $\varphi(y) - \varphi(x) = \varphi(y - x) > 0$, i.e. φ is order preserving.

Finally, let $a \in \mathbb{R}$ and choose rational numbers $x_n, y_n \in \mathbb{Q}$ such that $x_n < a < y_n$ with $x_n \uparrow a$ and $y_n \downarrow a$ as $n \to \infty$. Then

$$x_n = \varphi(x_n) < \varphi(a) < \varphi(y_n) = y_n \text{ for all } n.$$

Letting $n \to \infty$ in this last equation then shows, $a \leq \varphi(a) \leq a$, i.e. $\varphi(a) = a$. Since $a \in \mathbb{R}$ was arbitrary, we may conclude that φ is the identity map on $\mathbb{R}$. ■

8

Lecture 8

Remark 8.1 (Comments on ideals). Let me make two comments on ideals in a commutative ring, R.

1. To check that a non-empty subset, $S \subset R$, is an ideal, we should show $(S, +)$ is a subgroup of R and that $RS \subset S$. Since R is commutative, you do not have to also show $SR \subset S$. This is because $RS = SR$ in when R is commutative.
2. If $a \in R$, the **principle ideal generated by** a is defined by;

$$\langle a \rangle := Ra = \{ra : r \in R\}.$$

It is easy to check that this is indeed an ideal. So for example if $R = \mathbb{R}[x]$ then $\langle x \rangle = \mathbb{R}[x] \cdot x$ which is the same as the polynomials without a constant term, i.e. $p(x) = a_1 x + a_2 x^2 + \cdots + a_n x^n$. The coefficient $a_0 = 0$. Similarly, $\left\langle x^2+1 \right\rangle = \mathbb{R}[x]\left(x^2+1\right)$ is the collection of all polynomials which contain $\left(x^2+1\right)$ as a factor.

Recall from last time:

1. If $a \in R$ satisfies $a^2 = a$, then $\varphi(k) := k \cdot a$ is a ring homomorphism from $\mathbb{Z} \to R$.
2. If we further assume that $n \cdot a = 0$ for some $n \in \mathbb{Z}_+$, then $\varphi(k) := k \cdot a$ also defines a ring homomorphism from $\mathbb{Z}_n \to R$.

Example 8.2. For any $m > 1$, $\varphi : \mathbb{Z} \to \mathbb{Z}_m$ given by $\varphi(a) = a \cdot 1_{\mathbb{Z}_m} = a \bmod m$ is a ring homomorphism. This also follows directly from the properties of the $(\cdot) \bmod m-$ function. In this case $\ker(\varphi) = \langle m \rangle = \mathbb{Z}m$.

Example 8.3. If $n \in \mathbb{Z}_+$ and $m = kn$ with $k \in \mathbb{Z}_+$, then there is a unique ring homomorphisms, $\varphi : \mathbb{Z}_m \to \mathbb{Z}_n$ such that $\varphi(1_m) = 1_n$. To be more explicit,

$$\varphi(a) = \varphi(a \cdot 1_m) = a \cdot \varphi(1_m) = a \cdot 1_n = (a \bmod n) \cdot 1_n = a \bmod n.$$

Example 8.4. In $\mathbb{Z}_{10}$, the equation, $a^2 = a$ has a solutions $a = 5$ and $a = 6$. Notice that $|5| = 2$ and $|6| = |\gcd(10, 6)| = |2| = 5$. Thus we know that for any $k \geq 1$ there are ring homomorphisms, $\varphi : \mathbb{Z}_{5k} \to \mathbb{Z}_{10}$ and $\psi : \mathbb{Z}_{2k} \to \mathbb{Z}_{10}$ such that

$$\varphi(1_{5k}) = 6 \text{ and } \psi(1_{2k}) = 5.$$

As before, one shows that

$$\varphi(m) = m \cdot 6 = (6m) \bmod 10 \text{ and } \psi(m) = m \cdot 5 = (5m) \bmod 10.$$

Example 8.5 (Divisibility tests). Let $n = a_k a_{k-1} \dots a_0$ be written in decimal form, so that

$$n = \sum_{i=0}^{k} a_i 10^i. \tag{8.1}$$

Applying the ring homomorphism, $\bmod 3$ and $\bmod 9$ to this equation shows,

$$\begin{aligned} n \bmod 3 &= \sum_{i=0}^{k} a_i \bmod 3 \cdot (10 \bmod 3)^i \\ &= \left(\sum_{i=0}^{k} a_i\right) \bmod 3 \end{aligned}$$

and similarly,

$$n \bmod 9 = \sum_{i=0}^{k} a_i \bmod 9 \cdot (10 \bmod 9)^i = \left(\sum_{i=0}^{k} a_i\right) \bmod 9.$$

Thus we learn that $n \bmod 3 = 0$ iff $\left(\sum_{i=0}^{k} a_i\right) \bmod 3 = 0$ i.e. $3|n$ iff $3|\left(\sum_{i=0}^{k} a_i\right)$. Similarly, since $10 \bmod 9 = 1$, the same methods show $9|n$ iff $9|\left(\sum_{i=0}^{k} a_i\right)$. (See the homework problems for more divisibility tests along these lines. Also consider what this test gives if you apply $\bmod 2$ to Eq. (8.1).)

Theorem 8.6. *Let R be a commutative ring with $1 \in R$. To each $a \in R$ with $a^2 + 1 = 0$, there is a unique ring homomorphism $\varphi : \mathbb{Z}[i] \to R$ such that $\varphi(1) = 1$ and $\varphi(i) = a$.*

Proof. Since $\mathbb{Z}[i]$ is generated by i, we see that φ is completely determined by $a := \varphi(i) \in R$. Now we can not choose a arbitrarily since we must have

$$a^2 = \varphi(i)^2 = \varphi\left(i^2\right) = \varphi(-1) = -1_R,$$

i.e. $a^2+1=0$.

Conversely given $a \in R$ such that $a^2+1=0$, we should define

$$\varphi(x+iy) = x1 + ya \text{ for all } x,y \in \mathbb{Z},$$

where $ya = a + a + \cdots + a - y$ times. The main point in checking that φ is a homomorphism is to show it preserves the multiplication operation of the rings. To check this, let $x,y,u,v \in \mathbb{Z}$ and consider;

$$\varphi((x+iy)(u+iv)) = \varphi(xu - yv + i(xv+yu)) = (xu-yv)1_R + (xv+yu)a.$$

On the other hand

$$\begin{aligned}\varphi(x+iy)\varphi(u+iv) &= (x1_R + ya)(u1_R + va) \\ &= (x1_R + ya)(u1_R + va) \\ &= xu1_R + yva^2 + yua + xva \\ &= (xu-yv)1_R + (yu+xv)a \\ &= \varphi((x+iy)(u+iv)).\end{aligned}$$

Thus we have shown $\varphi(\xi\eta) = \varphi(\xi)\varphi(\eta)$ for all $\xi,\eta \in \mathbb{Z}[i]$. The fact that $\varphi(\xi+\eta) = \varphi(\xi)+\varphi(\eta)$ is easy to check and is left to the reader. ■

Remark 8.7. This could be generalized by supposing that $a,b \in R$ with $b^2 = b$ and $a^2 + b = 0$. Then we would have $\varphi(x+yi) = x\cdot b + y \cdot a$ would be the desired homomorphism. Indeed, let us observe that

$$\begin{aligned}\varphi(x+iy)\varphi(u+iv) &= (xb+ya)(ub+va) \\ &= (xb+ya)(ub+va) \\ &= xub^2 + yva^2 + yua + xva \\ &= (xu-yv)b + (yu+xv)a \\ &= \varphi((x+iy)(u+iv)).\end{aligned}$$

Example 8.8. Let $\varphi : \mathbb{Z}[i] \to \mathbb{Z}_3[i]$ be the unique homomorphism such that $\varphi(1) = 1$ and $\varphi(i) = i$, i.e.

$$\varphi(a+ib) = a\cdot 1 + b \cdot i = a \bmod 3 + (b \bmod 3)\, i \in \mathbb{Z}_3[i].$$

Notice that

$$\ker(\varphi) = \{a+bi : a,b \in \langle 3\rangle \subset \mathbb{Z}\} = \langle 3\rangle + \langle 3\rangle i.$$

Here is a more interesting example.

Example 8.9. In $\mathbb{Z}_{10}$ we observe that $3^2 = 9 = -1$ and also $7 = -3$ has this property, namely $7^2 = (-3)^2 = 3^2 = 9 = -1$. Therefore there exists a unique homomorphism, $\varphi : \mathbb{Z}[i] \to \mathbb{Z}_{10}$ such that $\varphi(1) = 1$ and $\varphi(i) = 7 = -3$. The explicit formula is easy to deduce,

$$\varphi(a+bi) = a\cdot 1 + b \cdot 7 = (a-3b) \bmod 10.$$

9

Lecture 9

Lemma 9.1. *If $\varphi : R \to S$ is a ring homomorphism, then* $\ker(\varphi)$ *is an ideal of* R.

Proof. We know from last quarter that $\ker(\varphi)$ is a subgroup of $(R,+)$. If $r \in R$ and $n \in \ker(\varphi)$, then

$$\begin{aligned} \varphi(rn) &= \varphi(r)\varphi(n) = \varphi(r)0 = 0 \text{ and} \\ \varphi(nr) &= \varphi(n)\varphi(r) = 0\varphi(r) = 0, \end{aligned}$$

which shows that rn and $nr \in \ker(\varphi)$ for all $r \in R$ and $n \in \ker(\varphi)$. ■

Example 9.2. If $\varphi : \mathbb{Z} \to \mathbb{Z}_m$ is the ring homomorphism defined by $\varphi(a) := a \bmod m$, then

$$\ker\varphi = \{a \in \mathbb{Z} : a \bmod m = 0\} = \mathbb{Z}m = \langle m \rangle.$$

We will see many more examples of Lemma 9.1 below.

9.1 Factor Rings

Definition 9.3. *Let R be a ring, $I \subset R$ an ideal. The factor ring R/I is defined to be*

$$R/I := \{r + I : r \in R\}$$

with operations

$$\begin{aligned} (a+I) + (b+I) &:= (a+b) + I \text{ and} \\ (a+I)(b+I) &:= (ab) + I. \end{aligned}$$

We may also write $[a]$ for $a+I$ in which cases the above equations become,

$$[a] + [b] := [a+b] \text{ and } [a][b] := [ab].$$

Theorem 9.4. *A factor ring really is a ring.*

Proof. The elements of R/I are the left cosets of I in the group $(R,+)$. There is nothing new here. R/I is itself a group with the operation $+$ defined by $(a+I) + (b+I) = (a+b) + I$. This follows from last quarter as $I \subset R$ is a normal subgroup of $(R,+)$ since $(R,+)$ is abelian. So we only need really to check that the definition of product makes sense.

Problem: we are multiplying coset representatives. We have to check that the resulting coset is independent of the choice of representatives. Thus we need to show; if $a, b, a', b' \in R$ with

$$a + I = a' + I \text{ and } b + I = b' + I,$$

then $ab + I = a'b' + I$. By definition of cosets, we have $i := a - a' \in I$ and $j := b - b' \in I$. Therefore,

$$ab = (a' + i)(b' + j) = a'b' + ib' + a'j + ij \in a'b' + I$$

since $ib' + a'j + ij \in I$ because I is an ideal. So indeed, $ab + I = a'b' + I$ and we have a well defined product on R/I. Checking that product is associative and the distributive laws is easy and will be omitted. ■

Example 9.5. Suppose that $I = \langle 4 \rangle = \mathbb{Z} \cdot 4 \subset \mathbb{Z}$. In this case, if $a \in \mathbb{Z}$ then $a - a \bmod 4 \in I$ and therefore,

$$[a] = a + I = a \bmod 4 + I = [a \bmod 4].$$

Moreover if $0 \le a, b \le 3$ with $a + I = b + I$ then $a - b \in I$, i.e. $a - b$ is a multiple of 4. Since $|a - b| < 4$, this is only possible if $a = b$. Thus if we let $\mathcal{S} = \{0, 1, 2, 3\}$, then

$$\mathbb{Z}/\langle 4 \rangle = \{[m] = m + \langle 4 \rangle : m \in \mathcal{S}\} = [\mathcal{S}].$$

Moreover, we have

$$[a][b] = [ab] = [(ab) \bmod 4]$$

and

$$[a] + [b] = [a+b] = [(a+b) \bmod 4].$$

Thus the induced ring structure on $\mathcal{S}$ is precisely that of $\mathbb{Z}_4$ and so we may conclude;

$$\mathbb{Z}_4 \ni a \to [a] = a + \langle 4 \rangle \in \mathbb{Z}/\langle 4 \rangle$$

is a ring isomorphism.

Lecture 10

Remark 10.1. Roughly speaking, you should think of R/I being R with the proviso that we identify two elements of R to be the same if they differ by an element from I. To understand R/I in more concrete terms, it is often useful to look for subset, $\mathcal{S} \subset R$, such that the map,

$$\mathcal{S} \ni a \to a + I \in R/I$$

is a bijection. (We will call such an $\mathcal{S}$ a slice.) This allows us to identify R/I with $\mathcal{S}$ and this identification induces a ring structure on $\mathcal{S}$. We will see how this goes in the examples below. **Warning:** the choice of a slice $\mathcal{S}$ is highly non-unique although there is often a "natural" choice in a given example. The point is to make $\mathcal{S}$ we need only choose one element from each of the cosets in R/I.

Example 9.5 easily generalizes to give the following theorem. We will give another proof shortly using the first isomorphism theorem, see 10.4 below.

Theorem 10.2 ($\mathbb{Z}_m \cong \mathbb{Z}/\langle m\rangle$). *For all $m \geq 2$, the map,*

$$\mathbb{Z}_m \ni a \to [a] = a + \langle m\rangle \in \mathbb{Z}/\langle m\rangle \tag{10.1}$$

is a ring isomorphism.

Proof. The distinct cosets of $\mathbb{Z}/\langle m\rangle$ are given by

$$\{[k] = k + \langle m\rangle : k = 0, 1, 2 \ldots, m-1\}$$

and therefore we may take $\mathcal{S} = \mathbb{Z}_m$. Since $[a] = [a \bmod m]$, it is easy to see that the map in Eq. (10.1) is a ring isomorphism. ■

10.1 First Isomorphism Theorem

Recall that two rings, R and S (written $R \cong S$) are isomorphic, if there is a ring isomorphism, $\varphi : R \to S$. That is φ should be a one-to-one and onto ring homomorphism.

Theorem 10.3 (First Isomorphism Theorem). *Let R and S be rings and $\varphi : R \to S$ be a homomorphism. Let*

$$\varphi(R) = \operatorname{Ran} \varphi = \{\varphi(r) : r \in R\} \subset S$$

and recall that $I = \ker \varphi := \{r \in R : \varphi(r) = 0\}$ is an ideal in R. Then $\varphi(R)$ is a subring of S and $\bar{\varphi} : R/I \to \varphi(R)$ defined by

$$\bar{\varphi}([r]) = \bar{\varphi}(r + I) := \varphi(r) \text{ for all } r \in R$$

is a ring isomorphism.

Proof. We have seen last quarter that $\bar{\varphi} : R/\ker\varphi \to \varphi(R)$ is an (additive) group isomorphism. So it only remains to show $\bar{\varphi}$ preserves the multiplication operations on $\varphi(R)$ and R/I which goes as follows;

$$\begin{aligned}\bar{\varphi}([a])\,\bar{\varphi}([b]) &= \varphi(a)\,\varphi(b) \\ &= \varphi(ab) = \bar{\varphi}([ab]) = \bar{\varphi}([a]\,[b]).\end{aligned}$$

■

Example 10.4 ($\mathbb{Z}/(\mathbb{Z}m) \cong \mathbb{Z}_m$). Let $m \in \mathbb{Z}_+$ and $\varphi : \mathbb{Z} \to \mathbb{Z}_m$ be the ring homomorphism, $\varphi(x) = x \bmod m$. Since $\varphi(\mathbb{Z}) = \mathbb{Z}_m$ and $\ker(\varphi) = \langle m\rangle = \mathbb{Z}m$, the first isomorphism theorem implies, $\bar{\varphi} : \mathbb{Z}/(\mathbb{Z}m) \to \mathbb{Z}_m$ is a ring isomorphism where $\bar{\varphi}([a]) = \varphi(a) = a \bmod m$ for all $a \in \mathbb{Z}$.

Example 10.5. Let us consider $R := \mathbb{Z}[i]/\langle i-2\rangle$. In this ring $[i-2] = 0$ or equivalently, $[i] = [2]$. Squaring this equation also shows,

$$[-1] = \left[i^2\right] = [i]^2 = [2]^2 = \left[2^2\right] = [4]$$

from which we conclude that $[5] = 0$, i.e. $5 \in \langle i-2\rangle$. This can also be seen directly since $5 = -(i+2)(i-2) \in \langle i-2\rangle$. Using these observations we learn for $a + ib \in \mathbb{Z}[i]$ that

$$[a + ib] = [a + 2b] = [(a + 2b) \bmod 5].$$

Thus, if we define $\mathcal{S} = \{0, 1, 2, 3, 4\}$, we have already shown that

$$R = \{[a] : a \in \mathcal{S}\} = [\mathcal{S}] \,.$$

Now suppose that $a, b \in \mathcal{S}$ with $[a] = [b]$, i.e. $0 = [a - b] = [c]$ where $c = (a - b) \bmod 5$. Since $c \in \langle i - 2 \rangle$ we must have

$$c = (i - 2)(a + bi) = -(2a + b) + (a - 2b)\, i$$

from which it follows that $a = 2b$ and

$$c = -(2a + b) = -5b.$$

Since $0 \leq c < 5$, this is only possible if $c = 0$ and therefore,

$$a = a \bmod 5 = b \bmod 5 = b.$$

Finally let us now observe that

$$\begin{aligned} [a] + [b] &= [a + b] = [(a + b) \bmod 5] \text{ and} \\ [a] \cdot [b] &= [ab] = [(ab) \bmod 5] \end{aligned}$$

so that the induced ring structure on $\mathcal{S}$ is the same a the ring structure on $\mathbb{Z}_5$. Hence we have proved,

$$\mathbb{Z}_5 \ni a \rightarrow [a] = a + \langle i - 2 \rangle \in \mathbb{Z}[i] / \langle i - 2 \rangle$$

is an isomorphism of rings.

Lecture 11

Example 11.1 (Example 10.5 revisited). In $\mathbb{Z}_5$, we see that $2^2 = 4 = -1$ and therefore there is a ring homomorphism, $\varphi : \mathbb{Z}[i] \to \mathbb{Z}_5$ such that $\varphi(1) = 1$ and $\varphi(i) = 2$. More explicitly we have,

$$\varphi(a + bi) = a \cdot 1 + b \cdot 2 = (a + 2b) \bmod 5.$$

Moreover, $(a + ib) \in \ker(\varphi)$ iff $a + 2b = 5k$ for some $k \in \mathbb{Z}$ and therefore,

$$\ker(\varphi) = \{-2b + 5k + ib : b, k \in \mathbb{Z}\} = \mathbb{Z}(2 - i) + \mathbb{Z} \cdot 5.$$

Since $(2 + i)(2 - i) = 5$ and $2 - i \in \ker(\varphi)$, we have, and

$$\langle 2 - i \rangle \subset \ker(\varphi) = \mathbb{Z}(2 - i) + \mathbb{Z} \cdot 5 \subset \langle 2 - i \rangle$$

from which it follows that $\ker(\varphi) = \langle 2 - i \rangle$. Thus by the first isomorphism theorem, $\bar{\varphi} : \mathbb{Z}[i] / \langle 2 - i \rangle \to \mathbb{Z}_5$ defined by

$$\bar{\varphi}([a + ib]) = \varphi(a + bi) = (a + 2b) \bmod 5$$

is a ring isomorphism. Notice that the inverse isomorphism is given by $\bar{\varphi}^{-1}(a) = [a]$ for all $a \in \mathbb{Z}_5$ which should be compared with Example 10.5 above.

For what follows recall that the evaluation maps are homomorphisms.

Theorem 11.2 (Evaluation homomorphism). *Let R be a subring of a commutative ring, $\bar{R}$, and $t \in \bar{R}$. Then there exists a ring homomorphism, $\varphi_t : R[x] \to \bar{R}$ such that*

$$\varphi_t(p) = \sum_{k=0}^{n} a_k t^k \text{ when } p(x) = \sum_{k=0}^{n} a_k x^k \in R[x].$$

We will usually simply write $p(t)$ for $\varphi_t(p)$.

The hole point of how we define polynomial multiplication is to make this theorem true. We will give the formal proof of this theorem a bit later in the notes.

Example 11.3. Let $I := \langle x \rangle = \mathbb{R}[x]\,x \subset \mathbb{R}[x]$ from which it follows that $[x] = 0 \in \mathbb{R}[x] / \langle x \rangle$. Therefore if $p(x) = a_0 + a_1 x + \dots + a_n x^n$, then

$$[p(x)] = [a_0 + a_1 x + \dots + a_n x^n] = [a_0].$$

Alternatively put, $p(x) + I = a_0 + I$ since $a_1 x + \dots + a_n x^n \in I$. Moreover, if $[a_0] = [b_0]$, then $a_0 - b_0 \in I$ which can happen iff $a_0 = b_0$. Therefore we may identify $\mathbb{R}[x] / \langle x \rangle$ with $\mathcal{S} = \mathbb{R}$ thought of as the constant polynomials inside of $\mathbb{R}[x]$. In fact it is easy to check that

$$\mathbb{R} \ni a \to a + I \in \mathbb{R}[x] / \langle x \rangle$$

is a ring isomorphism.

Alternatively we may use the first isomorphism theorem as follows. Let $\varphi(p) := p(0)$, then $\varphi : \mathbb{R}[x] \to \mathbb{R}$ is a ring homomorphism onto $\mathbb{R}$ with $\ker(\varphi) = \langle x \rangle$. Therefore, $\bar{\varphi} : \mathbb{R}[x] / \langle x \rangle \to \mathbb{R}$ is a ring isomorphism.

Theorem 11.4 (Division Algorithm). *Let $F[x]$ be a polynomial ring where F is a field. Given $f, g \in F[x]$ both nonzero, there exists a unique $q, r \in F[x]$ with $f = qg + r$ such that either $r = 0$ or $\deg r < \deg g$.*

Interpretation. We are dividing f by g and so g **goes into** f, q **times with remainder** r. This is really high school polynomial division which we will discuss in more detail a bit later. In the sequel we will sometimes denote the remainder, r by $f \bmod g$.

Corollary 11.5. *Suppose that F is a field, $p(x) = c_0 + \dots + c_n x^n \in F[x]$ is a polynomial with $c_n \neq 0$, and let*

$$\mathcal{S} := \left\{a_0 + a_1 x + \dots + a_{n-1} x^{n-1} : a_i \in F \text{ for } i = 0, 1, \dots, n-1\right\}.$$

Then the map, $\varphi : \mathcal{S} \to F[x] / \langle p \rangle$ defined by

$$\varphi(f) = [f] := f + \langle p \rangle \text{ for all } f \in \mathcal{S}$$

is a bijection. Moreover, $\mathcal{S}$ becomes a ring and φ a ring homomorphism provided we define

$$f(x) \cdot g(x) := [f(x)\, g(x)] \bmod p$$

and $f + g$ as usual polynomial addition.

Proof. 1. If $f \in F[x]$, then by the division algorithm

$$f = qp + r = qp + f \bmod p$$

and therefore,

$$[f] = [qp + r] = [q][p] + [r] = [q]\,0 + [r] = [r].$$

Thus we have shown

$$[f] = [f \bmod p] \text{ for all } f \in F[x]. \tag{11.1}$$

2. Equation (11.1) shows $\varphi : \mathcal{S} \to F[x] / \langle p \rangle$ is onto. To see φ is injective, suppose that $f, g \in \mathcal{S}$ and $\varphi(f) = \varphi(g)$. Then $[f - g] = 0$, i.e. $f - g \in \langle p \rangle$, i.e. $f - g = q \cdot p$ for some $q \in F[x]$. However this is impossible unless $q = 0$ and $f = g$ since otherwise,

$$n - 1 \geq \deg(f - g) = \deg(q) + \deg(p) = \deg(q) + n.$$

Thus we have shown φ is injective as well, i.e. $\varphi : \mathcal{S} \to F[x] / \langle p \rangle$ is a bijection.

3. Making use of Eq. (11.1) and the fact that φ is a bijection shows,

$$\begin{aligned} \varphi(f)\varphi(g) &= [f][g] = [fg] = [(fg) \bmod p] = \varphi((fg) \bmod p) \text{ and} \\ \varphi(f) + \varphi(g) &= [f] + [g] = [f + g] = \varphi(f + g) \end{aligned}$$

for all $f, g \in \mathcal{S}$. Thus $\mathcal{S}$ equipped with the operations described in the theorem makes $\mathcal{S}$ into a ring for which φ is a ring isomorphism. ■

Theorem 11.6 ($\mathbb{C}$ as a factor ring). $\mathbb{C} \cong \mathbb{R}[x] / \left\langle x^2 + 1 \right\rangle =: R$. *The maps,*

$$\begin{aligned} \mathbb{C} \ni (a + ib) &\to [a + bx] \in \mathbb{R}[x] / \left\langle x^2 + 1 \right\rangle \text{ and} \\ \mathbb{R}[x] / \left\langle x^2 + 1 \right\rangle \ni [p(x)] = p(x) + \left\langle x^2 + 1 \right\rangle &\to p(i) \in \mathbb{C} \end{aligned}$$

are ring isomorphisms which are inverses to one another.

Proof. We are going to give two proofs that $\mathbb{C} \cong \mathbb{R}[x] / \left\langle x^2 + 1 \right\rangle$. Our first proof gives rise to the first map while our second gives rise to the second map.

First Proof. Let $\mathcal{S} = \{a + bx : x, b \in \mathbb{R}\}$ so that

$$\mathcal{S} \ni (a + bx) \to [a + bx] \in \mathbb{R}[x] / \left\langle x^2 + 1 \right\rangle$$

is a bijection. Since $\left[x^2 + 1\right] = 0$, we have $\left[x^2\right] = [-1]$ and therefore,

$$\begin{aligned} [a + bx][c + dx] &= \left[ac + (bc + ad)x + bdx^2\right] \\ &= [ac + (bc + ad)x + bd(-1)] \\ &= [ac - bd + (bc + ad)x]. \end{aligned}$$

Moreover one easily shows,

$$[a + bx] + [c + dx] = [(a + c) + (b + d)x].$$

From these two facts it is now easy to check that

$$\mathbb{C} \ni (a + ib) \to [a + bx] \in R$$

is an isomorphism of rings.

Second Proof. Let $\varphi : \mathbb{R}[x] \to \mathbb{C}$ be the evaluation homomorphism, $\varphi(p) = p(i)$ where $i = \sqrt{-1} \in \mathbb{C}$. We then have $\varphi(\mathbb{R}[x]) = \mathbb{R}[i] = \mathbb{C}$ and so by the first isomorphism theorem, $\mathbb{R}[x] / \ker(\varphi) \cong \mathbb{C}$. So to finish the proof we must show,

$$\ker(\varphi) = \left\langle x^2 + 1 \right\rangle = \mathbb{R}[x]\left(x^2 + 1\right). \tag{11.2}$$

Suppose that $p \in \ker(\varphi)$ and use the division algorithm to write,

$$\begin{aligned} &p(x) = q(x)\left(x^2 + 1\right) + r(x) \text{ where} \\ &r(x) = a + bx \text{ for some } a, b \in \mathbb{R}. \end{aligned}$$

As $p(i) = 0$ and $i^2 + 1 = 0$, it follows that $r(i) = a + bi = 0$. But this happens iff $a = 0 = b$, and therefore we see that $r \equiv 0$ an hence that $p(x) \in \left\langle x^2 + 1 \right\rangle$. Thus we have shown $\ker(\varphi) \subset \left\langle x^2 + 1 \right\rangle$ and since $x^2 + 1 \in \ker(\varphi)$ we must have $\ker(\varphi) = \left\langle x^2 + 1 \right\rangle$ which completes the second proof of the theorem.

Alternative method for computing $\ker(\varphi)$.

If $p \in \ker(\varphi)$, then $p(i) = 0$. Taking the complex conjugates of this equation (using $\overline{z + w} = \bar{z} + \bar{w}$ and $\overline{zw} = \bar{z} \cdot \bar{w}$ for all $z, w \in \mathbb{C}$) we learn that $p(-i) = 0$ as well. As we will see in detail later, $p(i) = 0$ implies $p(x) = (x - i)u(x)$ for some $u \in \mathbb{C}[x]$. Moreover since,

$$0 = p(-i) = -2i \cdot u(-i)$$

we learn that $u(-i) = 0$ and therefore, $u(x) = (x + i)q(x)$ with $q \in \mathbb{C}[x]$. Therefore,

$$p(x) = (x - i)(x + i)q(x) = \left(x^2 + 1\right)q(x).$$

It is not too hard to see (use complex conjugation again) that in fact $q \in \mathbb{R}[x]$. Conversely if $p(x) = \left(x^2 + 1\right)q(x)$ with $q \in \mathbb{R}[x]$, then $p(i) = 0$. Therefore we have again proved Eq. (11.2). ■

12

Lecture 12

Example 12.1. Let $R := \mathbb{Q}[x] / \left\langle x^2 - 2 \right\rangle$ so that $\left[x^2\right] = [2]$ now. Again we take $\mathcal{S} = \{a + bx : a, b \in \mathbb{Q}\}$ and observe that

$$\begin{aligned} [a + bx][c + dx] &= \left[ac + (bc + ad)x + bdx^2\right] \\ &= [ac + (bc + ad)x + bd2] \\ &= [ac + 2bd + (bc + ad)x]. \end{aligned}$$

Recalling that, in $\mathbb{Q}\left[\sqrt{2}\right]$, that

$$\left(a + b\sqrt{2}\right)\left(c + d\sqrt{2}\right) = ac + 2bd + (bc + ad)\sqrt{2}$$

it follows that

$$\mathbb{Q}\left[\sqrt{2}\right] \ni a + b\sqrt{2} \to [a + bx] \in \mathbb{Q}[x] / \left\langle x^2 - 2 \right\rangle$$

is a ring isomorphism.

Example 12.2 (Example 12.1 revisited). Let $\varphi : \mathbb{Q}[x] \to \mathbb{Q}\left[\sqrt{2}\right]$ be the evaluation map, $\varphi(p) = p\left(\sqrt{2}\right)$. Then by the first isomorphism theorem, $\bar{\varphi} : \mathbb{Q}[x] / \ker(\varphi) \to \mathbb{Q}\left[\sqrt{2}\right]$ is an isomorphism of rings. We now claim that

$$\ker(\varphi) = \left\langle x^2 - 2 \right\rangle. \tag{12.1}$$

Since $x^2 - 2 \in \ker(\varphi)$ we know that $\left\langle x^2 - 2 \right\rangle \subset \ker(\varphi)$. Conversely, if $p \in \ker(\varphi)$ and $p(x) = q(x)\left(x^2 - 2\right) + r(x)$ for some $r(x) = a + bx$ with $a, b \in \mathbb{Q}$, then

$$0 = p\left(\sqrt{2}\right) = q\left(\sqrt{2}\right) \cdot 0 + r\left(\sqrt{2}\right) = a + b\sqrt{2}.$$

As $\sqrt{2}$ is irrational, this is only possible if $a = b = 0$, i.e. $r(x) = 0$. Thus we have shown $p \in \left\langle x^2 - 2 \right\rangle$ and therefore Eq. (12.1) is valid.

Example 12.3. Let $I := \left\langle x^2 \right\rangle = \mathbb{R}[x]\,x^2 \subset \mathbb{R}[x]$. If $p(x) = a_0 + a_1x + \cdots + a_nx^n$, then $p + I = a_0 + a_1x + I$ since $a_2x^2 + \cdots + a_nx^n \in I$. Alternatively, we now have $\left[x^2\right] = 0$ in $\mathbb{R}[x] / \left\langle x^2 \right\rangle$, so that

$$[p(x)] = [a_0 + a_1x + \cdots + a_nx^n] = [a_0 + a_1x].$$

Moreover $[a_0 + a_1x] = 0$ iff $a_0 = a_1 = 0$, so we may take $\mathcal{S} = \{a_0 + a_1x : a_0, a_1 \in \mathbb{R}\}$ – the polynomials of degree less than or equal to 1. Thus it follows that

$$\mathbb{R}[x] / \left\langle x^2 \right\rangle = \{(a_0 + a_1x) + I : a_0, a_1 \in \mathbb{R}\} \sim \mathbb{R}^2.$$

This induces a ring multiplication on $\mathbb{R}^2$ determined as follows;

$$\begin{aligned} [a_0 + a_1x][b_0 + b_1x] &= [(a_0 + a_1x)(b_0 + b_1x)] \\ &= \left[a_0b_0 + (a_1b_0 + a_0b_1)x + a_1b_1x^2\right] \\ &= [a_0b_0 + (a_1b_0 + a_0b_1)x]. \end{aligned}$$

Thus the multiplication rule on $\mathcal{S}$ should be defined by

$$(a_0 + a_1x)(b_0 + b_1x) = a_0b_0 + (a_1b_0 + a_0b_1)x.$$

Alternatively, if we identify $\mathcal{S}$ with $R := \mathbb{R}^2$ and equip R with the multiplication and addition rules,

$$\begin{aligned} (a_0, a_1) \cdot (b_0, b_1) &= (a_0b_0, a_1b_0 + a_0b_1) \text{ and} \\ (a_0, a_1) + (b_0, b_1) &= (a_0 + b_0, a_1 + b_1), \end{aligned} \tag{12.2}$$

then

$$R \ni (a_0, a_1) \to a_0 + a_1x + I \in \mathbb{R}[x] / \left\langle x^2 \right\rangle$$

is a ring isomorphism.

An important point to observe for later is that R in Example 12.3 is **not** a field and in fact not even an integral domain. For example, $(0,1) \cdot (0,1) = (0,0) = 0$. Alternatively, notice that $[x] \cdot [x] = \left[x^2\right] = 0$, so that $0 \neq [x] \in \mathbb{R}[x] / \left\langle x^2 \right\rangle$ is a zero divisor.

Example 12.4 (Example 12.3 revisited). We let R be the ring, $\mathbb{R}^2$, with usual addition and the multiplication rule in Eq. (12.2). Let $\varphi : \mathbb{R}[x] \to R$ be the map define by, $\varphi(p) = (p(0), p'(0))$ where $p'(x)$ is the derivative of $p(x)$ computed as usual for polynomials. Then one easily checks that φ is a ring homomorphism. Moreover if $p \in \ker(\varphi)$, then $p(0) = 0$ and therefore $p(x) = xg(x)$ for some polynomial $g(x)$. Since

$$0 = p'(0) = g(0) + 0 \cdot g'(0)$$

it follows that $g(x) = xq(x)$ for some polynomial $q(x)$. Thus $p(x) = x^2 q(x)$. Conversely if $p(x) = x^2 q(x)$, then $p(0) = 0$ and $p'(0) = \left[2xq(x) + x^2 q'(x)\right]_{x=0} = 0$. Therefore we have shown, $\ker(\varphi) = \langle x^2 \rangle$ and so by the first isomorphism theorem, it follows that $\mathbb{R}[x] / \langle x^2 \rangle \ni [p(x)] \to (p(0), p'(0)) \in R$ is a ring isomorphism.

12.1 Higher Order Zeros (Not done in class)

Remark 12.5. Example 12.4 generalizes in the following way. Let $n \in \mathbb{Z}_+$ and $\lambda \in \mathbb{R}$ and define $\varphi : \mathbb{R}[x] \to R := \mathbb{R}^{n+1}$ by

$$\varphi(p) := \left(p(\lambda), p'(\lambda), \dots, p^{(n)}(\lambda)\right) \in R. \tag{12.3}$$

We wish to define $+$ and $\cdot$ on R so that his map is a homomorphism. Since the derivative operation is linear we should use the ordinary vector addition on $\mathbb{R}^{n+1}$ in which case φ will be an additive group homomorphism. For the multiplication rule we have to use the product rule of differentiation in the following form,

$$(pq)^{(k)}(\lambda) = \sum_{j=0}^{k} \binom{k}{j} p^{(j)}(\lambda) q^{(k-j)}(\lambda).$$

Thus if $a = (a_0, \dots, a_n)$ and $b = (b_0, \dots, b_n)$, we should define

$$a \cdot b := \left((a \cdot b)_0, \dots, (a \cdot b)_n\right) \tag{12.4}$$

where

$$(a \cdot b)_k := \sum_{j=0}^{k} \binom{k}{j} a_j \cdot b_{k-j}. \tag{12.5}$$

Theorem 12.6. *Suppose that $R = \mathbb{R}^{n+1}$ with the addition and multiplication operations described in Remark 12.5. Then;*

1. *R is a ring with identity, $1 = (1, 0, \dots, 0)$.*
2. *$\varphi : \mathbb{R}[x] \to R$ defined in Eq. (12.3) is a ring homomorphism.*
3. $\ker(\varphi) = \left\langle (x-\lambda)^{n+1} \right\rangle = \mathbb{R}[x](x-\lambda)^{n+1}$.

Proof. Item 1. can be proved by a straight forward but tedious verification. However there is a better way! Consider the bijective map,

$$R \ni (a_0, \dots, a_n) \xrightarrow{\varphi} \left[\sum_{k=0}^{n} \frac{a_k}{k!} x^k\right] := \sum_{k=0}^{n} \frac{a_k}{k!} x^k + \langle x^{n+1} \rangle \subset \mathbb{R}[x] / \langle x^{n+1} \rangle.$$

Since,

$$\left[\sum_{k=0}^{n} \frac{a_k}{k!} x^k\right] \left[\sum_{l=0}^{n} \frac{b_l}{l!} x^l\right] = \left[\sum_{k=0}^{2n} \left(\sum_{j=0}^{k} \frac{a_j}{j!} \frac{b_{k-j}}{(k-j)!}\right) x^k\right]$$
$$= \left[\sum_{k=0}^{n} \frac{1}{k!} \left(\sum_{j=0}^{k} \binom{k}{j} a_j b_{k-j}\right) x^k\right],$$

that φ becomes a ring homomorphisms provided we use the multiplication rule in Eqs. (12.4) and (12.5).

Item 2. is easy since we defined the ring multiplication on R so that φ would be a homomorphism.

For item 3. let me only explain the case where $n = 1$ here. If $p \in \ker(\varphi)$, then

$$0 = (0, 0) = \varphi(p) = (p(\lambda), p'(\lambda)).$$

Since $p(\lambda) = 0$, we know $p(x) = (x - \lambda) u(x)$ for some $u \in \mathbb{R}[x]$. Differentiating this equation at $x = \lambda$ then implies, $0 = p'(\lambda) = u(\lambda)$ and therefore $u(x) = (x - \lambda) q(x)$ for some $q \in \mathbb{R}[x]$. Therefore $p(x) = (x - \lambda)^2 q(x)$ for some $q \in \mathbb{R}[x]$. Conversely if $p(x) = (x - \lambda)^2 q(x)$, then

$$\varphi(p) = \varphi((x - \lambda)) \varphi((x - \lambda)) \varphi(q) = (0, 1)(0, 1) \varphi(q) = (0, 0) \varphi(q) = 0.$$

Thus we have shown, $\ker(\varphi) = \mathbb{R}[x](x - \lambda)^2$ as claimed when $n = 1$. (We have also shown that $(0, 1)$ is a zero divisor in R and hence R is **not** an integral domain.) ■

12.2 More Example of Factor Rings

Example 12.7. Here is another example similar to Example 11.1. In $R := \mathbb{Z}[i] / \langle 3 + i \rangle$, we have $[i] = [-3]$ and therefore $[-1] = [9]$ or equivalently $[10] = 0$. Therefore for $a, b \in \mathbb{Z}$,

$$[a + ib] = [a - 3b] = [(a - 3b) \bmod 10].$$

Thus we should take $\mathcal{S} = \{0, 1, 2, \dots, 9\}$. If $a, b \in \mathcal{S}$ and $[a] = [b]$, then $[c] = 0$ where $c = (b - a) \bmod 10$. Since $[c] = 0$, we must have

$$c = (3 + i)(a + ib) = (3a - b) + (a + 3b) i$$

from which it follows that $a = -3b$ and $3(-3b) - b = -10b = c$. Since $0 \le c \le 9$, this is only possible if $c = 0$ and so as above if $a = b$. Therefore

$$\mathcal{S} \ni a \to [a] = a + \langle 3+i \rangle \in \mathbb{Z}[i] \,/\, \langle 3+i \rangle$$

is a bijection. Moreover it is easy to see that thinking of $\mathcal{S}$ as $\mathbb{Z}_{10}$, the above map is in fact a ring isomorphism.

Example 12.8 (Example 12.7 revisited). From Example 8.9 we have seen that $\varphi : \mathbb{Z}[i] \to \mathbb{Z}_{10}$ defined by $\varphi(a+bi) = (a-3b) \bmod 10$ is a ring homomorphism – recall that $3^2 = (-3)^2 = 9 = -1$. In this case,

$$a + ib \in \ker(\varphi) \iff a - 3b = 0 \text{ in } \mathbb{Z}_{10},$$

i.e. $a = 3b + 10k$ for some $k \in \mathbb{Z}$. Therefore,

$$\ker(\varphi) = \{3b + 10k + ib : b, k \in \mathbb{Z}\} = \mathbb{Z}(3+i) + \mathbb{Z} \cdot 10.$$

In particular it follows that $3 + i \in \ker(\varphi)$ and therefore

$$\langle 3+i \rangle \subset \ker \varphi = \mathbb{Z}(3+i) + \mathbb{Z} \cdot 10.$$

Moreover, since $(3-i)(3+i) = 10$, we see that

$$\mathbb{Z}(3+i) + \mathbb{Z} \cdot 10 \subset \mathbb{Z}(3+i) + \mathbb{Z}[i](3+i) = \mathbb{Z}[i](3+i) = \langle 3+i \rangle.$$

Hence we have shown,

$$\langle 3+i \rangle \subset \ker \varphi = \mathbb{Z}(3+i) + \mathbb{Z} \cdot 10 \subset \langle 3+i \rangle$$

and therefore

$$\ker \varphi = \mathbb{Z}(3+i) + \mathbb{Z} \cdot 10 = \langle 3+i \rangle = \mathbb{Z}[i](3+i).$$

Consequently, by the first isomorphism theorem, $\bar{\varphi} : \mathbb{Z}[i]/\langle 3+i \rangle \to \mathbb{Z}_{10}$, given by

$$\bar{\varphi}([a+bi]) = \varphi(a+bi) = (a-3b) \bmod 10$$

is a ring isomorphism. Again, by taking $b = 0$, we see that $\bar{\varphi}^{-1}(a) = [a] = a + \langle 3+i \rangle$ is the inverse isomorphism, compare with Example 12.7.

Theorem 12.9. *Let $\rho \in \mathbb{Z}_+$ and $a, b \in \mathbb{Z}$ such that $a+ib \neq 0$ and $1 = \gcd(a,b)$. Further let*

$$\mathcal{S} := \mathbb{Z}_{\rho(a^2+b^2)} + i\mathbb{Z}_\rho = \left\{x + iy : x \in \mathbb{Z}_{\rho(a^2+b^2)} \text{ and } y \in \mathbb{Z}_\rho\right\}$$

where $\mathbb{Z}_1 := \{0\}$ and in this case we may take $\mathcal{S} := \mathbb{Z}_{\rho(a^2+b^2)}$. Then the map,

$$\mathcal{S} \ni (x+iy) \xrightarrow{\varphi} [x+iy] \in \mathbb{Z}[i] \,/\, \langle \rho(a+ib) \rangle \tag{12.6}$$

is a bijection of sets. If we further assume that $\rho = 1$, then

$$\mathbb{Z}_{(a^2+b^2)} \ni x \to [x] \in \mathbb{Z}[i] \,/\, \langle a+ib \rangle \tag{12.7}$$

is an isomorphism of rings.

Proof. The proof is carried out in a number of steps.

1. First observe that
$$\begin{aligned}\langle \rho(a+ib) \rangle &= \{\rho(a+ib)(s+it) : s,t \in \mathbb{Z}\} \\ &= \{\rho[as - bt + i(bs+at)] : s,t \in \mathbb{Z}\}. \end{aligned} \tag{12.8}$$
2. There exists $s, t \in \mathbb{Z}$ such that $bs + at = 1$ and so from Eq. (12.8) it follows that $[\rho i] = [bt - as]$. Therefore every element of $\mathbb{Z}[i] \,/\, \langle \rho(a+ib) \rangle$ may be represented in the form $[x+iy]$ where $x \in \mathbb{Z}$ and $y \in \mathbb{Z}_\rho$. Notice that
$$\rho = \min\{\beta \in \mathbb{Z}_+ : \alpha + i\beta \in \langle \rho(a+ib) \rangle \text{ for some } \alpha \in \mathbb{Z}\}.$$
3. If $s, t \in \mathbb{Z}$ such that $bs + at = 0$, then $(s,t) = \lambda(a,-b)$ for some $\lambda \in \mathbb{Q}$. In fact λ can not be a fraction. If it were, since both $s, t \in \mathbb{Z}$, the denominator (in reduced form) of λ would have to divide both a and b and hence $\lambda = \pm 1$ as $\gcd(a,b) = 1$. Thus we have $\lambda \in \mathbb{Z}$.
For such $(s,t) = \lambda(a,-b)$ with $\lambda \in \mathbb{Z}$ we have
$$\rho[as - bt + i(bs+at)] = \lambda\rho\left(a^2+b^2\right).$$
So the smallest positive number this expression can take is $\rho\left(a^2+b^2\right)$ which occurs when $\lambda = 1$.
4. From item 3. it follows that $\left[\rho\left(a^2+b^2\right)\right] = 0$. Combining this observation with item 2. shows that the map, φ, in Eq. (12.6) is onto.
5. The last main thing to prove is that the map φ is one to one. Suppose that $x + iy$ and $x' + iy'$ are in $\mathcal{S}$ with $[x+iy] = [x'+iy']$. This happens iff
$$[x - x' + i(y-y')] = [0] \iff x - x' + i(y-y') \in \langle \rho(a+ib) \rangle. \tag{12.9}$$
Since $|y - y'| < \rho$, it follows form item 2. that if Eq. (12.9) holds then $y - y' = 0$. Since $|x - x'| < \rho\left(a^2+b^2\right)$, it now follows from item 3. that we must have $x - x' = 0$. Thus we have shown $x + iy = x' + iy'$ and hence φ is one to one.
6. The assertion that when $\rho = 1$ the map in Eq. (12.7) is a ring isomorphism is left to the reader.

■

Example 12.10. In this example, we wish to consider, $\mathbb{Z}[x] \,/\, \langle 2x-1 \rangle$. In this ring we have

$$[1] = [2x] = [2][x]$$

which suggests that roughly speaking, "$[x] = 1/2$." Thus we might guess that

$$\mathbb{Z}[x] \,/\, \langle 2x-1 \rangle \cong \mathbb{Z}[1/2]. \tag{12.10}$$

The general element of $\mathbb{Z}[1/2]$ is a rational number which has a denominator of the form 2^n for some $n \in \mathbb{N}$. In order to try to prove this, let $\varphi : \mathbb{Z}[x] \to \mathbb{Z}[1/2]$ be the evaluation map, $\varphi(p) = p(1/2)$. Since $\varphi(\mathbb{Z}[x]) = \mathbb{Z}[1/2]$ to prove Eq. (12.10) we need to show

$$\ker(\varphi) = \langle 2x-1 \rangle . \tag{12.11}$$

On one hand it is clear that $2x-1 \in \ker(\varphi)$ and therefore $\langle 2x-1\rangle \subset \ker(\varphi)$. For the opposite inclusion, suppose that $p \in \ker(\varphi)$, i.e. $p(1/2) = 0$. By the division algorithm, we may write $p(x) = q(x)(x - 1/2) + r$ where $r \in \mathbb{Q}$. Since $p(1/2) = 0$ it follows that $r = 0$. Let $g(x) := \frac{1}{2} q(x)$, then $g(x) \in \mathbb{Q}[x]$ satisfies,

$$p(x) = g(x)(2x-1).$$

I claim that $g(x) \in \mathbb{Z}[x]$. To see this look at the expressions,

$$p(x) = \sum_{k=0}^{n} a_k x^k = \left(\sum_{j=0}^{n-1} b_j x^j\right)(2x-1)$$

where $a_k \in \mathbb{Z}$ and $b_k \in \mathbb{Q}$. By looking at the coefficient of the x^k term we learn, $a_k = -b_k + 2b_{j-1}$ with the convention that $b_{-1} = 0 = b_n$. So for $k = 0$ we learn that $b_0 = -a_0 \in \mathbb{Z}$, and for general k, that $b_k = -a_k + 2b_{j-1}$. Thus it follows inductively that $b_k \in \mathbb{Z}$ for all k.

Hence we have shown if $p \in \ker(\varphi)$, then $p \in \langle 2x-1\rangle$, i.e. $\ker(\varphi) \subset \langle 2x-1\rangle$ which completes the proof of Eq. (12.11).

12.3 II. More on the characteristic of a ring

Let R be a ring with 1. Recall: the characteristic of R is the minimum $n > 1$ (if any exist) such that $n \cdot 1 = \overbrace{1 + \cdot + 1}^{n} = 0$. If no such n exists, we call $\operatorname{chr}(R) = 0$.

Theorem 12.11 (Characteristic Theorem). *Let R be a ring with 1. Then $\varphi(a) := a \cdot 1_R$ is a homomorphism from $\mathbb{Z} \to R$ and $\ker\varphi = \langle m \rangle$ where $m = \operatorname{chr}(R)$. Moreover, R contains a copy of $\mathbb{Z}/\langle m\rangle$ as a subring.*

Proof. Since $1_R^2 = 1_R$, we have already seen that $\varphi(a) = a \cdot 1_R$ defines a homomorphism Moreover it is clear that $a \cdot 1_R = 0$ iff $\operatorname{chr}(R)|a$, i.e. $\ker(\varphi) = \langle m\rangle$. The remaining statement follows by an application of the first isomorphism theorem; i.e. $\mathbb{Z}/\langle m\rangle \cong \varphi(Z) = \operatorname{Ran}\varphi$. So $\operatorname{Ran}\varphi$ is a subring of R, and it is isomorphic to $\mathbb{Z}/\langle m\rangle$. ■

So the rings $\mathbb{Z}$ and $\mathbb{Z}/\langle m\rangle \cong \mathbb{Z}_m$ are the "simplest" rings in the sense that every ring with 1 has a copy of one of these sitting inside of it.

Example 12.12. Let $m \geq 2$ and

$$R = M_2(\mathbb{Z}_m) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{Z}_m \right\}.$$

Then the homomorphisms above is $\varphi : \mathbb{Z} \to M_2(\mathbb{Z}_m)$ by

$$a \mapsto a \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$$

$\ker\varphi = \langle m\rangle$, and R has the subring

$$\left\{ \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} : a, b, c, d \in \mathbb{Z}_m \right\}$$

which is isomorphic is $\mathbb{Z}_m$.

Example 12.13. If $R = \mathbb{Z}_m$ the homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}_m$ constructed above is just the natural one $a \mapsto a \bmod m$ that we have been looking at all along and $\operatorname{chr}(\mathbb{Z}_m) = m$.

Example 12.14. If $R = \mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$. Then $\varphi : \mathbb{Z} \to \mathbb{Z}[i], a \mapsto a \cdot 1 = a + 0i$ has kernel $\ker\varphi = \langle 0\rangle$. So $\operatorname{chr}(\mathbb{Z}[i]) = 0$ and $\mathbb{Z}[i]$ has a copy of $\mathbb{Z}/\langle 0\rangle \cong \mathbb{Z}$ inside it, namely $\{a + 0i : a \in \mathbb{Z}\}$.

12.4 Summary

Let us summarize what we know about rings so far and compare this to the group theory of last quarter.

	Group	**Ring**
Definition	G with $\cdot$,associative, identity,multiplicative inverse.	R with $(+,\cdot)$ $\ni (R,+)$ is an abelian group (can add and subtract). Associative, distributive laws $a(b+c)=ab+ac$, $(b+c)a=ba+ca$.
Sub-structure	$H \subset G$ is a subgroup if $h_1h_2^{-1} \in H$ for all $h_1, h_2 \in H$, i.e. H is closed under the group operations	$S \subset R$ is a subring if $a-b \in S,\ ab \in S$ $\forall\ a, b \in S$.
Factor Structure	If $H \lhd G$ is a normal subgroup of G, then $G/H := \{gH : g \in G\}$ is the factor group of G by H.	If $I \subset R$ is an ideal, then $R/I = \{r + I : r \in R\}$ is the factor ring of R by I.
Homomorphisms: Functions Preserving Structure	$\varphi : G \to H$ a function between groups, G, and H is a homomorphism if $\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2)$ for all $g_1, g_2 \in G$.	$\varphi : R \to S$ is a function between two rings R and S is a homomorphism if $\varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2)$ and $\varphi(r_1r_2) = \varphi(r_1)\varphi(r_2)$ for all $r_1, r_2 \in R$.

13

Lecture 13

13.1 Ideals and homomorphisms

Example 13.1. Let $R = \mathbb{Z}$. We have already seen that the ideals of $\mathbb{Z}$ are exactly $\langle 0\rangle = \{0\}$, $\langle 1\rangle = \mathbb{Z}$, $\langle 2\rangle, \langle 3\rangle, \dots$, i.e., every ideal is a principle ideal of the form $\langle m\rangle$ for some $m \geq 0$. We already know that the ideal $\langle m\rangle$ is the kernel of the homomorphism $\mathbb{Z} \to \mathbb{Z}/\langle m\rangle \cong \mathbb{Z}_m$.

Theorem 13.2 (Kernels are ideals). *Let R be a ring and $I \subset R$ be an ideal. Then the that map, $\pi : R \to R/I$, defined by*

$$\pi(a) := [a] = a + I \text{ for all } a \in R,$$

is a homomorphism of rings called that ***natural homomorphism.*** *In particular as subset, $S \subset R$ is an ideal iff S is the kernel of some ring homomorphism.*

Proof. We have already seen in Lemma 7.6 that kernels of ring homomorphisms are ideals. We leave it to the reader to verify $\pi : R \to R/I$ is a homomorphism with $\ker \pi = I$. Once this is done, it follows that every ideal is also the kernel of some homomorphism – namely π. ■

Lemma 13.3. *Let $\varphi : R \to \bar{R}$ be a surjective homomorphism of rings and $J \subset R$ be an ideal. Then $\varphi^{-1}(\varphi(J)) = J + \ker(\varphi)$.*

Proof. Unwinding all the definitions implies;

$$\begin{aligned} a \in \varphi^{-1}(\varphi(J)) &\iff \varphi(a) \in \varphi(J) \iff \varphi(a) = \varphi(j) \text{ for some } j \in J \\ &\iff \varphi(a - j) = 0 \text{ for some } j \in J \\ &\iff a - j \in \ker(\varphi) \text{ for some } j \in J \\ &\iff a \in J + \ker(\varphi). \end{aligned}$$

■

Proposition 13.4. *Let $\varphi : R \to \bar{R}$ be a surjective homomorphism of rings and $I := \ker(\varphi)$. Then the two sided ideals of R which contain I are in one to one correspondence with the two sided ideals of $\bar{R}$. The correspondence is given by,*

$$\{J : I \subset J \subset R\} \ni J \to \varphi(J) \subset \bar{R} \text{ and} \tag{13.1}$$

$$\{J : I \subset J \subset R\} \ni \varphi^{-1}(\bar{J}) \longleftarrow \bar{J} \subset \bar{R}. \tag{13.2}$$

Proof. Let us begin by showing that $\varphi(J)$ and $\varphi^{-1}(\bar{J})$ are ideals whenever J and $\bar{J}$ are ideals. First off it is easy to verify that $\varphi(J)$ and $\varphi^{-1}(\bar{J})$ are sub-rings if J and $\bar{J}$ are subrings. Moreover, for $r \in R$, we have

$$\varphi\left(r\varphi^{-1}(\bar{J})\right) = \varphi(r)\,\varphi\left(\varphi^{-1}(\bar{J})\right) = \varphi(r)\,\bar{J} \subset \bar{J}$$

and similarly,

$$\varphi\left(\varphi^{-1}(\bar{J})\,r\right) = \varphi\left(\varphi^{-1}(\bar{J})\right)\varphi(r) = \bar{J}\varphi(r) \subset \bar{J}$$

wherein we have used φ is surjective to conclude that $\varphi\left(\varphi^{-1}(\bar{J})\right) = \bar{J}$. Similarly, if $\bar{r} \in \bar{R}$, there exists $r \in \varphi^{-1}(\{\bar{r}\})$ and therefore,

$$\begin{aligned} \bar{r}\varphi(J) &= \varphi(r)\,\varphi(J) = \varphi(rJ) \subset \varphi(J) \text{ and} \\ \varphi(J)\,\bar{r} &= \varphi(J)\,\varphi(r) = \varphi(Jr) \subset \varphi(J) \end{aligned}$$

which shows that $\varphi(J)$ is an ideal as well.

Lastly we show that the maps in Eqs. (13.1) and (13.2) are inverses to one another. Indeed,

$$\begin{aligned} \varphi\left(\varphi^{-1}(\bar{J})\right) &= \bar{J} \text{ and} \\ \varphi^{-1}(\varphi(J)) &= J + I = J. \end{aligned}$$

In the first line we have used that fact that φ is surjective while in the second we used Lemma 13.3 and the assumption that $I = \ker(\varphi) \subset J$ so that $I + J = J$. ■

Example 13.5. Let $\varphi : \mathbb{Z}[x] \to \mathbb{Z}$ be the evaluation homomorphism, $\varphi(f(x)) = f(0)$ so that

$$\ker(\varphi) = \langle x\rangle = \{f(x) \in \mathbb{Z}[x] : f(0) = 0\}.$$

For any $m \in \mathbb{Z}$, $\langle m\rangle \subset \mathbb{Z}$ is an ideal and therefore,

$$\varphi^{-1}(\langle m\rangle) = \{f(x) \in \mathbb{Z}[x] : f(0) \in \langle m\rangle\}$$

is an ideal in $\mathbb{Z}[x]$ which contains $\langle x\rangle$. In fact this is the list of all ideals of $\mathbb{Z}[x]$ which contain $\langle x\rangle = \varphi^{-1}(\{0\})$.

Corollary 13.6. *If $I \subset R$ is an ideal than the ideals of R/I are in one to one correspondence with the ideals of R containing I. The correspondence is given by $J \subset R$ with $I \subset J$ is sent to $\pi(J) = \{[a] = a + I : a \in J\}$.*

Example 13.7. Let us find all of the ideals of $\mathbb{Z}/\langle 10\rangle$. Since that ideals of $\mathbb{Z}$ containing $\langle 10\rangle$ are;

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}, \ \ \langle 2\rangle = 0, \pm 2, \pm 4, \dots\},$$
$$\langle 5\rangle = \{0, \pm 5, \pm 10, \dots\}, \text{ and } \langle 10\rangle = \{0, \pm 10, \pm 20, \dots\},$$

it follows by Corollary 13.6 that the ideals of $\mathbb{Z}/\langle 10\rangle$ are given by

$$\begin{aligned} \mathbb{Z}/\langle 10\rangle &= \{0,1,2,\dots,9\} + \langle 10\rangle, \\ \langle 2\rangle / \langle 10\rangle &= \{0,2,4,6,8\} + \langle 10\rangle, \\ \langle 5\rangle / \langle 10\rangle &= \{0,5\} + \langle 10\rangle, \text{ and} \\ \langle 10\rangle / \langle 10\rangle &= \{0\} + \langle 10\rangle. \end{aligned}$$

13.2 Maximal and Prime Ideals

Let R be a ring, and I an ideal. How do we tell if R/I has properties we like? For example, when is R/I a domain or when is R/I a field? It turns out that this has nothing to do with R itself but rather only depends on properties of the ideal, I. We explore these connections in the context where R is a commutative ring with $1 \in R$.

Definition 13.8. *Let R be a commutative ring with identity 1. An ideal I is called **prime** if given any $a, b \in R$ with $ab \in I$, either $a \in I$ or $b \in I$. An ideal I is called **maximal** if given an ideal J with $I \subset J \subset R$ either $I = J$ or $J = R$.*

Example 13.9. In $\mathbb{Z}$, the ideal $\langle m\rangle$ is neither prime nor maximal when m is composite. For example, $\langle 6\rangle$ is not prime. We see that $2 \cdot 3 = 6 \in \langle 6\rangle$, but $2 \notin \langle 6\rangle$, and $3 \notin \langle 6\rangle$. Moreover $\langle 6\rangle \subsetneq \langle 2\rangle \subsetneq \mathbb{Z}$ and $\langle 6\rangle \subsetneq \langle 3\rangle \subsetneq \mathbb{Z}$ which shows that $\langle 6\rangle$ is not prime. On the other hand, if p is prime, then the ideal $\langle p\rangle$ is both prime and maximal. Let us also observe that $\langle 0\rangle = \{0\}$ is prime, but not maximal. To see that it is not maximal simply observe, for example, that $\langle 0\rangle \subsetneq \langle 2\rangle \subsetneq \mathbb{Z}$. The fact that $\langle 0\rangle$ is prime is equivalent to the statement that $\mathbb{Z}$ is an integral domain, see the next lemma.

Example 13.10. Consider the ideal, $I := \langle x\rangle \subset \mathbb{Z}[x]$. This ideal is prime. Indeed, if $p, q \in \mathbb{Z}[x]$, then

$$\begin{aligned} p(x)\, q(x) \in I &\implies p(0)\, q(0) = 0 \\ &\implies p(0) = 0 \text{ or } q(0) = 0 \implies p(x) \in I \text{ or } q(x) \in I. \end{aligned}$$

On the other hand, because of Example 13.5 we know that $\langle x\rangle$ is not a maximal ideal.

Lecture 14

Lemma 14.1. *Let R be a commutative ring with identity $1 = 1_R$. Then R is an integral domain iff $\langle 0\rangle = \{0\}$ is a prime ideal and R is a field iff $\langle 0\rangle = \{0\}$ is a maximal ideal.*

Proof. Suppose that $a, b \in R$. Then $ab \in \langle 0\rangle$ iff $ab = 0$. If $\langle 0\rangle$ is a prime ideal then $a \in \langle 0\rangle$ or $b \in \langle 0\rangle$, i.e. either $a = 0$ or $b = 0$. This shows that R is an integral domain. Similarly if R is an integral domain and $ab \in \langle 0\rangle$, then $ab = 0$ and hence either $a = 0$ or $b = 0$, i.e. either $a \in \langle 0\rangle$ or $b \in \langle 0\rangle$. Thus $\langle 0\rangle$ is a prime ideal.

You are asked to prove on your homework that R is a field iff the only ideals of R are $\{0\}$ and R. Now $\{0\}$ is maximal iff the only other ideal of R is R itself. ■

Theorem 14.2. *Suppose that R and $\bar{R}$ are commutative rings with identities and $\varphi : R \to \bar{R}$ is a surjective homomorphism. Then an ideal, $J \subset R$ such that* $\ker(\varphi) \subset J$ *is prime (maximal) iff $\varphi(J)$ is prime (maximal) in $\bar{R}$. Similarly, an ideal $\bar{J} \subset \bar{R}$ is prime (maximal) iff $\varphi^{-1}\left(\bar{J}\right)$ is prime (maximal) in R.*

Proof. We begin by proving the statements referring to J. In what follows below J will always be an ideal of R containing $\ker(\varphi)$.

1. Suppose that $J \subset R$ is prime and let $\bar{a} = \varphi(a)$ and $\bar{b} = \varphi(b)$ are generic elements of $\bar{R}$. Then $\bar{a}\bar{b} \in \varphi(J)$ iff $\varphi(a)\varphi(b) = \varphi(j)$ for some $j \in J$ which happens iff $ab - j \in \ker(\varphi)$, i.e. $ab \in J + \ker(\varphi) = J$. Since J is prime it follows that either a or $b \in J$ and therefore either $\bar{a}$ or $\bar{b}$ in $\varphi(J)$.
   Now suppose that $\varphi(J)$ is prime. If $a, b \in R$ such that $ab \in J$, then $\varphi(a)\varphi(b) = \varphi(ab) \in \varphi(J)$. Since $\varphi(J)$ is prime it follows that either $\varphi(a)$ or $\varphi(b)$ is in $\varphi(J)$, i.e. either a or $b \in \varphi^{-1}(\varphi(J)) = J + \ker(\varphi) = J$.
2. If $J \subset R$ is not a maximal ideal then there exists an ideal K such that $J \subsetneq K \subsetneq R$. Since $\varphi^{-1}(\varphi(J)) = J$, $\varphi^{-1}(\varphi(K)) = K$, $\varphi^{-1}\left(\bar{R}\right) = R$, it follows that $\varphi(J) \subsetneq \varphi(K) \subsetneq \bar{R}$ which shows $\varphi(J)$ is not maximal. Conversely if $\varphi(J)$ is not maximal, there exists an ideal, $\bar{K}$ of $\bar{R}$, such that $\varphi(J) \subsetneq \bar{K} \subsetneq \bar{R}$. Then $K := \varphi^{-1}\left(\bar{K}\right)$ is an ideal of R that $J \subset K \subset R$. Since $\varphi(J) \subsetneq \bar{K} = \varphi(K) \subsetneq \bar{R}$, it follows that $J \subsetneq K \subsetneq R$ and hence J is not maximal.

The statements referring to $\bar{J}$ now follow from what we have already proved. Indeed, let $\bar{J} \subset \bar{R}$ be an ideal and $J := \varphi^{-1}\left(\bar{J}\right)$ which is an ideal of R containing $\ker(\varphi)$. By what we have already proved we know that $\varphi^{-1}\left(\bar{J}\right) = J$ is prime (maximal) in R iff $\bar{J} = \varphi\left(\varphi^{-1}\left(\bar{J}\right)\right) = \varphi(J)$ is prime (maximal) in $\bar{R}$. ■

The following theorem is now an easy corollary of Lemma 14.1 and Theorem 14.2.

Theorem 14.3. *Let R be commutative with 1, I a prime ideal. Then*

1. *R/I is an integral domain $\Leftrightarrow$ I is a prime.*
2. *R/I is a field $\Leftrightarrow$ I is maximal.*

Proof. Easy Proof. As usual we will write $[a]$ for $a + I$ and recall from your homework that R/I is a commutative ring with identity, $[1] = 1 + I$. Let $\pi : R \to R/I$ be the natural homomorphism, $\pi(a) = [a] = a + I$. Then $I = \ker(\pi) = \pi^{-1}(\{0\})$. Therefore, by Theorem 14.2, $I \subset R$ is prime (maximal) iff $\{0\} \subset R/I$ is prime (maximal). But by Lemma 14.1 we know that $\{0\} \subset R/I$ is prime iff R/I is an integral domain and $\{0\} \subset R/I$ is maximal iff R/I is a field.

Second Proof. To help the reader understand this theorem better, let us also give a second more direct proof of the theorem.

1. Suppose I is prime and $[ab] = [a][b] = [0]$ in R/I, then $ab \in I$. Since I is prime it follows that $a \in I$ or $b \in I$, i.e. $[a] = 0$ or $[b] = 0$. Therefore R/I has no zero divisors, i.e. R/I is an integral domain.
   Conversely, if I is not prime there exists $a, b \in R \setminus I$ with $ab \in I$. Therefore $[a] \neq 0 \neq [b]$ while $[a][b] = [ab] = 0$ which shows that R/I has zero divisors and hence is not an integral domain.
2. Suppose that I is maximal and let $0 \neq [a] \in R/I$ so that $a \in R$ but $a \notin I$. Let J be the ideal generated by a and I, i.e.
   $$J := Ra + I = \{ra + b : r \in R, b \in I\}.$$
   (You should check that J is an ideal.) Since $a \in J$ it follows that $I \subsetneq J$ and since I was maximal we may conclude that $J = R$. In particular $1 \in J$ and hence there exists $r \in R$ and $b \in I$ such that $1 = ra + b$. Therefore, $1 = [1] = [ra] = [r] \cdot [a]$ which shows $[a]^{-1}$ exists and is equal to $[r]$. Therefore $U(R/I) = (R/I) \setminus \{0\}$ which shows R/I is a field.
   Conversely if I is not maximal, then there exists another ideal, J, of R such that $I \subsetneq J \subsetneq R$. Let $b \in J \setminus I$ so that $[b] = b + I \neq 0$. If $[b]^{-1}$ exists, then

there exist $a \in R$ such that $[a][b] = 1 = [1]$, i.e. $i := ab - 1 \in I$. Solving for 1 gives,

$$1 = i + ab \in I + J \subset J.$$

This however contradicts the fact that J is proper since $R = R \cdot 1 \subset RJ = J$. Thus $[b]$ is not invertible for all $b \in J \setminus I$.

■

Corollary 14.4. *In a commutative ring R with* 1, *every maximal ideal is also a prime ideal.*

Proof. First Proof. This follows from the Theorem 14.3 since a field is always an integral domain.

Second Proof. Suppose I is a maximal ideal that $a \notin I$ and let $J := I + Ra = J + \langle a \rangle$. Then J is an ideal in R which properly contains I and therefore we must have $J = R$. Hence it follows that $1 = ra + i$ for some $r \in R$ and $i \in I$. So if $ab \in I$ then

$$b = rab + ib \in I$$

showing I is a prime ideal. ■

Example 14.5 (Example 13.9 revisited). Since $\mathbb{Z}/\langle m \rangle \cong \mathbb{Z}_m$ and $\mathbb{Z}_m$ is an integral domain (field) iff m is prime, we see that the following are equivalent,

1. m is prime,
2. $\langle m \rangle$ is a prime ideal of $\mathbb{Z}$, and
3. $\langle m \rangle$ is a maximal ideal of $\mathbb{Z}$.

Example 14.6. In $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$, $I = \{0, 3\}$ is a maximal ideal. To see this let $\varphi : \mathbb{Z}_6 \to \mathbb{Z}_3$ be the homomorphism, $\varphi(x) = x \bmod 3$. Then $\ker(\varphi) = I$ and since $\mathbb{Z}_3$ is a field it follows that $\langle 0 \rangle$ is a maximal ideal and hence $\varphi^{-1}(\{0\}) = \ker(\varphi)$ is a maximal ideal.

Example 14.7. In $\mathbb{R}[x]$, $\langle x \rangle$ is maximal since $\mathbb{R}[x] / \langle x \rangle \cong \mathbb{R}$. Notice also that $\langle 0 \rangle$ is prime since $\mathbb{R}[x]$ is an integral domain but not maximal since $\mathbb{R}[x]$ is not a field.

Example 14.8. $\langle 2 - i \rangle$ is maximal inside of $\mathbb{Z}[i]$. This is hard to see without our earlier result that $\mathbb{Z}[i]/\langle 2 - i \rangle \simeq \mathbb{Z}_5$, which is a field.

Example 14.9. Let $I := \langle 2, x \rangle = \mathbb{Z}[x] \cdot 2 + \mathbb{Z}[x] \cdot x$ – an ideal of $\mathbb{Z}[x]$. Let $\varphi : \mathbb{Z}[x] \to \mathbb{Z}$ be the evaluation homomorphism, $\varphi(p(x)) = p(0)$ with $\ker\varphi = \langle x \rangle$. Notice that $\ker\varphi \subset I$ and that $\varphi(I) = 2\mathbb{Z} = \langle 2 \rangle$. Since $\langle 2 \rangle$ is maximal in $\mathbb{Z}$ we know that $I = \langle 2, x \rangle$ is a maximal ideal of $\mathbb{Z}[x]$. In fact for any prime, $p \in \mathbb{N}$, the same argument shows that $\langle p, x \rangle$ is a maximal ideal of $\mathbb{Z}[x]$ and these are precisely all of the maximal ideals of $\mathbb{Z}[x]$ which contain $\langle x \rangle$.

Example 14.10. $\langle x \rangle \subset \mathbb{Z}[x]$ is a prime ideal which is not maximal. Indeed from the previous example, $\langle x \rangle \subsetneq \langle 2, x \rangle \subsetneq \mathbb{Z}[x]$ which implies $\langle x \rangle$ is not maximal. To see that it is prime observe that $\mathbb{Z}[x] / \langle x \rangle \cong \mathbb{Z}$ is an integral domain. Moreover since $\mathbb{Z}$ is not a field it again follows that $\langle x \rangle$ is not a maximal ideal.

Alternative 1. Observe that $\varphi(\langle x \rangle) = \{0\}$ and $\{0\}$ is prime but not maximal in $\mathbb{Z}$.

Alternative 2. If $f \in \mathbb{Z}[x]$, we have $f(x) = xq(x) + a_0$ and using $[x] = 0 = [2]$ we find

$$[f(x)] = [x][q(x)] + [a_0] = [a_0 \bmod 2].$$

Moreover if $a, b \in \mathbb{Z}_2$ and $[a] = [b]$, then $a - b \in \langle 2, x \rangle$ which is only possible if $a - b = 0$. Thus it follows that we may take $\mathcal{S} = \mathbb{Z}_2$. We may now work as we have done many times before to see that $\mathbb{Z}[x] / \langle 2, x \rangle \cong \mathbb{Z}_2$.

Definition 14.11 (Principle ideal domains). *A* ***principle ideal domain*** *(PID for short) is an integral domain, R, such that every ideal $I \subset R$ is a principle ideal.*

Example 14.12. $\mathbb{Z}$ and $\mathbb{Z}_m$ for all $m \in \mathbb{Z}_+$ are principle ideal domains. We will also see later that $F[x]$ is a principle ideal domain for every field F.

Example 14.13. In this example we show that $\mathbb{Z}[x]$ is **not** a principle ideal domain. For example, consider that ideal generated by $\langle 2, x \rangle$, i.e.

$$\begin{aligned} \langle 2, x \rangle &= \mathbb{Z}[x] \cdot 2 + \mathbb{Z}[x] \cdot x \\ &= \{2a + xq(x) : a \in \mathbb{Z} \text{ and } q(x) \in \mathbb{Z}[x]\} \qquad (14.1) \\ &= \{f(x) \in \mathbb{Z}[x] : f(0) \in \langle 2 \rangle\}. \qquad (14.2) \end{aligned}$$

If $\langle 2, x \rangle = \langle p(x) \rangle$ for some $p \in \mathbb{Z}[x]$, then $2 = q(x)p(x)$ for some $q(x) \in \mathbb{Z}[x]$. However, this is only possible if both $q(x)$ and $p(x)$ are constant polynomials in which case we must have $p(x) = a_0 \in \{\pm 1, \pm 2\}$. We can rule out $a_0 = \pm 1$ since $\langle 2, x \rangle$ is a proper ideal and hence we may assume that $p(x) = 2$. Noting that $x \notin \langle 2 \rangle$ we learn that $\langle 2 \rangle \subsetneqq \langle 2, x \rangle$ and therefore $\langle 2, x \rangle$ is not a principle ideal in $\mathbb{Z}[x]$.

15

Lecture 15

- We first went over Quiz #4 in class.

Lemma 15.1. *Suppose that R is an integral domain and $a, b \in R$ with $a \neq 0 \neq b$. Then $\langle a \rangle = \langle b \rangle$ iff a and b are* ***associates****, i.e. $a = ub$ for some $u \in U(R)$.*

Proof. If $\langle a \rangle = \langle b \rangle$ then $a \in \langle b \rangle$ and $b \in \langle a \rangle$ and therefore there exists $u_1, u_2 \in R$ such that $a = u_1 b$ and $b = u_2 a$. Thus we may conclude that $b = u_2 u_1 b$ and hence by cancellation that $u_2 u_1 = 1$. This shows that $a = ub$ with $u = u_1 \in U(R)$. Conversely if $a = ub$ with $u \in U(R)$ then $a \in \langle b \rangle$ and $b \in \langle a \rangle$ since $b = u^{-1}a$. Therefore, $\langle a \rangle \subset \langle b \rangle$ and $\langle b \rangle \subset \langle a \rangle$, i.e. $\langle a \rangle = \langle b \rangle$. ■

Proposition 15.2 (maximal $\iff$ prime in PIDs). *If R is a principle ideal domain and $I \subset R$ be a non-zero ideal. Then I is maximal iff I is prime.*

Proof. By Corollary 14.4, we know in general that maximal ideals are prime ideals. So we need only show that if I is a prime ideal then I is a maximal ideal. Suppose that $I = \langle a \rangle \subset R$ (with $a \neq 0$) is a prime ideal and $J = \langle b \rangle$ is another ideal such that $I \subset J$. We will finish the proof by showing either $J = I$ or $J = R$.

As $a \in \langle b \rangle$, we have $a = bc$ for some $c \in R$. Since $bc \in I = \langle a \rangle$ and I is prime, we must have either $b \in \langle a \rangle$ or $c \in \langle a \rangle$. Case 1: if $b \in \langle a \rangle$, then $J = \langle b \rangle \subset \langle a \rangle = I$ and hence $J = I$. Case 2: if $c \in \langle a \rangle$, then $c = au$ for some $u \in R$ and we then have $a = bc = bau$. Cancelling a (here is where we use $a \neq 0$) from this equation shows that $1 = bu$ and therefore b and 1 are associates and so, by Lemma[1] 15.1, $J = \langle 1 \rangle = R$. ■

15.1 The rest this section was not covered in class

Lemma 15.3. *Suppose that R_1 and R_2 are two commutative rings with identities.. Then every ideal, $J \subset R_1 \oplus R_2$, is of the form $J = I_1 \oplus I_2$ where I_1 and I_2 are ideals of R_1 and R_2 respectively.*

Proof. It is easy to check that $J := I_1 \oplus I_2 \subset R_1 \oplus R_2$ is an ideal whenever I_1 and I_2 are ideals of R_1 and R_2 respectively. So let us concentrate on the converse assertion.

Suppose that $J \subset R_1 \oplus R_2$. If $(a, b) \in J$, then $(a, 0) = (1, 0)(a, b)$ and $(0, b) = (0, 1)(a, b)$ are in J. It is now a simple matter to check that

$$I_1 := \{a \in R_1 : (a, 0) \in J\} \text{ and } I_1 := \{b \in R_2 : (0, b) \in J\}$$

are ideals of R_1 and R_2 respectively. If $a \in I_1$ and $b \in I_2$, then $(a, b) = (a, 0) + (0, b) \in J$ showing $I_1 \oplus I_2 \subset J$. Similarly if $(a, b) \in J$, then as noted above $a \in I_1$ and $b \in I_2$ which implies $J \subset I_1 \oplus I_2$. ■

Corollary 15.4. *Let R_1, R_2, be as in Lemma 15.3. Then the maximal ideals of $R_1 \oplus R_2$ are of the form $J = I_1 \oplus R_2$ or $J = R_1 \oplus I_2$ where I_1 is a maximal ideal of R_1 and I_2 is a maximal ideal of R_2.*

Example 15.5 (Book problem 14.30). Find the maximal ideals, I, in $R := \mathbb{Z}_8 \oplus \mathbb{Z}_{30}$ and for each maximal ideal find size of the field, R/I. The only maximal ideal of $\mathbb{Z}_8$ is $\langle 2 \rangle$ and the maximal ideals of $\mathbb{Z}_{30}$ are $\langle 2 \rangle$, $\langle 3 \rangle$, and $\langle 5 \rangle$. Thus the maximal ideals of R are

$$\langle 2 \rangle \oplus \mathbb{Z}_{30}, \quad \mathbb{Z}_8 \oplus \langle 2 \rangle, \quad \mathbb{Z}_8 \oplus \langle 3 \rangle, \text{ and } \mathbb{Z}_8 \oplus \langle 5 \rangle.$$

The respective fields have size, by Lagrange's theorem or other means, $2 = 8/\frac{8}{\gcd(2,8)}$, $2 = 30/\frac{30}{\gcd(2,30)}$, $3 = 30/\frac{30}{\gcd(3,30)}$, and $5 = 30/\gcd(5, 30)$.

[1] More directly, $1 = bu$ implies b^{-1} exists and therefore $1 = b^{-1}b \in \langle b \rangle = J$ and hence that $J = R$.

16

Lecture 16

16.1 The Degree of a Polynomial

Recall the following definition of being a polynomial.

Definition 16.1. *Let R be any commutative ring with identity. The polynomial ring $R[x]$ is defined to be*

$$R[x] = \{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 : a_i \in R \text{ and } n \geq 0\}$$

equipped with the usual rules for addition and multiplication of polynomials.

Recall that the multiplicative identity in $R[x]$ is 1, and the additive identity is 0.

Notation 16.2 *Suppose that*

$$f = a_n x^n + \cdots + a_1 x + a_0 \in R[x]$$

is a polynomial with $a_n \neq 0$. The we say a_n is the ***leading coefficient*** *and f has* ***degree*** *n, written $\deg(f) = n$. (It is convenient to use the convention that $\deg(0) = -\infty$.) If $a_n = 1$, then f is called* ***monic.*** *When $\deg f = 0$, i.e. $f = a_0$, we say f is a* ***constant*** *polynomial.*

Example 16.3. In $\mathbb{C}[x]$, $\deg(4x^2 + (i+3)x + 5) = 2$. The coefficient of the largest power of x is called the leading coefficient. The leading coefficient of $4x^2 + (i+3)x + 5$ is 4. The polynomial $4x^2 + (i+3)x + 5$ is not monic while $g = x^2 + 6$ is monic. $f = 5$ is constant, but $g = x + 5$ is not.

Example 16.4. Let $R = \mathbb{Z}_6 = \{0, 1, \ldots, 5\}$,

$$f = x^2 + 5, \quad g = 2x + 1, \quad \text{and } h = 3x.$$

Then

$$\begin{aligned} fg &= (2x^3 + x^2 + 10x + 5) = 2x^3 + x^2 + 4x + 5, \\ f + g &= x^2 + 2x + 6 = x^2 + 2x, \text{ and} \\ gh &= 6x^2 + 3x = 3x. \end{aligned}$$

In this example, $\deg f = 2$, $\deg g = 1$, $\deg h = 1$, and $\deg(gh) = 1\,(\neq 2)$. So it is not always true in a polynomial ring that $\deg(fg) = \deg g + \deg h$. Let us compute the "values" of $g(x)$;

x	0	1	2	3	4	5
$g(x)$	0	3	0	3	0	3

Thus we see that $\deg(g) = 1$ yet $g(x)$ has three roots over $\mathbb{Z}_6$.

Theorem 16.5. *Let R be an integral domain. Then for any $f, g \in R[x]$,*

$$\deg(fg) = \deg(f) + \deg(g),$$

and $R[x]$ is an integral domain.

Proof. If $f = 0$ then $fg = 0$ and we will have $\deg(fg) = -\infty$ and $\deg(f) + \deg(g) = -\infty + \deg(g) = -\infty$. So we may now assume that f and g are non-zero polynomials which we write as,

$$\begin{aligned} f &= a_n x^n + \cdots + a_1 x + a_0, \text{ and} \\ g &= b_m x^n + \cdots + b_1 x + b_0 \end{aligned}$$

with $a_i, b_i \in R$, $a_n \neq 0$, and $b_m \neq 0$. Then $\deg f = n$ and $\deg g = m$. So

$$fg = a_n b_m x^{n+m} + \cdots + (a_1 b_0 + a_0 b_1)x + a_0 b_0,$$

and $a_n b_m \neq 0$ so that fg is not the zero polynomial and $\deg(fg) = m + n = \deg f + \deg g$. ■

Lemma 16.6. *If R is an integral domain, then $U(R[x]) = U(R)$. In particular if $R = F$ is a field and $F^\times := F \setminus \{0\}$, then $U(F[x]) = F^\times$.*

Proof. If $p(x) \in U(R[x])$ then there exists $q(x) \in R[x]$ such that $p(x)\,q(x) = 1$. Therefore, $0 = \deg(p) + \deg(q)$, showing $\deg(p) = 0 = \deg(q)$. Thus $p(x) = p_0 \in R$ and p_0 is invertible in $R[x]$ iff it is invertible in R because we have seen above that $\deg(q) = 0$ where $q(x)$ is the inverse to $p(x) = p_0$. ■

16.2 The evaluation homomorphism (review)

Theorem 16.7 (Evaluation homomorphism). *Let R be a subring of a commutative ring, $\bar{R}$, and $t \in \bar{R}$. Then there exists a ring homomorphism, $\varphi_t : R[x] \to \bar{R}$ such that*

$$\varphi_t(p) = \sum_{k=0}^{n} a_k t^k \text{ when } p(x) = \sum_{k=0}^{n} a_k x^k \in R[x].$$

We will usually simply write $p(t)$ for $\varphi_t(p)$.

Proof. Let $q(x) = \sum_{l=0}^{n} b_l x^l$, then

$$\begin{aligned}\varphi_t(p+q) &= \varphi_t\left(\sum_{l=0}^{n}(a_l + b_l)x^l\right) = \sum_{l=0}^{n}(a_l+b_l)t^l \\ &= \sum_{l=0}^{n}\left(a_l t^l + b_l t^l\right) = \sum_{l=0}^{n} a_l t^l + \sum_{l=0}^{n} b_l t^l \\ &= \varphi_t(p(x)) + \varphi_t(q(x)).\end{aligned}$$

Similarly,

$$\begin{aligned}\varphi_t(pq) &= \varphi_t\left(\sum_m \left(\sum_{l+k=m} a_l b_k\right) x^m\right) \\ &= \sum_m \left(\sum_{l+k=m} a_l b_k\right) t^m = \sum_m \left(\sum_{l+k=m} a_l b_k t^m\right) \\ &= \sum_m \left(\sum_{l+k=m} a_l t^l b_k t^k\right) = \sum_{l,k} a_l t^l b_k t^k \\ &= \left(\sum_l a_l t^l\right)\left(\sum_k b_k t^k\right) = \varphi_t(p) \cdot \varphi_t(q).\end{aligned}$$

The point is that the multiplication and addition rules for polynomials was chosen precisely so as to make this theorem true. ■

Example 16.8. Suppose that $\varphi := \varphi_1 : \mathbb{R}[x] \to \mathbb{R}$ is the evaluation homomorphism, $\varphi(p) = p(1)$. Then

$$\varphi(fg) = fg(1) = f(1)g(1) \text{ and}$$
$$\varphi(f+g) = [f+g](1) = f(1) + g(1).$$

For example suppose that

$$f = x^2 + 5 \text{ and } g = 2x^3 - 5x + 2.$$

Then

$$\begin{aligned}f + g &= 2x^3 + x^2 - 5x + 7, \\ fg &= 2x^5 + 5x^3 + 2x^2 - 25x + 10, \\ f(1) &= 6, \quad g(1) = -1,\end{aligned}$$

and so

$$(f+g)(1) = 5 = f(1) + g(1) \text{ and}$$
$$(fg)(1) = -6 = f(1) \cdot g(1).$$

Example 16.9 (Evaluation example). Suppose that $R = \mathbb{Z}_6 = \{0,1,2,3,4,5\}$, $a = 3$, and $\varphi : R[x] \to R$ is the evaluation map, $f \mapsto f(3)$. For example, if $f = 3x^2 + 5x + 2$, and $g = x + 3$, then

$$\varphi(f) = f(3) = 3(3)^2 + 5(3) + 2 = 44 = 2 \text{ and}$$
$$\varphi(g) = g(3) = 3 + 3 = 6 = 0,$$

from which it follows that $f \notin \ker(\varphi)$ while $g \in \ker(\varphi)$.

Example 16.10. Suppose that $\lambda \in \mathbb{R}$ and $\varphi = \mathrm{eval}_\lambda : \mathbb{R}[x] \to \mathbb{R}$, i.e. $\varphi(p) = p(\lambda)$. Then $p \in \ker\varphi$ iff $p(\lambda) = 0$ which happens (as we will see shortly) iff $p(x) = (x - \lambda)q(x)$ for some $q \in \mathbb{R}[x]$. Therefore,

$$\ker(\varphi) = \langle x - \lambda \rangle = \mathbb{R}[x](x - \lambda)$$

for this homomorphism.

16.3 The Division Algorithm

Definition 16.11. *Let R be an integral domain and $f, g \in R[x]$. We say that g **divides** f if $f = kg$ for some $k \in R[x]$. We also say that g is a **factor** of f.*

Example 16.12. In $\mathbb{Z}[x]$, $(2x-4)$ does not divide $\left(x^2 - 4\right)$. Indeed, if it did then

$$x^2 - 4 = (a + bx)(2x - 4) = -4a + (2a - 4b)x + 2bx^2$$

which would imply $2b = 1$ which is impossible in $\mathbb{Z}$. On the other hand, working in $\mathbb{Q}[x]$, we have

$$x^2 - 4 = (x-2)(x+2) = (2x-4)\left(\frac{1}{2}x + 1\right)$$

which shows that $(2x - 4)$ is a factor of $x^2 - 4$ in $\mathbb{Q}[x]$.

Theorem 16.13 (Division Algorithm). *Let $F[x]$ be a polynomial ring where F is a field. Given $f, g \in F[x]$ both nonzero, there exists a unique $q, r \in F[x]$ with $f = qg + r$ such that either $r = 0$ or $\deg r < \deg g$. (We will give the proof of this theorem later.)*

Interpretation. We are dividing f by g and so g **goes into** f, q **times with remainder** r. This is really high school polynomial division which we will discuss in more detail a bit later. In the sequel we will sometimes denote the remainder, r by $f \bmod g$.

Example 16.14. Let $f := 3x^3 + 5$ and $g = 2x + 3$ in $\mathbb{Q}[x]$, then

$$\begin{array}{r l}
 & \frac{3}{2}x^2 \;\; - \frac{9}{4}x + \frac{27}{8} \\ \cline{2-2}
2x+3 \big) & 3x^3 \qquad\qquad\quad + 5 \\
 & -3x^3 - \frac{9}{2}x^2 \\ \cline{2-2}
 & \qquad - \frac{9}{2}x^2 \\
 & \qquad \frac{9}{2}x^2 + \frac{27}{4}x \\ \cline{2-2}
 & \qquad\qquad \frac{27}{4}x \;\; + 5 \\
 & \qquad\qquad - \frac{27}{4}x - \frac{81}{8} \\ \cline{2-2}
 & \qquad\qquad\qquad - \frac{41}{8}
\end{array}$$

which shows,

$$3x^3 + 5 = \left(\frac{3}{2}x^2 - \frac{9}{4}x + \frac{27}{8}\right)(2x+3) + \left(-\frac{41}{8}\right)$$

so that

$$q(x) = \left(\frac{3}{2}x^2 - \frac{9}{4}x + \frac{27}{8}\right) \text{ and } r(x) = -\frac{41}{8}$$

in this example.

Example 16.15. Consider $f(x) = x^2 + x + 2$ and $g(x) = 2x + 1$ inside of $\mathbb{Z}_3[x]$. Then, using $2 \cdot 2 = 4 \bmod 3 = 1$, we find

$$\begin{array}{r l}
 & 2x - \;\; 2 \\ \cline{2-2}
2x+1 \big) & x^2 + \;\; x \;\; + 2 \\
 & x^2 + \; 2x \\ \cline{2-2}
 & \qquad -x + 2 \\
 & \qquad -x - 2 \\ \cline{2-2}
 & \qquad\qquad 4 = 1
\end{array},$$

which implies

$$\begin{aligned} x^2 + x + 2 &= (2x - 2)(2x + 1) + 1 \\ &= (2x + 1)(2x + 1) + 1. \end{aligned}$$

Example 16.16 (Example 16.15). Here is alternate way to do the last example. First use the division algorithm over $\mathbb{Q}$ to find;

$$\begin{array}{r l}
 & \qquad\quad \frac{1}{2}x + \frac{1}{4}, \\ \cline{2-2}
2x+1 \big) & x^2 \;\; + x + 2 \\
 & -x^2 - \frac{1}{2}x \\ \cline{2-2}
 & \qquad \frac{1}{2}x \;\; + 2 \\
 & \quad - \frac{1}{2}x - \frac{1}{4} \\ \cline{2-2}
 & \qquad\qquad \frac{7}{4}
\end{array}$$

that is over $\mathbb{Q}$ we have

$$x^2 + x + 2 = \left(\frac{1}{2}x + \frac{1}{4}\right)(2x + 1) + 7/4.$$

Multiplying this equation through by 4 gives,

$$4\left(x^2 + x + 2\right) = (2x + 1)(2x + 1) + 7$$

and then apply the "mod 3 homomorphism" to the coefficients implies,

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

which is the result above again.

Example 16.17. Let $f(x) = 2x^3$ and $g(x) = ix^2 + 5x + 2$ in $\mathbb{C}[x]$, then

$$\begin{array}{r l}
 & -2ix + \;\; 10 \\ \cline{2-2}
ix^2 + 5x + 2 \big) & 2x^3 \; + \; 0x^2 \; + \qquad 0x \qquad + \; 0 \\
 & 2x^3 \; - 10ix^2 - \qquad 4ix \\ \cline{2-2}
 & \qquad + 10ix^2 + \qquad 4ix \qquad + \; 0 \\
 & \qquad + 10ix^2 + \qquad 50x \qquad + 20 \\ \cline{2-2}
 & \qquad\qquad\qquad (-50 + 4i)\,x - 20
\end{array}$$

so that

$$2x^3 = (-2ix + 10)\left(ix^2 + 5x + 2\right) + (-50 + 4i)\,x - 20,$$

that is

$$q(x) = (-2ix + 10) \text{ and } r(x) = (-50 + 4i)\,x - 20.$$

Corollary 16.18. *Let F be a field, $a \in F$, and $f(x) \in F[x]$. Then $f(a)$ is the remainder in the division of $f(x)$ by $(x - a)$.*

Proof. By the division algorithm, there exists $k(x), r(x) \in F[x]$ such that $\deg(r) = 0 < 1$ and

$$f(x) = k(x)(x-a) + r(x). \tag{16.1}$$

Since $\deg(r) = 0$, $r(x) = b$ for some $b \in F$ and hence evaluating Eq. (16.1) at $x = a$ implies,

$$f(a) = k(a)(a-a) + b = b.$$

■

Example 16.19. Let $f(x) = x^2 + 5$ in $\mathbb{R}[x]$. If we divide $(x+1) = (x - (-1))$ into $f(x)$ the remainder will be $f(-1) = 6$.

Example 16.20. If we divide $x - 1$ into $x^2 + 2$ we find,

$$\begin{array}{r} x+1 \\ x-1\overline{\big)\; x^2 \qquad +2} \\ \underline{-\,x^2+x\quad} \\ x+2 \\ \underline{-\,x+1} \\ 3 \end{array}$$

which gives $x^2 + 2 = (x+1)(x-1) + 3$. Notice that the remainder, $3 = (1)^2 + 2$ as it should be.

Theorem 16.21 ($F[x]$ is a PID). *Let F be a field, then $F[x]$ is a principle ideal domain. Moreover the map,*

$$\{\textit{monic polynomials}\} \ni p \to \langle p \rangle \in \{\textit{non-zero ideals of } F[x]\} \tag{16.2}$$

is a one to one correspondence. The inverse map is given by associating to a non-zero ideal, $I \subset F[x]$, the unique monic polynomial, $p \in I$, with lowest degree.

Proof. Let us first show the map in Eq. (16.2) is one to one. So suppose that p and q are monic polynomials such that $\langle p \rangle = \langle q \rangle$. Then by Lemmas 15.1 and 16.6, we know that $p(x) = kq(x)$ for some $k \in U(F) = F \setminus \{0\}$. Since both p and q are monic, we must in fact have $k = 1$, i.e. $p(x) = q(x)$.

Suppose that $I \subset F[x]$ is an ideal. If $I = \{0\}$ then $I = \langle 0 \rangle$ so that $\{0\}$ is a principle ideal (as always). So now suppose that $I \neq \{0\}$ and let $p(x) \in I$ be a non-zero polynomial in I with minimal degree. By dividing $p(x)$ by its leading order coefficient, we may further assume that $p(x)$ is monic. If $f(x) \in I$, use the division algorithm to write, $f(x) = k(x)p(x) + r(x)$ where $\deg(r) < \deg(p)$. Since $r = f - kq \in I$, we must have $r = 0$ showing that $I = \langle p \rangle$ as claimed. If $q \in I$ is another monic polynomial such that $\deg(q) = \deg(p)$, then $q(x) = k(x)p(x)$ for some $k \in F[x]$. A simple degree argument then shows that $\deg(k) = 0$ so that $k(x) = k_0$ is a constant polynomial. Since q and p are both monic, it follows that $k_0 = 1$, i.e. $q(x) = p(x)$. ■

Example 16.22. $\mathbb{Z}[x]$ is not an principle ideal domain. For example consider the ideal,

$$I := \langle 2, x \rangle = \mathbb{Z}[x] \cdot 2 + \mathbb{Z}[x] \cdot x.$$

This ideal is proper since if $1 \in I$, then $1 = 2p(x) + xq(x)$ which would imply that $2p_0 = 1$ for some $p_0 \in \mathbb{Z}$. But this is impossible. If there exists $q \in \mathbb{Z}[x]$ such that $I = \langle q \rangle$, then $2 = q(x)p(x)$ for some p. However this would imply $0 = \deg(2) = \deg(p) + \deg(q)$ from which it follows that $\deg(q) = 0$. Therefore $q(x) = q_0$ for some $q_0 \in \mathbb{Z}$. As I is proper we know that $q_0 \neq \pm 1$ and since $2 \in \langle q_0 \rangle$ we must have $q_0 = \pm 2$. However, it should be clear that $x \in I$ while $x \notin \langle 2 \rangle = \langle -2 \rangle$. Thus I is not a principle ideal.

16.4 Appendix: Proof of the division algorithm

Let us now give the formal proof of Theorem 16.13.

Proof. Proof of Theorem 16.13. Suppose that $f, g \in F[x]$ with $g \neq 0$. We break the proof into the existence and uniqueness assertions.

Uniqueness. Suppose that we have two decompositions,

$$f = qg + r = q'g + r'$$

where $\deg r < \deg g$ and $\deg r' < \deg g$. Then

$$(q - q')g + (r - r') = 0,$$

or equivalently,

$$(q - q')g = (r' - r).$$

If $r - r' \neq 0$, we may take degrees of this equation to conclude,

$$\deg(q - q') + \deg g = \deg(r - r') < \deg g$$

which is a contradiction. Therefore $r = r'$ which then forces $q = q'$ since $F[x]$ is an integral domain and $g \neq 0$.This proves uniqueness.

Existence. Write

$$\begin{aligned} f(x) &= a_n x^n + \dots + a_1 x + a_0 \text{ and} \\ g(x) &= b_m x^m + \dots + b_1 x + b_0. \end{aligned}$$

If $m = \deg g > \deg f = n$ then we $q = 0$, and $r = f$ so that $f = 0 \cdot g + f$ with $\deg f < \deg g$.

If $m = \deg g \leq \deg f = n$, we will use induction on the degree of f. In the base case, $\deg f = 0$, we have $\deg g = 0$ so that $f = a_0$ and $g = b_0$ and therefore, $f = \frac{a_0}{b_0} g + 0$, so $r = 0$ in this case. Now suppose that $n = \deg f \geq 1$ and the existence has been established for all lower n. Let

$$\begin{aligned} f'(x) &= f(x) - \frac{a_n}{b_m} x^{n-m} g(x) \\ &= a_n x^n + \cdots + a_1 x + a_0 - \frac{a_n}{b_m} x^{n-m} \left(b_m x^m + \cdots + b_1 x + b_0 \right) \\ &= c_{n-1} x^{n-1} + \cdots + c_0. \end{aligned}$$

Thus $\deg f' < n$ and hence by the induction hypothesis, $f'(x) = q'(x) g(x) + r(x)$ where $r = 0$ or $\deg(r) < \deg(g)$. Therefore,

$$\begin{aligned} f(x) &= \frac{a_n}{b_m} x^{n-m} g(x) + f'(x) = \frac{a_n}{b_m} x^{n-m} g(x) + q'(x) g(x) + r(x) \\ &= \left[\frac{a_n}{b_m} x^{n-m} + q'(x) \right] g(x) + r(x) \\ &= q(x) g(x) + r(x) \end{aligned}$$

where $q(x) = \frac{a_n}{b_m} x^{n-m} + q'(x)$ and $r = 0$ or $\deg(r) < \deg g$. ■

17

Lecture 17

We reviewed Corollary 11.5 in preparation for the next quiz.

17.1 Roots of polynomials

Definition 17.1. *Let R be an integral domain and $f(x) \in R[x]$. Then $a \in R$ is a **zero or root** of $f(x)$ if $f(a) = 0$.*

Example 17.2. Let $f(x) = 2x^2 + 5x - 7$ in $\mathbb{Z}[x]$. Then $f(1) = 0$ so 1 is a root while $f(2) = 11 \neq 0$ so 2 is not a root of f.

Corollary 17.3. *Let F be a field, $a \in F$, and $f(x) \in F[x]$. Then $(x-a)\,|\,f(x) \iff f(a) = 0$, i.e. iff a is a root of $f(x)$.*

Remark 17.4. Corollary 17.3 holds more generally in that we may replace F by any commutative ring with identity. Indeed, if $f(x) \in R[x]$ and $f(a) = 0$ for some $a \in R$. Let $g(x) := f(x+a)$, so that $g(x) \in R[x]$ with $g(0) = 0$. Since $g(0) = 0$, $g(x)$ has no constant term which means that we may factor x out of $g(x)$, i.e. $g(x) = xk(x)$ for some $k(x) \in R[x]$. This translates into the statement about $f(x)$;

$$f(x) = g(x-a) = (x-a)\,k(x-a),$$

which shows $x - a$ is a factor of $f(x)$.

Example 17.5. Let $f(x) = x^2 + 5$ has no roots over $\mathbb{R}$ but two roots over $\mathbb{C}$. Indeed,

$$f(x) = \left(x - i\sqrt{5}\right)\left(x + i\sqrt{5}\right)$$

so that $f\left(\pm i\sqrt{5}\right) = 0$.

Example 17.6 (Book Problem 17.13). Consider the polynomial, $f(x) = x^3 + 6$ on $\mathbb{Z}_7[x]$ and observe that

x	0	1	2	3	4	5	6
$f(x)$	6	0	0	5	0	5	5

.

Thus we know that $(x-1)$, $(x-2)$, and $(x-4)$ are all factors of $f(x)$. For example,

$$\begin{array}{r@{\,}l} & \qquad\quad x^2 + x + 1 \\ \cline{2-2} x-1\big) & x^3 + 0x^2 + 0x + 6 \\ & \underline{x^3 - x^2} \\ & \quad + x^2 + \\ & \quad \underline{+ x^2 - x} \\ & \qquad\qquad x + 6 \\ & \qquad\qquad \underline{x - 1} \\ & \qquad\qquad\quad 7 = 0 \end{array}$$

from which it follows that

$$f(x) = (x-1)\left(x^2 + x + 1\right).$$

Let $g(x) = x^2 + x + 1$ and notice that $0 = f(2) = g(2)$ so that $(x-2)$ must divide $g(x)$. Indeed this is the case,

$$\begin{array}{r@{\,}l} & \qquad x + 3 \\ \cline{2-2} x-2\big) & x^2 + x + 1 \\ & \underline{x^2 - 2x} \\ & \quad + 3x + 1 \\ & \quad \underline{+ 3x - 6} \\ & \qquad\quad 7 = 0 \end{array}$$

from which it follows that

$$g(x) = (x-2)(x+3) = (x-2)(x-4).$$

Thus as expected, we have

$$f(x) = (x-1)(x-2)(x-4).$$

Corollary 17.7. *Let F be a field, $f(x) \in F[x]$, and suppose that $\{a_i\}_{i=1}^n \subset F$ is a list of n – distinct zeros of f. Then $\prod_{i=1}^n (x - a_i)$ divides $f(x)$. Alternatively put, there exists $k(x) \in F[x]$ such that*

$$f(x) = k(x)(x-a_1)(x-a_2)\dots(x-a_n). \tag{17.1}$$

Proof. The proof goes by induction on n. When $n = 1$ this is the content of Corollary 17.3. Suppose the result holds for all $k \leq n$ for some $n \geq 1$. Then if $\{a_i\}_{i=1}^{n+1}$ is a list of distinct zeros of f, by corollary 17.3, there exists $g(x) \in F[x]$ such that

$$f(x) = g(x)(x - a_{n+1}). \tag{17.2}$$

Since

$$0 = f(a_i) = g(a_i)(a_i - a_n + 1)$$

and $a_i - a_{n+1} \neq 0$ for all $i \leq n$, it follows that $\{a_i\}_{i=1}^{n}$ are distinct zeros of g. Therefore by the induction hypothesis, there exists, $k(x) \in F[x]$ such that

$$g(x) = k(x)(x - a_1)(x - a_2) \dots (x - a_n). \tag{17.3}$$

Thus it follows from Eqs. (17.2) and (17.3) that

$$f(x) = k(x)(x - a_1)(x - a_2) \dots (x - a_n)(x - a_{n+1})$$

which completes the induction step and the proof. ∎

Corollary 17.8. *Let F be a field and $f(x) \in F[x]$ with $N := \deg(f)$. Then f has at most N distinct roots in F.*

Proof. If $a_1, \dots, a_n$ be distinct zeros of $f(x)$. Then we may write $f(x)$ as in Eq. (17.1) from which it follows that

$$N = \deg(f) = n + \deg(k) \geq n.$$

∎

Example 17.9. Consider

$$f(x) = (x-4)(x-5) = (x+2)(x+1) = x^2 + 3x + 2$$

in $\mathbb{Z}_6[x]$. Clearly $f(4) = f(5) = 0$ but this is not all. Indeed, $f(0) = 2 \neq 0$, $f(1) = 3 \cdot 2 = 6 = 0$, $f(2) = 4 \cdot 3 = 12 = 0$, $f(3) = (-1)(-2) = 2 \neq 0$. Thus see that $f(x)$ has four zeros in $\mathbb{Z}_6$, namely $\{1, 2, 4, 5\}$.

Example 17.10 (Zeros of $x^n - 1$). Suppose that $z = re^{i\theta}$ is a zero of $x^n - 1$ in $\mathbb{C}$. Then we must have

$$r^n e^{in\theta} = 1$$

which implies $r = 1$ and $in\theta = k2\pi$ for some $k \in \mathbb{Z}$. Thus the zeros, $\mathcal{Z}$, of $x^n - 1$ are,

$$\mathcal{Z} = \left\{e^{ik2\pi/n} : k \in \mathbb{Z}\right\} = \left\{e^{ik2\pi/n} : k \in \mathbb{Z}_n\right\}.$$

If we let $\omega := e^{i2\pi/n}$, then we may write

$$\mathcal{Z} = \left\{\omega^k : k \in \mathbb{Z}_n\right\}$$

and ω is called a primitive n^{th} – root of unity.

17.2 Roots with multiplicities

Definition 17.11. *Let F be a field and $f(x) \in F[x]$. A root, $a \in F$, of $f(x)$ is said to have **multiplicity** $k \geq 1$ if $(x-a)^k$ divides $f(x)$ but $(x-a)^{k+1}$ does not.*

Example 17.12. In $\mathbb{R}[x]$, 3 is a root of order 2 for $f(x) = x^2 - 6x + 9$. Indeed, $f(x) = (x-3)^2$. If $f(x) = x^2 - 7x + 10$, then $f(x) = (x-2)(x-5)$ and the only zeros of f are $x = 2$ and $x = 5$ each of which have multiplicity 1.

Example 17.13. Since $f(x) = x^3 - 2x^2 + x \in \mathbb{Q}[x]$ factors as,

$$f(x) = x\left(x^2 - 2x + 1\right) = x(x-1)^2,$$

the roots of $f(x)$ are 0 and 1 with multiplicities 1 and 2 respectively.

Example 17.14. In $\mathbb{R}[x]$ the polynomial, $f(x) = x^4 + 2x^2 + 1$ has no roots. While in $\mathbb{C}[x]$ it has two distinct roots each with multiplicity 2. To find these roots observe that

$$f(x) = \left(x^2 + 1\right)^2 = [(x-i)(x+i)]^2 = (x-i)^2(x+i)^2.$$

Thus the roots are $\pm i$.

Example 17.15 (17.23). Find all of the zeros and multiplicity of

$$f(x) = x^5 + 4x^4 + 4x^3 - x^2 - 4x + 1 \in \mathbb{Z}_5[x].$$

We start by finding all of the roots;

x	0	1	2	3	4
$f(x)$	1	0	2	0	3

.

Then we divide $(x-1)$ into $f(x)$ to find,

$$\begin{array}{r|l}
 & x^4 + 4x^2 + 3x - 1 \\
\hline
x - 1 & x^5 + 4x^4 + 4x^3 - x^2 - 4x + 1 \\
 & \underline{x^5 - x^4} \\
 & \qquad 4x^3 - x^2 - 4x + 1 \\
 & \qquad \underline{4x^3 - 4x^2} \\
 & \qquad\qquad 3x^2 - 4x + 1 \\
 & \qquad\qquad \underline{3x^2 - 3x} \\
 & \qquad\qquad\qquad - x + 1 \\
 & \qquad\qquad\qquad \underline{- x + 1} \\
 & \qquad\qquad\qquad\qquad 0
\end{array}$$

to find

$$f(x) = (x-1)\,g(x)$$

with $g(x) = x^4 + 4x^2 + 3x - 1$. Let us check we got this right,

$$\begin{aligned}(x-1)\left(x^4 + 4x^2 + 3x - 1\right) &= x^5 - x^4 + 4x^3 - x^2 - 4x + 1 \\ &= x^5 + 4x^4 + 4x^3 - x^2 - 4x + 1.\checkmark\end{aligned}$$

Notice that $g(1) = 7 \bmod 5 = 2 \neq 0$ so that 1 is a root with multiplicity 1. We now divide $(x-3)$ into $g(x)$ to find;

$$\begin{array}{r|rrrrr}
\multicolumn{1}{r}{} & & x^3 & +\,3x^2 & +\,3x & +\,2 \\ \cline{2-6}
x-3 & x^4 & +\,0x^3 & +\,4x^2 & +\,3x & -\,1 \\
\multicolumn{1}{r}{} & \underline{x^4} & \underline{-\,3x^3} & & & \\
\multicolumn{1}{r}{} & & 3x^3 & +\,4x^2 & +\,3x & -\,1 \\
\multicolumn{1}{r}{} & & \underline{3x^3} & \underline{-\,4x^2} & & \\
\multicolumn{1}{r}{} & & & +\,3x^2 & +\,3x & -\,1 \\
\multicolumn{1}{r}{} & & & +\,\underline{3x^2} & \underline{-\,4x} & \\
\multicolumn{1}{r}{} & & & & +\,2x & -\,1 \\
\multicolumn{1}{r}{} & & & & +\,\underline{2x} & \underline{-\,1} \\
\multicolumn{1}{r}{} & & & & & 0
\end{array}$$

so that

$$g(x) = (x-3)\left(x^3 + 3x^2 + 3x + 2\right).$$

Let $h(x) := x^3 + 3x^2 + 3x + 2$, then $h(3) = 2 \cdot 3^3 + 3^2 + 2 = 65 \bmod 5 = 0$ so that $(x-3)$ goes into $h(x)$. Here is the computation,

$$\begin{array}{r|rrrr}
\multicolumn{1}{r}{} & & x^2 & +\;x\;+ & 1 \\ \cline{2-5}
x-3 & x^3 & +\,3x^2 & +\,3x\;+ & 2 \\
\multicolumn{1}{r}{} & \underline{x^3} & \underline{-\,3x^2} & & \\
\multicolumn{1}{r}{} & & x^2 & +\,3x\;+ & 2 \\
\multicolumn{1}{r}{} & & \underline{x^2} & \underline{-\,3x} & \\
\multicolumn{1}{r}{} & & & +\;x\;+ & 2 \\
\multicolumn{1}{r}{} & & & +\;\underline{x\;-} & \underline{3} \\
\multicolumn{1}{r}{} & & & & +\,5 = 0
\end{array}.$$

Thus we have shown,

$$x^5 + 4x^4 + 4x^3 - x^2 - 4x + 1 = (x-1)(x-3)^2\left(x^2 + x + 1\right).$$

so that 3 has multiplicity 2. This is rather painful way to carry this out. We will later develop a derivative test to make finding multiplicities easier to determine.

18

Lecture 18

Corollary 17.7 has the following useful refinement.

Theorem 18.1. *Suppose that $a_1, \dots, a_n$ are distinct zeros of $f(x) \in F[x]$ with multiplicities, $l_1, \dots, l_n$. Then there exists $k(x) \in F[x]$ such that*

$$f(x) = k(x)(x - a_1)^{l_1} \dots (x - a_n)^{l_n}. \tag{18.1}$$

Proof. The proof will be by induction on $N := \sum_{i=1}^n l_i$. If $N = 1$, then we have $n = 1$, $a_1 = a \in F$ and $l_1 = 1$. In this case it follows by definition of a root with multiplicity 1 that $f(x) = k(x)(x - a)$ for some $k(x) \in F[x]$.

Now suppose $N \geq 2$ and the theorem holds whenever $\sum_{i=1}^n l_i < N$. By the induction hypothesis there exists $k_1(x) \in F[x]$ such that

$$f(x) = k_1(x)(x - a_1)^{l_1 - 1}(x - a_2)^{l_2} \dots (x - a_n)^{l_n}. \tag{18.2}$$

Moreover since $(x - a_1)^{l_1} | f(x)$ it follows that $(x - a_1)$ divides

$$f_1(x) = k_1(x)(x - a_2)^{l_2} \dots (x - a_n)^{l_n}$$

which implies $f_1(a_1) = 0$. Since $f_1(a_1) = 0$ while

$$(x - a_2)^{l_2} \dots (x - a_n)^{l_n} |_{x = a_1} \neq 0,$$

we may conclude that $k_1(a_1) = 0$. Therefore that $(x - a_1) | k_1(x)$, i.e. $k_1(x) = k(x)(x - a_1)$ for some $k(x) \in F[x]$. Using this expression for $k_1(x)$ back in Eq. (18.2) completes the proof. ∎

Corollary 18.2. *Let F be a field and $f(x) \in F[x]$ with $N := \deg(f)$. Then f has at most N roots in F when counted with multiplicities. In particular, there can be at most N distinct roots of $f(x)$ in F.*

Proof. If $a_1, \dots, a_n$ be zeros of $f(x)$ with multiplicities, $l_1, \dots, l_n$. Then from Theorem 18.1 there exists $k(x) \in F[x]$ such that Eq. (18.1) holds. In particular it follows that

$$N := \deg(f) = \deg(k) + \sum_{i=1}^n l_i \geq \sum_{i=1}^n l_i.$$

This inequality is precisely what the Corollary states. ∎

18.1 Irreducibles and Maximal Ideals

Definition 18.3. *Let R be an integral domain and $a \in R^\times \setminus U(R)$. We say that a is* ***reducible*** *if it admits a* ***non-trivial factorization****, i.e. $a = bc$ for come $b, c \in R^\times \setminus U(R)$. Otherwise we say that a is* ***irreducible.*** *So a is irreducible iff $a \neq 0$, $a \notin U(R)$, and whenever $a = bc$ then either b or c is in $U(R)$.*

Let F be a field and recall from Lemma 16.6 that $U(F[x]) = U(F) = F^\times$. Therefore the associates to $f(x) \in F[x]$ are $\{af(x) : a \in F^\times\}$. So $f(x) = a \cdot h(x)$ with $a \in F^\times$ and $h(x) \in F[x]$ is a trivial factorization.

Example 18.4. If F is a field then $f(x) \in F[x]$ is reducible iff there is a factorization of the form $f(x) = g(x)h(x)$ where $\deg g \geq 1$ and $\deg h \geq 1$.

Lemma 18.5. *If F is a field and $f(x) = g \in F[x]$ is reducible, then* $\deg(f(x)) \geq 2$. *In particular if* $\deg(f(x)) = 1$ *then $f(x)$ is irreducible.*

Proof. If $f(x)$ admits a non-trivial factorization, $f(x) = g(x)h(x)$, then

$$\deg f(x) = \deg h(x) + \deg g(x) \geq 1 + 1 \geq 2.$$

∎

Example 18.6. In $\mathbb{Z}[x]$;

1. $x^2 - 1$ is reducible since $x^2 - 1 = (x - 1)(x + 1)$.
2. $2x + 4 = 2(x + 2)$ is reducible in $\mathbb{Z}[x]$ since both 2 and $x + 2$ are not units in $\mathbb{Z}[x]$. Similarly $5x \in \mathbb{Z}[x]$ is reducible since $f(x) = 5 \cdot x$ where both 5 and x are not units.

Notice that $2x + 4 = 2(x + 2)$ is irreducible in $\mathbb{Q}[x]$ by Lemma 18.5. The difference now is that 2 is invertible in $\mathbb{Q}[x]$ while it is not in $\mathbb{Z}[x]$.

Proposition 18.7. *Suppose that D is an integral domain and $0 \neq a \in D \setminus U(D)$. Then a is irreducible iff $\langle a \rangle$ is maximal among all principle ideals in D. To say $\langle a \rangle$ is maximal among all principle ideals in D we mean if $b \in D$ satisfies $\langle a \rangle \subset \langle b \rangle \subset R$, then either $\langle a \rangle = \langle b \rangle$ or $\langle b \rangle = R$.*

Proof. You will prove this in the homework. ■

The following theorem is a direct consequence of Proposition 18.7 and Theorem 14.3.

Theorem 18.8 (p irreducible $\iff$ $\langle p\rangle$ maximal in a PID). *Suppose that D is a PID and $p \in D^{\times} \setminus U(D)$. Then the following are equivalent;*

1. *p is irreducible,*
2. *$\langle p\rangle$ is a maximal ideal, and*
3. *$D/\langle p\rangle$ is a field.*

Example 18.9. In $\mathbb{R}[x]$, x^2+1 is irreducible. Indeed if x^2+1 where to factor non-trivially the factors would have to be linear and x^2+1 would have to have a root which it does not. Consequently we know that $\mathbb{R}[x]/\left\langle x^2+1\right\rangle$ is a field. We already know this since we have seen, $\mathbb{R}[x]/\left\langle x^2+1\right\rangle$ is isomorphic to $\mathbb{C}$. Nevertheless, we have now shown that $\mathbb{R}[x]/\left\langle x^2+1\right\rangle$ is a field without knowledge about $\mathbb{C}$ and hence we may view this as a fresh construction of $\mathbb{C}$.

19

Lecture 19

19.1 Irreducibles Polynomials I

Let F be a field – later it will be $\mathbb{Q}$. Our goal is to determine when a polynomial, $p(x) \in F[x]$, is irreducible. We start with the following general results.

Example 19.1 (Second proof of Lemma 18.5). Let F be a field and $\lambda \in F$. Then $\varphi : F[x] \to F$ be the evaluation homomorphism, $\varphi(p) = p(\lambda)$. Then $\varphi(F[x]) = F$, $I = \ker(\varphi) = \langle x - \lambda \rangle = F[x] \cdot (x - \lambda)$, and therefore, $F[x] / \langle x - \lambda \rangle \cong F$ and therefore $x - \lambda$ is irreducible in $F[x]$ for all $\lambda \in F$ by Theorem 18.8.

Lemma 19.2. *If F is a field, $f(x) \in F[x]$ with $\deg(f) \geq 2$, and f has a zero in F, then f is reducible over F.*

Proof. Let $a \in F$ be a root of f, then $f(x) = (x - a)k(x)$ for some $k(x) \in F(x)$ with $\deg(k) \geq 1$. Since both $(x - a)$ and $k(x)$ are not units, it follows that f is reducible. ■

We have the following partial converse to this lemma.

Theorem 19.3. *If F is a field, $f(x) \in F[x]$ with $\deg(f) = 2$ or $\deg(f) = 3$, then f is irreducible over F iff f has no zeros in F.*

Proof. Lemma 19.2 shows that f has zero implies f is reducible. Conversely if f is reducible, then $f(x) = p(x)q(x)$ for some polynomials p and q with $\deg(p), \deg(q) \geq 1$. Since $\deg(p) + \deg(q) = \deg(f) \leq 3$, it follows that $\deg(p) = 1$ or $\deg(q) = 1$, say $\deg(p) = 1$. Thus $p(x) = ax + b$ for some $a, b \in F$ with $a \neq 0$ and $f(x) = (ax + b)q(x)$. Therefore

$$f\left(\frac{-b}{a}\right) = \left(a\frac{-b}{a} + b\right) q\left(\frac{-b}{a}\right) = 0 \cdot q\left(\frac{-b}{a}\right)$$

which shows that f has a zero in F. ■

Example 19.4 (Construction of finite fields). Observe that the x^3 function on $\mathbb{Z}_5$ is given by

x	0	1	2	3	4
x^3	0	1	3	2	4

.

Let $f(x) := x^3 + x + 1 \in \mathbb{Z}_5[x]$, then

x	0	1	2	3	4
$f(x)$	1	3	1	1	4

showing $f(x)$ has not roots and hence is irreducible. Therefore $\langle f(x) \rangle \subset \mathbb{Z}_5[x]$ is a maximal ideal and therefore,

$$F := \mathbb{Z}_5[x] / \langle f(x) \rangle$$

is a field. As we have,

$$F = \left\{ \left[a + bx + cx^2\right] : a, b, c \in \mathbb{Z}_5 \right\},$$

we see that the number of elements in F is $5^3 = 125$. Thus we have constructed a field with 125 elements and characteristic 5.

Proposition 19.5. *Suppose that $f(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{Z}[x]$. If and $f(r/s) = 0$ for some $r \in \mathbb{Z}$ and $s \in \mathbb{Z}_+$ with $\gcd(r, s) = 1$, .then $s | a_n$ and $r | a_0$. In particular if $f(x)$ is monic, i.e. $a_n = 1$, the only possible rational roots of $f(x)$ must actually be in $\mathbb{Z}$ and in fact must be a divisor of a_0.*

Proof. See problem **17.25** in Gallian which you are assigned for homework. ■

Example 19.6. Let $k \geq 2$ and p be a prime in $\mathbb{Z}_+$, then $\sqrt[k]{p}$ is irrational. Indeed if $\sqrt[k]{p} = r$ were rational, being a root of $x^k - p$, r would have to be an integer which divides p by Proposition 19.5. Thus r would have to be $\pm p$ or ± 1 none of which will work. Hence we may conclude that $x^2 - p$ and $x^3 - p$ are irreducible polynomials over $\mathbb{Q}$ (or $\mathbb{Z}$) for any prime number p. (In fact by Eisenstein's Criterion, Theorem 22.6 below, it is easy to see that $x^k - p$ is irreducible over $\mathbb{Q}$ for all primes, p, and $k \geq 2$.) Of course

$$x^2 - 2 = \left(x - \sqrt{2}\right)\left(x + \sqrt{2}\right) \text{ in } \mathbb{R}[x]$$

and therefore $x^2 - 2$ is reducible over $\mathbb{R}$.

Example 19.7. In this example we consider the polynomial, $x^2 + 1 \in F[x]$ for a number of fields, F.

1. If $F = \mathbb{R}$, then $f(x) = x^2 + 1 \geq 1 > 0$ for $x \in \mathbb{R}$ and so x^2+1 is irreducible in $\mathbb{R}[x]$. Consequently, $\mathbb{R}[x] / \langle x^2+1 \rangle$ is a field. (We already have seen this isomorphic to $\mathbb{C}$.)
2. If $F = \mathbb{C}$, then $x^2 + 1 = (x-i)(x+i)$ is reducible over $\mathbb{C}[x]$.
3. If $F = \mathbb{Z}_3$, then $f(0) = 1$, $f(1) = 2$ and $f(2) = 5 \bmod 3 = 2$ and therefore f is irreducible over $\mathbb{Z}_3$. Consequently $\mathbb{Z}_3[x] / \langle x^2+1 \rangle$ is a field which is easily seen to be isomorphic to $\mathbb{Z}_3[i]$.
4. If $F = \mathbb{Z}_5$, then $f(2) = 2^2 + 1 = 5 \bmod 5 = 0$ and $f(3) = 10 \bmod 5 = 0$ and therefore $f(x)$ is reducible over $\mathbb{Z}_5$. Notice that
$$f(x) = (x-2)(x-3) = (x+3)(x+2).$$
In this case $\mathbb{Z}_5[i] \cong \mathbb{Z}_5[x] / \langle x^2+1 \rangle$ is not a field and in fact not even an integral domain since $[x+2][x+3] = 0$.

Example 19.8. The polynomial $f(x) = x^3 + x + 1$ is irreducible over $\mathbb{Z}_2[x]$ since $f(0) = 1 = f(1)$.

Example 19.9. The polynomial, $f(x) = x^4 + 2x^2 + 1$ has no roots over $\mathbb{R}$ yet it is reducible over $\mathbb{R}$. Indeed we have the non-trivial factorizations,
$$f(x) = \left(x^2+1\right)\left(x^2+1\right).$$
This shows the hypothesis that $\deg(f) \in \{2,3\}$ is necessary in Theorem 19.3.

Example 19.10. Let $f(x) = 3x^3 + 2x + 1$ in $\mathbb{Z}_5[x]$. Then

x	0	1	2	3	4
$f(x)$	1	1	4	3	1

and it follows that f is irreducible over $\mathbb{Z}_5$.

Example 19.11. Suppose that $f(x) = x^4 + x + 1 \in \mathbb{Z}_2[x]$. In this case $f(0) = 1 = f(1)$ so that f has no roots and hence not linear factors. Thus if $f(x)$ is to factor it must be of the form $f(x) = p(x)q(x)$ where $\deg p(x) = 2 = \deg q(x)$ and both $p(x)$ and $q(x)$ have no roots. By dividing $p(x)$ by its leading order coefficient we may assume that p is monic. This forces q to be monic as well since f was monic. Thus we have, $p(x) = x^2 + ax + b$ form some $a, b \in \mathbb{Z}_2$. Since $0 \neq p(0) = b$ we must have $p(x) = x^2 + ax + 1$. Similarly, $0 \neq p(1) = a$ so that $a = 1$. Thus the only monic degree 2 polynomial which is itself irreducible is $x^2 + x + 1$. Thus if there is going to be a factorization of $f(x)$ it must be given by $f(x) = \left(x^2+x+1\right)^2$. However a simple computation shows,
$$\left(x^2+x+1\right)^2 = 2x + 3x^2 + 2x^3 + x^4 + 1 = x^4 + x^2 + 1 \neq x^4 + x + 1.$$
Alternatively we could divide $x^2 + x + 1$ into $f(x)$ as follows,

$$\begin{array}{r|l}
 & x^2 + x \\ \hline
x^2+x+1 \,\big) & x^4 + 0x^3 + 0x^2 + x + 1 \\
 & \underline{x^4 + x^3 + x^2} \\
 & x^3 + x^2 + x + 1 \\
 & \underline{x^3 + x^2 + x} \\
 & 1
\end{array}$$

which shows that
$$x^4 + x + 1 = \left(x^2+x\right)\left(x^2+x+1\right) + 1$$
and therefore $\left(x^2+x+1\right)$ is not a factor of $f(x)$.

Theorem 19.12 (Fundamental theorem of algebra). *The complex number field, $\mathbb{C}$, is* ***algebraically closed,*** *i.e. every non-constant polynomial, $p(x) \in \mathbb{C}[x]$ has a root.*

Proof. This is a standard result proved in a course on complex variables. We will not give the proof in this class. ■

Corollary 19.13. *Let $p(x) \in \mathbb{C}[x]$ and $\{\lambda_i\}_{i=1}^m$ be the distinct zeros of $p(x)$ and $\{k_i\}_{i=1}^m$ be the corresponding multiplicities. Then*
$$p(x) = c \prod_{i=1}^m (x-\lambda_i)^{k_i} \text{ for some } c \in \mathbb{C}^\times. \tag{19.1}$$

Proof. From Theorem 18.1 we know that exists $k(x) \in \mathbb{C}[x]$ such that
$$p(x) = k(x) \prod_{i=1}^m (x-\lambda_i)^{k_i}.$$
However, $k(x)$ must not have any roots for otherwise we would not have accounted for all the zeros with multiplicities of $p(x)$. According to Theorem 19.12, $k(x)$ must be a constant polynomial c for some $c \in \mathbb{C}$. ■

20

Lecture 20

Corollary 20.1. *The only irreducible polynomials in $\mathbb{C}[x]$ are those with degree equal to one. The only irreducible polynomials in $\mathbb{R}[x]$ are the degree one polynomials and the degree two polynomials with no roots.*

Proof. The first assertion should be clear from Corollary 19.13. For the second assertion, suppose that $p(x) \in \mathbb{R}[x]$ is any polynomial. If $\lambda \in \mathbb{C}$ is a root of $p(x)$, then $0 = \overline{p(\lambda)} = p(\bar{\lambda})$ because $p(x)$ has all real coefficients. Let us now suppose that $p(x)$ is irreducible over $\mathbb{R}$.

If $\lambda \in \mathbb{R}$ is a root of $p(x)$, then $p(x) = (x - \lambda) k(x)$ for some $k(x) \in \mathbb{R}[x]$. Since $p(x)$ is irreducible we must have $\deg(k(x)) = 0$, i.e. $.k(x)$ is a constant and $p(x)$ is linear. On the other hand if $\lambda \in \mathbb{C} \setminus \mathbb{R}$ is a root of $p(x)$ then so is $\bar{\lambda}$. Therefore we know

$$p(x) = (x - \lambda)\left(x - \bar{\lambda}\right) k(x) = \left(x^2 - 2\operatorname{Re}\lambda + |\lambda|^2\right) k(x)$$

for some $k(x) \in \mathbb{R}[x]$. Since $p(x)$ is irreducible, we must again have $\deg k(x) = 2$. This observation along with Theorem 19.3 completes the proof of the corollary. ■

Proposition 20.2. *Let $p(x) \in \mathbb{R}[x]$, then $p(x)$ may be factored into irreducibles of $\mathbb{R}[x]$.*

Proof. If $p(x)$ is not irreducible, there exists $h(x), k(x) \in \mathbb{R}[x]$ with $\deg h \geq 1$, $\deg k \geq 1$, and $p(x) = h(x) k(x)$. Since $\deg h(x) < \deg p(x)$ and d $\deg k(x) < \deg p(x)$, the result follows by a simple induction argument. ■

20.1 Two more homomorphisms involving polynomials

Lemma 20.3. *Let R and T be commutative rings, $\varphi : R[x] \to T$ be a ring homomorphism, and let $t := \varphi(x)$. Then*

$$\varphi\left(\sum_{k=0}^{n} a_k x^k\right) = \sum_{k=0}^{n} \varphi(a_k) t^k \text{ for all } \sum_{k=0}^{n} a_k x^k \in R[x].$$

Proof. Using the ring homomorphism properties we find,

$$\begin{aligned}\varphi(p(x)) &= \sum_{k=0}^{n} \varphi\left(a_k x^k\right) = \sum_{k=0}^{n} \varphi(a_k) \varphi\left(x^k\right) \\ &= \sum_{k=0}^{n} \varphi(a_k) \varphi(x)^k = \sum_{k=0}^{n} \varphi(a_k) t^k.\end{aligned}$$

■

Proposition 20.4 (Changing coefficients). *Let R and T be commutative rings and $\psi : R \to T$ be a ring homomorphism. Further define $\bar{\psi} : R[x] \to T[x]$ by the formula,*

$$\bar{\psi}\left(\sum_{k=0}^{n} a_k x^k\right) := \sum_{k=0}^{n} \psi(a_k) x^k \text{ for all } \sum_{k=0}^{n} a_k x^k \in R[x]. \tag{20.1}$$

Then $\bar{\psi}$ is a ring homomorphism. Moreover this is the unique ring homomorphism from $R[x] \to T[x]$ such that $\bar{\psi}(x) = x$ and $\bar{\psi}(a) = \psi(a)$ for all $a \in R \subset R[x]$.

Proof. You are asked to prove this for homework. ■

Example 20.5. Suppose that $\psi_k : \mathbb{Z} \to \mathbb{Z}_k$ is the homomorphism, $\psi_k(a) := a \bmod k$ so that $\bar{\psi}_k : \mathbb{Z}[x] \to \mathbb{Z}_k[x]$. For example,

$$\begin{aligned}\bar{\psi}_3\left(5x^3 + 7x + 3\right) &= 2x^3 + x \in \mathbb{Z}_3[x], \\ \bar{\psi}_5\left(5x^3 + 7x + 3\right) &= 2x + 3 \in \mathbb{Z}_5[x], \text{ and} \\ \bar{\psi}_7\left(5x^3 + 7x + 3\right) &= 5x^3 + 3 \in \mathbb{Z}_7[x].\end{aligned}$$

Corollary 20.6. *Let R and T be commutative rings and $\varphi : R \to T$ be a ring homomorphism. Then for each $t \in T$ there exists a ring homomorphism $\varphi_t : R[x] \to T$ such that $\varphi_t(a) = \varphi(a)$ for all $a \in R$ and $\varphi_t(x) = t$. More specifically, if*

$$\varphi_t\left(\sum_{k=0}^{n} a_k x^k\right) = \sum_{k=0}^{n} \varphi(a_k) t^k \text{ for all } \sum_{k=0}^{n} a_k x^k \in R[x].$$

Proof. Let $\psi : R[x] \to T[x]$ be the homomorphism in Proposition 20.4 and $ev_t : T[x] \to T$ be the evaluation homomorphism at t. Then $\varphi_t = ev_t \circ \psi : R[x] \to T$ is the desired homomorphism. ■

Example 20.7. As an application of Corollary 20.6, let R be a commutative ring, $\varphi : R \to R[x]$ be the homomorphism taking $a \in R$ to the constant polynomial $a \in R[x]$, and $t(x) \in R[x]$ be any polynomial. Then there is a unique homomorphism, $\varphi_t : R[x] \to R[x]$ such that $\varphi_t(a) = a$ for all $a \in R$ and $\varphi_t(x) = t(x)$ which is given formulaically as,

$$\varphi(p(x)) = p(t(x)) \text{ for all } p(x) \in R[x].$$

In particular if $t(x) = x - a$ for some $a \in R$ the map, $p(x) \to p(x-a)$ is a ring isomorphism from $R[x]$ to $R[x]$ with inverse being given by $p(x) \to p(x+a)$.

20.2 Gauss' Lemma

We now go back to finding more ways to determine when $f(x) \in \mathbb{Q}[x]$ is irreducible. Are first goal is to change the question into one involving a polynomial $g(x) \in \mathbb{Z}[x]$ being irreducible or not – see Corollary 21.5 below. The main tool here will be Gauss' Lemma 20.10 below.

Definition 20.8. *Let $f(x) = \sum_{k=0}^{n} a_n x^n \in \mathbb{Z}[x]$, then $c(f) := \gcd(a_0, \dots, a_n)$ is called the* ***content*** *of f. A polynomial $f \in \mathbb{Z}[x]$ is said to be* ***primitive*** *if $c(f) = 1$.*

Example 20.9. If $f(x) = 14x^2 + 10x + 6$, then $c(f) = 2$ and $\frac{1}{2}f(x) = 7x^2 + 5x + 3$ is a primitive polynomial.

Lemma 20.10 (Gauss' Lemma). *If $f(x), g(x) \in \mathbb{Z}[x]$ are primitive polynomials, then $f(x)g(x)$ is primitive as well, i.e. $c(f) = 1 = c(g) \implies c(fg) = 1$.*

Proof. Suppose that $c(fg) > 1$ and let p be a prime divisor of $c(fg)$. Further let $\bar{f}(x) := \sum_{k=0}^{n} (a_n \bmod p) x^n$ and similarly for $\bar{g}(x)$. Observe that $f \to \bar{f}$ is a ring homomorphism from $\mathbb{Z}[x] \to \mathbb{Z}_p[x]$ so that $\overline{fg} = \bar{f}\bar{g}$. Since p divides all of the coefficients of fg we know $\overline{fg} = 0$ and since $\mathbb{Z}_p[x]$ is an integral domain, we may conclude that either $\bar{f} = 0$ or $\bar{g} = 0$, i.e. $p|c(f)$ or $p|c(g)$. Thus at least one of the polynomials, f or g, is not primitive. ■

Lemma 20.11. *Suppose that $a_0, \dots, a_n \in \mathbb{Z}$ (not all zero) and $m \in \mathbb{Z}_+$. Then $\gcd(ma_0, \dots, ma_n) = m \gcd(a_0, \dots, a_n)$. In particular if $f(x) \in \mathbb{Z}[x]$ and $m \in \mathbb{Z}^\times$, then $c(mf) = |m| c(f)$.*

Proof. First Proof. One way to compute the gcd of a bunch of numbers is by looking at their prime number decompositions. In terms of this decomposition,

$$\gcd(a_0, \dots, a_n) = \prod_{i=1}^{N} p_i^{k_i}$$

where $\{p_i\}_{i=1}^{N}$ are the distinct primes which appear in these decomposition and $k_i \in \{0, 1, 2, \dots\}$ is chosen so that $p_i^{k_i} | a_j$ for all j but $p_i^{k_i+1}$ does not divide one of the a_j. Using this description it is easy to see the truth of the lemma.

Second Proof. Let $d := \gcd(a_0, \dots, a_n)$ and $c := \gcd(ma_0, \dots, ma_n)$ then we know there exists $s_i \in \mathbb{Z}$ such that

$$d = s_0 a_0 + \dots + s_n a_n.$$

From this it follows that

$$md = s_0 m a_0 + \dots + s_n m a_n.$$

Thus every common divisor of ma_i for all i is a divisor of md and in particular $c|(md)$. Since md is a divisor of ma_i for all i we also know that $md|c$. Thus we may conclude that $md = c$ as claimed. ■

Corollary 20.12. *If $f, g \in \mathbb{Z}[x]$ then $c(fg) = c(f) \cdot c(g)$. In particular, $f(x)g(x)$ is primitive iff both $f(x)$ and $g(x)$ are primitive.*

Proof. Let $f_0(x) := \frac{1}{c(f)} f(x)$ and $g_0(x) := \frac{1}{c(g)} g(x)$ so that both $f_0(x)$ and $g_0(x)$ are primitive in $\mathbb{Z}[x]$. Then by Gauss' lemma, $f_0(x)g_0(x)$ is primitive and since $f(x)g(x) = c(f)c(g)f_0(x)g_0(x)$, it follows from Lemma 20.11 that $c(fg) = c(f)c(g)$. ■

Corollary 20.13. *Suppose that $f(x), g(x) \in \mathbb{Z}[x]$, $g(x)$ is primitive, and $k(x) \in \mathbb{Q}[x]$. If $f(x) = k(x)g(x)$ in $\mathbb{Q}[x]$, then $k(x) \in \mathbb{Z}[x]$. In words if $g(x)$ divides $f(x)$ in $\mathbb{Q}[x]$ then the quotient is in fact back in $\mathbb{Z}[x]$.*

Proof. Let $t \in \mathbb{Z}_+$ be chosen so that $h(x) := tk(x) \in \mathbb{Z}[x]$ – for example let t be the least common multiple of all of the denominators in the coefficients of $k(x)$. We then have $tf(x) = h(x)g(x)$ therefore that $tc(f) = c(tf) = c(h)c(g) = c(h)$. It now follows that

$$\frac{1}{c(f)} f(x) = \frac{1}{c(tf)} tf(x) = \frac{1}{c(h)} tf(x) = \frac{1}{c(h)} h(x)g(x) = \frac{1}{c(h)} h(x)g(x).$$

Therefore

$$f(x) = c(f)\frac{1}{c(h)} h(x)g(x) = k(x)g(x)$$

and hence

$$k(x) = c(f)\left(\frac{1}{c(h)}h(x)\right) \in \mathbb{Z}[x].$$

■

Example 20.14. If $g(x)$ is not primitive, the results in Corollary 20.13 may fail. For example take $f(x) = x^2$, $g(x) = 2x$ and $k(x) = \frac{1}{2}x$.

21

Lecture 21

Lemma 21.1. *Suppose that $f(x), g(x) \in \mathbb{Z}[x]$ are primitive polynomials and there exists $\alpha, \beta \in \mathbb{Q}^\times$ such that $\alpha f(x) = \beta g(x)$ as polynomials in $\mathbb{Q}[x]$. Then $|\alpha| = |\beta|$ and $f(x) = \pm g(x)$.*

Proof. Apply Corollary 20.13 with $k(x) = \beta/\alpha$ to learn that $\beta/\alpha = t \in \mathbb{Z}$. Since $f(x) = tg(x)$ it follows that $1 = c(f) = |t|\, c(g) = |t|$ so that $\beta/\alpha = t = \pm 1$.

Alternative Proof. Let $\alpha := \frac{a}{b}$ and $\beta := \frac{u}{v}$ with $a, b, u, v \in \mathbb{Z}\setminus\{0\}$. Then $vaf(x) = ubg(x)$. From Lemma 20.11 we know

$$|va| = c(vaf(x)) = c(ubg(x)) = |ub|$$

from which it follows that $f(x) = \pm g(x)$. ■

Example 21.2. The polynomial, $f(x) = x^2 - 5x + 6 \in \mathbb{Z}[x]$ has a factorization over $\mathbb{Q}$ as

$$f(x) = \left(\frac{1}{2}x - \frac{3}{2}\right)(2x - 4).$$

Of course if we move the factor of 2 from the second factor to the first we also have the factorizations of $f(x)$ over $\mathbb{Z}$ as,

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

The next theorem shows this sort of rejiggering can always be done.

Theorem 21.3. *Let $f(x) \in \mathbb{Z}[x]$. If $f(x)$ is reducible over $\mathbb{Q}$ then $f(x)$ is reducible over $\mathbb{Z}$. In fact $f(x)$ is reducible over $\mathbb{Q}[x]$ iff $f(x)$ factors in $\mathbb{Z}[x]$ in the form, $f(x) = u(x)v(x)$, where $u(x),\ v(x) \in \mathbb{Z}[x]$ with $\deg u(x) \geq 1$ and $\deg v(x) \geq 1$.*

Consequently: *if $f(x)$ is irreducible over $\mathbb{Z}$ then $f(x)$ is also irreducible over $\mathbb{Q}$.*

Proof. We need to show that any non-trivial factorization over $\mathbb{Q}$ gives rise to a non-trivial factorization over $\mathbb{Z}$. So suppose $f(x) \in \mathbb{Z}[x]$ and $h(x), g(x) \in \mathbb{Q}[x]$ with $\deg(h) \geq 1$, $\deg(g) \geq 1$, and $f(x) = g(x)h(x)$ in $\mathbb{Q}[x]$. Choose $s, t \in \mathbb{Z}_+$ such that $g_1(x) := sg(x)$ and $h_1(x) := th(x)$ and in $\mathbb{Z}[x]$. Then we have

$$stf(x) = g_1(x)h_1(x) \implies stc(f) = c(g_1) \cdot c(h_1)$$

and therefore

$$\frac{1}{c(f)}f(x) = \frac{1}{stc(f)}stf(x) = \frac{1}{c(g_1)}g_1(x) \cdot \frac{1}{c(h_1)}h_1(x).$$

Therefore we have found a non-trivial factorization of $f(x)$ over $\mathbb{Z}$ in the desired form, namely;

$$\begin{aligned} f(x) &= \left[c(f)\frac{1}{c(g_1)}g_1(x)\right] \cdot \left[\frac{1}{c(h_1)}h_1(x)\right] \\ &= \left[c(f)\frac{s}{c(g_1)}g(x)\right] \cdot \left[\frac{t}{c(h_1)}h(x)\right] =: u(x)v(x). \end{aligned}$$

■

Example 21.4. Notice that $x \in \mathbb{Z}[x]$ is irreducible. Indeed if $x = f(x)g(x)$ for some $f(x), g(x) \in \mathbb{Z}[x]$, then $1 = \deg(f(x)) + \deg(g(x))$ and we may suppose that $\deg(f) = 1$ and $\deg(g) = 0$. Moreover we know that $f(x)$ and $g(x)$ are primitive and therefore $g(x) = \pm 1 \in U(\mathbb{Z}[x])$. Thus the factorization is trivial and $x \in \mathbb{Z}[x]$ is irreducible.

On the other hand, $f(x) = 5x \in \mathbb{Z}[x]$ is reducible over $\mathbb{Z}$ but irreducible over $\mathbb{Q}$. The point being that $5 \in U(\mathbb{Q})$ while $5 \notin U(\mathbb{Z})$.

However we do have the following corollary to Theorem 21.3.

Corollary 21.5. *If $f(x) \in \mathbb{Z}[x]$ is* ***primitive****, then $f(x)$ is irreducible over $\mathbb{Z}[x]$ iff $f(x)$ is irreducible over $\mathbb{Q}[x]$.*

Proof. 1. First observe that if $g(x) \notin \mathbb{U}(\mathbb{Z}[x])$ and $g(x)$ is primitive then $\deg g(x) \geq 1$. Indeed if $\deg g(x) = 0$ and $g(x)$ is primitive, then $g(x) = \pm 1$, i.e. $g(x) \in \mathbb{U}(\mathbb{Z}[x])$.

2. Suppose that $f(x)$ is primitive and reducible over $\mathbb{Z}$, i.e. $f(x) = g(x)h(x)$ for some $g(x), h(x) \notin U(\mathbb{Z}[x])$. Since $1 = c(f) = c(g)c(h)$ it follows that $c(g) = 1 = c(h)$ and therefore by the first part we must have $\deg g(x) \geq 1$ and $\deg h(x) \geq 1$. Therefore $g(x), h(x) \notin U(\mathbb{Q}[x]) \cong \mathbb{Q}^\times$ and so the factorization $f(x) = g(x)h(x)$ is nontrivial over $\mathbb{Q}$ as well, i.e. $f(x)$ is reducible in $\mathbb{Q}[x]$. Since the converse was already proven in Theorem 21.3, the proof is complete. ■

21.1 mod p Irreducibility Tests

Theorem 21.6 (mod p **test).** *Let $f(x) \in \mathbb{Z}[x]$, $p \in \mathbb{Z}_+$ be a prime, and $\bar{f}(x) \in \mathbb{Z}_p[x]$ be the reduction of $f(x)$ mod p, i.e. reduce all of the coefficients of $f(x)$ mod p. If $\deg \bar{f} = \deg f$,[1] and $\bar{f}(x)$ is irreducible over $\mathbb{Z}_p$, then $f(x)$ is irreducible over $\mathbb{Q}$.*

Proof. If $f(x)$ is reducible over $\mathbb{Q}$, then by Theorems 21.3, there $f(x) = g(x)h(x)$ with $\deg(g), \deg(h) \geq 1$, then $\bar{f}(x) = \bar{g}(x)\bar{h}(x)$. Since $\deg \bar{g} \leq \deg g$ and $\deg \bar{h} = \deg h$ and $\deg f = \deg \bar{f}$ by assumption,

$$\deg f = \deg \bar{f} = \deg \bar{g} + \deg \bar{h} \leq \deg g + \deg h = \deg f.$$

From this equation it follows that we must have $\deg \bar{g} = \deg g \geq 1$ and $\deg \bar{h} = \deg h \geq 1$. Therefore $\bar{f}(x)$ is reducible over $\mathbb{Z}_p$ as well. ■

*Remark 21.7 (*mod p *test failures).* We will see in examples below if $\bar{f}(x) \in \mathbb{Z}_p[x]$ is reducible or $\deg \bar{f} < \deg f$, then the only thing we may conclude is that the test fails and we need to try another p.

Example 21.8. Consider $f(x) = x^4 + x + 1 \in \mathbb{Z}[x]$. We have seen in Example 19.11 that

$$\bar{f}(x) = f(x) \bmod 2 = x^4 + x + 1 \in \mathbb{Z}_2[x]$$

is irreducible and therefore $f(x) = x^4 + x + 1$ is irreducible in $\mathbb{Z}[x]$ and hence also on $\mathbb{Q}[x]$.

Notice that if we had decided test this using the mod 3 test, we would have seen that $\bar{f}(1) = 1 + 1 + 1 = 3 \bmod 3 = 0$ so that $\bar{f}(x)$ is reducible in $\mathbb{Z}_3[x]$. The **only** thing we can conclude from this is that the **test has failed.**

Example 21.9. Now suppose that $f(x) = x^3 + x^2 + x - 1 \in \mathbb{Z}[x]$. In this case the mod 2 test fails but the mod 3 test wins the day. So suppose that $\bar{f}(x) = f(x) \bmod 3$ and observe that

$$\bar{f}(0) \bmod 3 = 2, \ \bar{f}(1) \bmod 3 = 2, \text{ and } \bar{f}(2) \bmod 3 = 1.$$

Since $\bar{f}$ is third order and has not roots it is irreducible and therefore $f(x)$ is irreducible over $\mathbb{Z}$ and hence also $\mathbb{Q}$.

[1] Alternatively put, we are assuming that the leading the order coefficient of $f(x)$ is not divisible by p.

22

Lecture 22

Let $f(x) \in \mathbb{Z}[x]$ and for any prime $p \geq 2$, let $f_p(x) := f(x) \bmod p$. Then combining Theorems 21.3 and 21.6 we have the following implication;

$$\begin{pmatrix} \exists\, p \text{ such that} \\ \deg f_p(x) = \deg f(x) \ \ \& \\ f_p(x) \in \mathbb{Z}_p[x] \ \text{ is irreducible} \end{pmatrix} \implies \begin{pmatrix} f(x) \ \text{ is irreducible} \\ \text{over } \mathbb{Q} \end{pmatrix}.$$

Example 22.1. Let us consider the polynomial,

$$f(x) := 5x^4 - \frac{8}{3}x^3 + 3x + \frac{11}{3} \in \mathbb{Q}[x].$$

Then $f(x)$ is irreducible iff

$$g(x) := 3f(x) = 15x^4 - 8x^3 + 9x + 11$$

is irreducible over $\mathbb{Q}$. As $g(x) \in \mathbb{Z}[x]$ and $c(g) = 1$ we need only show that $g(x)$ is irreducible over $\mathbb{Z}[x]$. To test this out, let $\bar{g}(x)$ be $g(x) \bmod 2$, so that

$$\bar{g}(x) = x^4 + x + 1 \in \mathbb{Z}_2[x].$$

We have already seen in Example 19.11 that $\bar{g}(x)$ is irreducible and therefore by the mod 2 test, $g(x)$ is irreducible and therefore so is $3f(x) \in \mathbb{Z}[x]$ and hence $f(x) \in \mathbb{Q}[x]$.

Example 22.2 (Converse of mod p *test is false).* Let $f(x) := x^2+5x+25 \in \mathbb{Z}[x]$ and $\bar{f}(x) = x^2$ be $f(x)$ mod 5. Notice that $\bar{f}$ is reducible over $\mathbb{Z}_5$ yet $f(x)$ is irreducible over $\mathbb{Q}$. Indeed, it is easy to see that $f(x)$ has not roots in $\mathbb{Z}$ and hence not roots in $\mathbb{Q}$. Alternatively, apply the mod 2 test in which case we consider x^2+x+1 which is irreducible over $\mathbb{Z}_2$ as it has no roots. Alternatively, by the quadratic formula the roots of $f(x)$ are given by

$$\frac{-5 \pm \sqrt{25-100}}{2} = \frac{-5 \pm i\sqrt{75}}{2}$$

so there are not even any real roots. This can also be seen by completing the squares to see,

$$\begin{aligned} f(x) &= \left(x + \frac{5}{2}\right)^2 + 25 - \left(\frac{5}{2}\right)^2 \\ &= \left(x + \frac{5}{2}\right)^2 + 25 \cdot \frac{3}{4} \geq \frac{75}{4}. \end{aligned}$$

Example 22.3 (Degree hypothesis is necessary). Let

$$f(x) = 2x^2 + 3x + 1 = (2x+1)(x+1) \in \mathbb{Q}[x]. \tag{22.1}$$

Clearly this polynomial is reducible. On the on the other hand if we reduce f mod 2 we get the polynomial, $\bar{f}(x) = x+1$ which is linear and hence irreducible. Notice that under the mod 2 reduction, the factorization in Eq. (22.1) becomes,

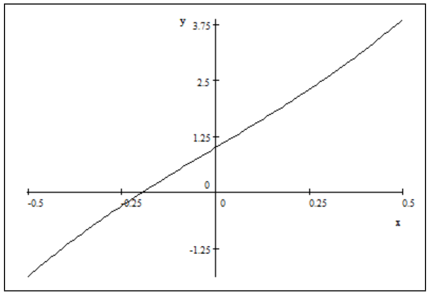
$$x + 1 = 1 \cdot (x+1)$$

which is a trivial factorization of $x+1$.

Example 22.4. Let us determine if

$$f(x) := \frac{3}{7}x^3 + \frac{5}{7}x + \frac{1}{7} \in \mathbb{Q}[x]$$

is irreducible or not. We need only consider $7f(x) = 3x^3 + 5x + 1$ which is primitive in $\mathbb{Z}[x]$.



Plot of $7f(x)$.

Let us consider

$$g(x) := 7f(x) \bmod 2 = x^3 + x + 1.$$

Since $g(0) = g(1) = 1$, it follows that g is irreducible and since $\deg(g) = \deg(7f)$, we learn that $7f$ and hence f is irreducible over $\mathbb{Q}$.

Alternatively we may use Proposition 19.5 to argue that $7f(x)$ and hence $f(x)$ has no roots in $\mathbb{Q}$. According to Proposition 19.5 we need only look for roots of the form $a/3$ where $a \in \{\pm 1\}$ that is we need only consider $\{0, \pm 1/3\}$. By direct calculation we learn

$$7f(-1/3) = -7/9 \neq 0 \text{ and } 7f(1/3) = \frac{25}{9} \neq 0.$$

This again shows that $f(x)$ is irreducible over $\mathbb{Q}[x]$.

Example 22.5. Let us show again the $f(x) = x^3 + 5$ is irreducible over $\mathbb{Q}$.

1. Method 1. By Proposition 19.5 the only possible roots of $f(x)$ are for $x \in \{\pm 5\}$ and it is clear that neither of these are roots of $f(x)$.
2. Method 2. Let us try the $\operatorname{mod} p$ tests.
   a) For $p = 2$ and $p = 3$ the tests fails since 1 is a root in each of these cases.
   b) For $p = 5$, $\bar{f}(x) = x^3$ is clearly reducible so the test fails again.
   c) For $p = 7$ we find
   $$\begin{aligned} &f(0) \operatorname{mod} 7 = 5,\ f(1) \operatorname{mod} 7 = 6,\ f(2) \operatorname{mod} 7 = 6,\ f(3) \operatorname{mod} 7 = 4, \\ &f(4) \operatorname{mod} 7 = 6,\ f(5) \operatorname{mod} 7 = 4,\ f(6) \operatorname{mod} 7 = 4 \end{aligned}$$
   and the test has succeeded in showing that $x^3 + 5$ is irreducible over $\mathbb{Q}$.
3. Method 3. Use Eisenstein's criteria in Theorem 22.6 below.

22.1 Eisenstein's Criterion

Theorem 22.6 (Eisenstein's Criterion). *Let $f(x) := \sum_{k=0}^{n} a_k x^k \in \mathbb{Z}[x]$ with $n \geq 2$. If there is a prime p such that $p \nmid a_n$, $p^2 \nmid a_0$ while $p|a_j$ for $j = 0, 1, 2, \dots, n-1$ then $f(x)$ is irreducible. over $\mathbb{Q}$.*

Proof. If $f(x)$ were reducible over $\mathbb{Q}[x]$ then, by Theorem 21.3, there exists $g(x), h(x) \in \mathbb{Z}[x]$ with $\deg(g), \deg(h) \geq 1$ such that $f(x) = g(x)h(x)$. Let

$$g(x) = \sum_{k=0}^{r} g_k x^k \text{ and } h(x) = \sum_{k=0}^{s} h_k x^k$$

where $r + s = n$.

In order to absorb the general proof better let us first suppose that $r = s = 1$ so that

$$f(x) = (g_0 + g_1 x)(h_0 + h_1 x) = g_0 h_0 + (g_0 h_1 + g_1 h_0) x + g_1 h_1 x^2.$$

Since $p|(g_0 h_0)$ and $p^2 \nmid (g_0 h_0)$, we may assume that $p|g_0$ but $p \nmid h_0$. Combining this with the assumption that $p|(g_0 h_1 + g_1 h_0)$ implies that $p|g_1$. But this then implies $p|(g_1 h_1)$ which contradicts the assumption that p does not divide the leading order coefficient of $f(x)$.

With this as a warm-up we go to the general case. Notice that $a_0 = g_0 h_0$ and $a_n = g_r h_s$. Since $p|a_0$ while $p^2 \nmid a_0$ it follows that p divides exactly one of g_0 or h_0 but not both. Suppose that $p|g_0$ and $p \nmid h_0$. Also notice that $p \nmid a_n$ implies $p \nmid g_r$ so there is a first $0 < t \leq r < n$ such that $p \nmid g_t$. Since,

$$a_t = g_0 h_t + g_1 h_{t-1} + \dots + g_{t-1} h_1 + g_t h_0,$$

and $p|a_t$, it follows that $p|g_t h_0$ which then implies $p|h_0$ which is a contradiction. ■

For a second proof of this theorem, see Theorem 24.8 below.

Example 22.7. The polynomial, $x^5 + 9x^4 + 12x^2 + 6$ is irreducible over $\mathbb{Q}[x]$ since $p = 3$ divides all of the coefficients except the leading order one and $3^2 \nmid 6$.

For $n \in \mathbb{N}$, let

$$\Phi_n(x) = 1 + x + x^2 + \dots + x^{n-1}.$$

This geometric series may be summed in the usual way,

$$x\Phi_n(x) - \Phi_n(x) = x^n - 1 \implies \Phi_n(x) = \frac{x^n - 1}{x - 1}.$$

Corollary 22.8. *For any prime, $p \in \mathbb{N}$, $\Phi_p(x)$ is irreducible over $\mathbb{Q}$.*

Proof. Observe that

$$(x+1)^p - 1 = \sum_{k=0}^{p} \binom{p}{k} x^k - 1 = \sum_{k=1}^{p} \binom{p}{k} x^k,$$

$(x+1) - 1 = x$, and hence

$$\begin{aligned} \Phi_p(x+1) &= \frac{(x+1)^p - 1}{(x+1) - 1} = \frac{\sum_{k=1}^{p} \binom{p}{k} x^k}{x} = \sum_{k=1}^{p} \binom{p}{k} x^{k-1} \\ &= x^{p-1} + \binom{p}{p-1} x^{p-2} + \binom{p}{p-2} x^{p-3} + \dots \binom{p}{2} x + \binom{p}{1}. \end{aligned}$$

For $1 \leq k < p$, since there is no factor of p in $k!$ or $(p-k)!$, it follows that

$$\binom{p}{p-k} = \frac{p(p-1)!}{k!(p-k)!} = p \cdot \frac{(p-1)!}{k!(p-k)!}$$

is an integer which is divisible by p. Thus $p|\binom{p}{p-k}$ for $1 \leq k \leq p-1$ while $p^2 \nmid \binom{p}{1} = p$ and hence the Eisenstein's Criterion, $\Phi_p(x+1)$ is irreducible. This then is easily seen to imply $\Phi_p(x)$ is irreducible. ■

22.2 Summary of irreducibility tests

Theorem 22.9 (Reducibility tests). *Let F be a field and $f(x) \in F[x]$. If* $\deg f(x) = 1$, *then $f(x)$ is irreducible so suppose that* $\deg f(x) \geq 2$. *Then;*

$$\begin{pmatrix} f(x) \ \text{ has a root} \\ \text{in } F \end{pmatrix} \Longrightarrow \begin{pmatrix} f(x) \ \text{ is reducible} \\ \text{over } F \end{pmatrix}.$$

Moreover if $\deg f(x) = 2$ or $\deg f(x) = 3$, then

$$\begin{pmatrix} f(x) \ \text{ has a root} \\ \text{in } F \end{pmatrix} \Longleftrightarrow \begin{pmatrix} f(x) \ \text{ is reducible} \\ \text{over } F \end{pmatrix}.$$

Theorem 22.10 (Finding roots). *Suppose that* $f(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{Z}[x]$ *with* $a_n \neq 0 \neq a_0$. *Then the rational roots of $f(x)$ are contained in*

$$\left\{ \frac{r}{s} : r \in \mathbb{N},\ s \in \mathbb{Z}_+ \ni \ s|a_n \ \textit{and}\ r|a_0 \right\}.$$

Theorem 22.11 ($\mathbb{Q}$ − Irreducibility tests). *Suppose* $f(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{Z}[x]$ *with* $n := \deg f(x) \geq 2$. *Then;*

$$\begin{tikzcd}
 & \begin{pmatrix} \deg f(x) = 1 \ \textit{or} \\ \deg f(x) = 2 \ \textit{or}\ 3 \\ \&\ f(x) \ \textit{has no} \\ \textit{roots in}\ \mathbb{Q} \end{pmatrix} \arrow[d, Rightarrow] & \\
\begin{matrix} \text{mod}\,\mathbf{p}\ \boldsymbol{\textit{test}} \\ \begin{pmatrix} \exists \ \textit{a prime}\ p \ \ni \\ \deg f_p(x) = \deg f(x) \\ \&\ f_p(x) \ \textit{is irreducible} \\ \textit{over}\ \mathbb{Z}_p \end{pmatrix} \end{matrix} \arrow[r, Rightarrow] & \begin{pmatrix} f(x) \ \textit{is irreducible} \\ \textit{over}\ \mathbb{Q} \end{pmatrix} & \begin{matrix} \boldsymbol{\textit{Eisenstein's}} \\ \boldsymbol{\textit{Criterion}} \\ \begin{pmatrix} \exists \ \textit{a prime}\ p \\ \textit{such that} \\ p|a_k \ \textit{for}\ k < n \\ p \nmid a_n \ \&\ p^2 \nmid a_0 \end{pmatrix} \end{matrix} . \arrow[l, Rightarrow] \\
 & \begin{pmatrix} f(x) \ \textit{is irreducible} \\ \textit{over}\ \mathbb{Z} \end{pmatrix} \arrow[u, Rightarrow] &
\end{tikzcd}$$

Moreover if $f(x)$ is primitive (i.e. $c(f) = 1$) then

$$\begin{pmatrix} f(x) \ \textit{is irreducible} \\ \textit{over}\ \mathbb{Q} \end{pmatrix} \Longleftrightarrow \begin{pmatrix} f(x) \ \textit{is irreducible} \\ \textit{over}\ \mathbb{Z} \end{pmatrix}.$$

23

Lecture 23

23.1 Irreducibles and Primes II

Definition 23.1. *Let D be an integral domain (i.e. a commutative ring with 1 which has no zero divisors, i.e. cancellation holds in D) and $a, b, c \in D$. Then;*

1. *a is* ***associated*** *to b if $a = b\,(\mathrm{mod})\,U(R)$, i.e. $a \in bU(R)$ or equivalently $b \in aU(R)$.*
2. *$a \in D$ is* ***irreducible*** *iff $a \neq 0$, $a \notin U(R)$ and whenever $a = bc$ then either $b \in U(R)$ or $c \in U(R)$. (So $a \in D$ is irreducible if it can not be factored in a non-trivial way.) Similarly we say that $0 \neq a \in D$ is reducible if it admits a factorization of the form $a = bc$ with $b, c \notin U(R)$.*
3. *$a \in D$ is* ***prime*** *if $0 \neq a \notin U(R)$ and if a satisfies; whenever $a|bc$ then either $a|b$ or $a|c$. (So $a \in D$ is prime if the conclusion of Euclid's lemma holds.)*

Example 23.2. Let $R = \mathbb{Z}$ so that $U(\mathbb{Z}) = \{\pm 1\}$ and the associates to m are $\{\pm m\}$. Now $0 \neq a \in \mathbb{Z}$ is irreducible iff whenever $a = bc$ we must have b or c is ± 1. (That is $|a|$ is prime in our old sense of the word.) It therefore follows by Euclid's lemma (also see Theorem 23.7)) that if $a|bc$ then $a|b$ or $a|c$ which means that a is prime in our new sense of being prime as well. As will see in Theorem 23.4 below, prime always implies irreducible. Thus we are new notion of prime gives $\{\pm 2, \pm 3, \pm 5, \pm 7, \dots\}$ for the primes in $\mathbb{Z}$.

Proposition 23.3 ($\langle$prime$\rangle$ is prime). *Let D be an integral domain and $a \in D$ with $0 \neq a \notin U(R)$. Then the following are equivalent;*

1. *a is prime.*
2. *$\langle a \rangle = Ra$ is a prime ideal in D.*
3. *$R/\langle a \rangle$ is an integral domain.*

Proof. See the homework. ■

Theorem 23.4 (Prime $\implies$ irreducible). *If D is an integral domain and $a \in D$ is prime then a is irreducible.*

Proof. Suppose that $a \in D$ is a prime. If $a = bc$ then $a|b$ or $a|c$. Say that $a|b$, i.e. $b = at$. Then $a = atc$ and since D is a domain we may cancel a form this equation to see that $1 = tc$. This shows that $c \in U(R)$ and so a only admits trivial factorizations, i.e. a is irreducible. ■

The next example shows that the converse of Theorem 23.4 is false in general, i.e. it is not always true that irreducibles are primes.

Example 23.5 (Example 1, p. 321). In this example we will show that

$$\alpha := 1 + \sqrt{-3} \in R := \mathbb{Z}\left[\sqrt{-3}\right]$$

is **irreducible but not prime** which we will now verify. First observe that R is a sub-ring of $\mathbb{C}$ which will simplify our computations. Given $x = a + b\sqrt{-3} \in R$, let $\bar{x} := a - b\sqrt{-3}$ and $N(x) := |x|^2 = x\bar{x} = a^2 + 3b^2 \in \mathbb{N}$. Since $|xy| = |x|\,|y|$ for all $x, y \in \mathbb{C}$ A direct calculation shows[1] that $N(xy) = N(x)\,N(y)$ for all $x, y \in R$.

- $U(R) = \{\pm 1\}$. If $xy = 1$, then $1 = N(1) = N(x)\,N(y)$ implies $N(x) = 1 = N(y)$. Conversely if $N(x) = 1$, then $x^{-1} = \bar{x}$ and therefore
$$U(R) = \{x \in R : N(x) = 1\} = \{\pm 1\}.$$
- α **is irreducible.** If $\alpha = xy$ with $x, y \in R$, then $4 = N(\alpha) = N(x)\,N(y)$. If neither x nor y is a unit, we must have $N(x) = 2 = N(y)$ which is impossible since $N(x) = a^2 + 3b^2$ which is never 2. Thus no such decomposition exists and α is irreducible.
- α **is not prime.** Since $\alpha\bar{\alpha} = N(\alpha) = 4$, $\alpha|4$ while $4 = 2 \cdot 2$. So if α were prime we would have $\alpha|2$, i.e. $2 = \alpha x$ for some $x = a + b\sqrt{-3}$. However if $2 = \alpha x$ then $2\bar{\alpha} = \alpha \cdot \bar{\alpha} x = N(\alpha)\,x = 4x$. Thus we must have $2x = \bar{\alpha}$, i.e.
$$2a + 2b\sqrt{-3} = 1 - \sqrt{-3}$$
which is impossible to do unless we take $a = 1/2$ and $b = -1/2$ neither of which are in R. Thus we have shown α does **not** divide 2 and so α is **not** prime.
- Similarly, one shows that 2 irreducible but not prime. To see that 2 is not prime, notice that $2|\alpha\bar{\alpha} = 4$ while $2 \nmid \alpha$. If 2 were reducible, then $2 = xy$ with $x, y \notin U(R)$. Since $4 = N(2) = N(x)\,N(y)$ and neither $N(x) = 1$ or $N(y) = 1$ we must have $N(x) = 2$ which is impossible.

[1] You should check this by direct calculation if this is not familiar to you.

The next theorem gathers in one place all of the results we know involving maximal and prime principle ideals and prime and irreducible elements of an integral domain R.

Theorem 23.6. *Let R be an integral domain and $\langle a\rangle = Ra$ where $a \in R$ with $a \neq 0$ and $a \notin U(R)$. Then;*

$$\begin{array}{c}
\langle a\rangle \text{ is maximal ideal} \iff R/\langle a\rangle \text{ is a field} \\
\Downarrow \ (1) \\
\langle a\rangle \text{ is a prime ideal} \iff R/\langle a\rangle \text{ is an integral domain} \\
\Updownarrow \ (2) \\
a \text{ is a prime element in } R \\
\Downarrow \ (3) \\
a \text{ is an irreducible element in } R.
\end{array}$$

If we further assume that R is a PID, then the converse of all of the above implications hold.[2]

Proof. The horizontal equivalences are contained in Theorem 14.3. The vertical implication (1) has been explained in Corollary 14.4. The vertical equivalence (2) is Proposition 23.3 and (3) is Theorem 23.4.

If R is now a PID, we know from Theorem 18.8 that a is irreducible implies $\langle a\rangle$ is maximal. Therefore all statements in the theorem are now equivalent. ■

If time permits we will consider some other interesting integral domains which are not PID's and for which some of the equivalences in Theorem 23.6 no longer hold.

Remark 23.7 (Prime $\iff$ irreducible in a PID). Here is an alternative proof that if D is a PID (principle ideal domain) and $a \in D$ is irreducible then a is prime. Suppose that $a \in D$ is irreducible and $a|bc$ for some $b, c \in D$. Let

$$I := aD + bD := \{ax + by : x, y \in D\}.$$

Notice that I is an ideal of D and hence $I = \langle d\rangle$ for some $d \in D$. Since $a \in I$ we have $a = dr$ for some $r \in D$. As a is irreducible, either $d \in U(R)$ or $a \in U(R)$. If $d \in U(R)$, then $I = \langle d\rangle = R$ and in particular $1 \in I$. Thus there exists $x, y \in D$ such that $ax + by = 1$. Therefore,

$$axc + ybc = c$$

from which it follows that $a|c$. On the other hand of $r \in U(R)$, then $I = \langle d\rangle = \langle a\rangle$. Since $b \in I$ it follows that $b = at$ for some $t \in D$, i.e. $a|b$. So we have shown that $a|bc$ implies $a|b$ or $a|c$, i.e. a is prime.

Example 23.8. We have $x \in \mathbb{Z}[x]$ is prime and hence irreducible. Nevertheless, $\langle x\rangle$ is not a maximal ideal as we have seen above. For example, $\langle x\rangle \underset{\neq}{\subsetneq} \{p(x) : p(0) \text{ is even}\} \underset{\neq}{\subsetneq} \mathbb{Z}[x]$. Thus we can not drop the assumption that D is a PID in Theorem 23.4. Nevertheless, we will see below in Theorem 24.1 that $f(x) \in \mathbb{Z}[x]$ is prime iff it is irreducible in $\mathbb{Z}[x]$.

Example 23.9. In $\mathbb{Z}_5$, $x^2 + 1$ has two zeros, namely $x = 2$ and $x = 3$. Therefore, $x^2 + 1$ factors as $(x - 2)(x - 3)$ from which it follows that $\langle x^2 + 1\rangle$ is not a maximal ($\iff$ prime) ideal in $\mathbb{Z}_5[x]$. Therefore it follows that $\mathbb{Z}_5[x] / \langle x^2 + 1\rangle \cong \mathbb{Z}_5[i]$ is not a field and not even an integral domain. Explicitly we have $(i - 2)(i - 3) = 6 - 1 - 5i = 0$ in $\mathbb{Z}_5[i]$.

Remark 23.10. On the other hand, it is easy to verify that $x^2 + 1$ has no zeros over $\mathbb{Z}_3$ and therefore $x^2 + 1$ is irreducible over $\mathbb{Z}_3$. From this we may conclude that $\mathbb{Z}_3[i] \cong \mathbb{Z}_3[x] / \langle x^2 + 1\rangle$ is a field. This was already proved with some effort in Example 5.3.

[2] More succinctly, if R is a PID and $a \in R^\times \setminus U(R)$ is irreducible, then $\langle a\rangle$ is a maximal ideal in R.

24

Lecture 24

It is possible for all irreducibles to be prime even if the ring is not a PID as the next theorem shows.

Theorem 24.1 (Irreducible $\iff$ prime in $\mathbb{Z}[x]$). *If $f(x) \in \mathbb{Z}[x]$ is irreducible then $f(x)$ is prime in $\mathbb{Z}[x]$.*

Proof. In the proof to follow we suppose that $g(x), h(x) \in \mathbb{Z}[x]$ and that $f(x) | g(x) h(x)$, i.e. there exists $k(x) \in \mathbb{Z}[x]$ such that

$$g(x) h(x) = k(x) f(x). \tag{24.1}$$

Case 1. Suppose that $\deg(f) = 0$, i.e. $f(x)$ is constant. Since $f(x)$ is irreducible we must have $f(x) = p$ where p is a prime in $\mathbb{Z}$ which we assume to be positive for simplicity. Applying the content function, $c(\cdot)$, to Eq. (24.1) gives

$$c(g) c(h) = c(k) c(f) = c(k) \cdot p.$$

Therefore $p|c(g) c(h)$ which implies $p|c(g)$ or $p|c(h)$, i.e. $p|g(x)$ or $p|h(x)$.

Case 2. Suppose that $\deg(f) \geq 1$. In this case $c(f) = 1$ for otherwise

$$f(x) = c(f) \left[\frac{1}{c(f)} f(x)\right]$$

would be a non-trivial factorization of $f(x)$ over $\mathbb{Z}[x]$. Since $f(x)$ is primitive and irreducible over $\mathbb{Z}[x]$ it is irreducible over $\mathbb{Q}[x]$ by Corollary 21.5. Therefore $f(x)$ is prime in $\mathbb{Q}[x]$ and it follows that $f(x)|g(x)$ or $f(x)|h(x)$ in $\mathbb{Q}[x]$. But by Corollary 20.13 this implies that $f(x)|g(x)$ or $f(x)|h(x)$ in $\mathbb{Z}[x]$, i.e. $f(x)$ is prime in $\mathbb{Z}[x]$. ■

24.1 Unique Factorization Domains

Definition 24.2 (UFD). *A* ***unique factorization domain*** *(UFD for short) is an integral domain, R, such that*

1. *Every non-zero element of R which is not a unit may be written as a product of irreducibles in R.*
2. *The factorization into irreducibles is unique up to associates and the order in which the factors appear. To be more precise, if $a_1, \dots, a_n$ and $b_1, \dots, b_m$ are irreducibles in R such that $a_1 \dots a_n = b_1 \dots b_m$ then $m = n$ and there exists a permutation, $\sigma \in S_n$, such that a_i and $b_{\sigma(i)}$ are associates for each i.*

Example 24.3. The fundamental theorem of arithmetic asserts that $\mathbb{Z}$ is a unique factorization domain.

Example 24.4 ($\mathbb{Z}[\sqrt{-3}]$ is not a UFD). Example 23.5 easily allows us to see that $R := \mathbb{Z}\left[\sqrt{-3}\right] \subset \mathbb{C}$ is **not** a unique factorization domain. (see Definition 24.2 below). We have see that $4 \in R$ may be factored as $4 = 2 \cdot 2$ and also as $4 = \alpha \cdot \bar{\alpha}$ where

$$\alpha := 1 + \sqrt{-3} \text{ and } \bar{\alpha} := 1 - \sqrt{-3}.$$

Since the units of R are $\{\pm 1\}$ it is clear neither α nor $\bar{\alpha}$ is associated to 2. (See Chapter 18 of the Gallian [2] and in particular the comments on p. 333 as to when $\mathbb{Z}\left[\sqrt{d}\right]$ is a UFD.)

Lemma 24.5. *Suppose that R is an integral domain, $a \in R$ is prime, and $b_1, \dots, b_n \in R$ are irreducibles. If $a|(b_1 \dots b_n)$ then a is associated to b_i for some i.*

Proof. Since a is prime, $a|b_i$ for some i. As b_i is irreducible this can only happen if $b_i = au$ where $u \in U(R)$, i.e. a is associated to b_i. ■

Proposition 24.6. *Suppose that R is an integral domain, $1 \leq m \leq n$, $\{a_i\}_{i=1}^m$ are primes in R and $\{b_j\}_{j=1}^n$ are irreducibles in R. If*

$$a_1 \dots a_m = b_1 \dots b_n,$$

then $m = n$ and after a possible reordering of the $\{b_i\}_{i=1}^n$, we have a_i and b_i are associates for each i.

Proof. Before proving the general case let first do two special cases as a warm up.

1) Suppose that $m = 2$ and $n = 3$ so that $a_1 a_2 = b_1 b_2 b_3$. By Lemma 24.5, $b_i = u_1 a_1$ for some i and $u_1 \in U(R)$. By relabeling the b_i if necessary we may

assume that $i = 1$ and we now have $a_1 a_2 = a_1 u_1 b_2 b_3$ and so by cancellation, $a_2 = (u_1 b_2) b_3$. Another application of Lemma 24.5 shows a_2 is associated to $u_1 b_2$ (hence b_2) or b_3. Again we may relabel the b_i if necessary and suppose that a_2 is associated to b_2, i.e. $b_2 = u_2 a_2$ for some $u_2 \in U(R)$. Thus we may conclude that $a_2 = a_2 u_1 u_2 b_3$ which implies $1 = u_1 u_2 b_3$. This shows that b_3 must be a unit which contradicts the assumption that b_3 was prime and in particular not a unit. Thus this case can not happen.

2) Suppose that $m = n = 2$ so that $a_1 a_2 = b_1 b_2$. Working as in case 1), we easily show, after relabeling the b_i if necessary, that a_1 is associated to b_1 and a_2 is associated to b_2.

The formal proof goes by induction on n. When $n = m = 1$ there is nothing to prove. Now suppose that $n \geq 2$ and the result holds for lower n. Then working as above, after relabeling the b's is necessary we may assume that $b_1 = a_1 u_1$ for some $u_1 \in U(R)$. If $m = 1$, this would imply that $1 = b_2 \dots b_n$ showing that all the remaining b_i are units which is impossible. Therefore we must have $m \geq 2$ and we now have

$$\left(u_1^{-1} a_2\right) a_3 \dots a_m = b_2 b_3 \dots b_n.$$

Since $u_1^{-1} a_2$ is still prime, it follows by the induction hypothesis that $m = n$ and that after relabeling the $\{b_i\}_{i=2}^n$ if necessary that

$$u_1^{-1} a_2 = u_2 b_2, \ a_3 = u_3 b_3, \dots, a_n = u_n b_n$$

for some $u_i \in U(R)$. This completes the proof since $a_2 = u_1 u_2 b_2$ and $u_1 u_2 \in U(R)$. ■

Because of this proposition we see that the uniqueness of factorizations into irreducibles will hold for an integral domain R when R has the property that irreducibles and primes are the same, i.e. irreducible implies prime. This is not always the case as was shown in Example **??**. However we have seen that irreducibles and primes are the same in principle ideal domains (Theorem 23.6) like $F[x]$ and in $R = \mathbb{Z}[x]$, see Theorem 24.1.

Theorem 24.7 ($F[x]$ is a UFD). *If F is a field, then $F[x]$ is a unique factorization domain.*

Proof. Let $f(x) \in F[x]$ with $\deg(f) \geq 1$. If $f(x)$ is not irreducible, then $f(x) = g(x) h(x)$ for some $g(x), h(x) \in F[x]$ with $\deg g \geq 1$ and $\deg h \geq 1$. Since $\deg g + \deg h = \deg f$ we must have $\deg g < \deg f$ and $\deg h < \deg f$. By further decomposing $g(x)$ and $h(x)$ is possible we must eventually arrive at a decomposition of $f(x)$ into irreducibles which gives the desired existence. (You may add the formal the induction proof if you wish.) The uniqueness of this factorization follows from Proposition 24.6 and the fact that primes and irreducibles are the same in a PID like $F[x]$, see Theorem 23.7. ■

As an application of this theorem let us give another proof of Eisenstein's Criterion – see Theorem 22.6. For the readers convenience we repeat the statement here.

Theorem 24.8 (Eisenstein's Criterion). *Let $f(x) := \sum_{k=0}^n a_k x^k \in \mathbb{Z}[x]$. If there is a prime p such that $p \nmid a_n$, $p^2 \nmid a_0$ while $p | a_j$ for $j = 0, 1, 2, \dots, n-1$ then $f(x)$ is irreducible. over $\mathbb{Q}$.*

Proof. If $f(x)$ were reducible over $\mathbb{Q}[x]$ then, by Theorem 21.3, it would be reducible over $\mathbb{Z}[x]$ and in fact there would exists $g(x), h(x) \in \mathbb{Z}[x]$ with $\deg(g), \deg(h) \geq 1$ such that $f(x) = g(x) h(x)$. Letting $\bar{f}$, $\bar{g}$ and $\bar{h}$ be the reductions of f, g, and h mod p we find, $\bar{a}_n x^n = \bar{g}(x) \bar{h}(x)$ wherein we have use $\bar{a}_n := a_n \bmod p \neq 0$ in $\mathbb{Z}_p$. Since $\mathbb{Z}_p$ is a field, $\mathbb{Z}_p[x]$ is a UFD and because $\deg \bar{g}$ and $\deg \bar{h}$ are both less than n we must have x divides both $\bar{g}(x)$ and $\bar{h}(x)$. (In fact there must exist $b, c \in \mathbb{Z}_p^\times$ such that $\bar{g}(x) = b x^{\deg g}$ and $\bar{h}(x) = c x^{\deg h}$.) This then implies that $p | g(0)$ and $p | h(0)$ and therefore $p^2 | a_0$ as $a_0 = g(0) h(0)$. However this contradicts the assumption that $p^2 \nmid a_0$. ■

Theorem 24.9 ($\mathbb{Z}[x]$ is a UFD). *$\mathbb{Z}[x]$ is a unique factorization domain.*

Proof. Existence. Let $f(x) \in \mathbb{Z}[x]$ and set $f_0(x) := \frac{1}{c(f)} f_0(x)$ so that $f(x) = c(f) f_0(x)$ where $f_0(x)$ is primitive. Since we may factor $c(f)$ into prime integers which are irreducible in $\mathbb{Z}[x]$ it suffices to factor $f_0(x)$ into irreducibles. So from now on we assume that $f(x)$ is primitive.

If $f(x)$ is not irreducible, the $f(x) = h(x) g(x)$ with neither $h(x)$ or $g(x)$ being invertible. Since $c(h) c(g) = c(f) = 1$ we know that both h and g are primitive and therefore $\deg h \geq 1$ and $\deg g \geq 1$. Since $\deg f = \deg h + \deg g$, we must have $\deg h < \deg f$ and $\deg g < \deg f$. Thus the existence of a factorization of $f(x)$ when $f(x)$ is primitive goes just as in the proof of Theorem 24.7.

Uniqueness. The uniqueness of the factorization follows by combining Proposition 24.6 with Theorem 24.1. ■

Fact 24.10 (Facts about UFD's) *Here are two more facts about unique factorization domains which we will not prove in class.*

1. *Every PID is a UFD. This is proved in Gallian [2, p. 327 – 329] and is fairly similar to the proof of Theorem 24.7.*
2. *If R a UFD then $R[x]$ is a UFD. Our proof that $\mathbb{Z}[x]$ is a UFD is rather similar to the general case – see Knapp [3, pgs. 384-396] for details. (This proof would make use of the field of fractions, see Subsection 24.2.4 below if you are interested.)*

24.2 Extra Topics (not covered in class)

24.2.1 Greatest Common Divisors (not covered in class)

Definition 24.11. *Let R be an integral domain. We say that $c \in R$ is a common divisor of $a_1, \dots, a_n \in R$ if $c|a_i$ for all i. We say that $d \in R$ is a greatest common divisor of $a_1, \dots, a_n \in R$ if d is a divisor and if $c|d$ for any other divisor of $a_1, \dots, a_n \in R$.*

If c and d are two greatest common divisors of $a_1, \dots, a_n \in R$, then $c = ud$ and $d = vc$ for some $u, v \in R$. Therefore, $c = uvc$ and since R is an integral domain, $uv = 1$ by cancellation. Thus c and d are associates and the gcd of $a_1, \dots, a_n$ is well defined modulo $U(R)$.

Proposition 24.12. *If R is a PID, then every list of elements, $a_1, \dots, a_n \in R^\times$, has a greatest common divisor, d. Moreover there exists $r_1, \dots, r_n \in R$ such that*

$$d = r_1 a_1 + \dots + r_i a_i + \dots + r_n a_n. \tag{24.2}$$

Proof. Let $J = \langle a_1, \dots, a_n\rangle = \langle a_1\rangle + \dots + \langle a_n\rangle$ be the ideal of R generated by $(a_1, \dots, a_n)$ and choose $c \in R$ such that $J = \langle c\rangle$. Then by definition of J, c may be expressed as in Eq. (24.2) and $a_i \in \langle c\rangle$ for all i. The latter condition asserts that $c|a_i$ for all i so that c is a common divisor and it follows from Eq. (24.2) that in fact c is a greatest common divisor. ■

Definition 24.13. *As in the case $R = \mathbb{Z}$ we will say that elements, $a_1, \dots, a_n$ in a PID, R, are relatively prime if the greatest common divisors are units. Alternatively, this is equivalent to saying $1 \in \gcd(a_1, \dots, a_n)$.*

Here is the analogue of Euclid's Lemma in this more general context.

Lemma 24.14 (Euclid's Lemma II). *Let R be a PID and $a, b, c \in R$ with a, b being relatively prime. If $a|(bc)$ then $a|c$.*

Proof. From Proposition 24.12 we know that $1 = sa + tb$ for some $s, t \in R$. Therefore $c = sac + tbc$ from which it follows that $a|c$. ■

Remark 24.15. If F is a field and $p(x), q(x) \in F[x]$ are two non-zero polynomials, one may find $\gcd(p, q)$ using the division algorithm just as we did for the integers, $\mathbb{Z}$. The point is to repeatedly make use of the fact that

$$\gcd(p, q) = \gcd(q, q \bmod p)$$

where $q \bmod p$ denotes the remainder from dividing p into q.where $q \bmod p$ denotes the remainder from dividing p into q.

The next result has already been proved in Theorem 18.8. Nevertheless it is instructive to give another proof.

Corollary 24.16. *Suppose that F is a field and $p(x) \in F[x]$ is irreducible. Then $F[x] / \langle p(x)\rangle$ is a field.*

Proof. If $f(x) \in F[x]$ with $\deg f(x) < \deg p(x)$, then $f(x)$ and $p(x)$ are relatively prime since the only possible non-trivial common factor would be $p(x)$ which is impossible since $\deg f(x) < \deg p(x)$. Therefore by Proposition 24.12, there exists $s(x), t(x) \in F[x]$ such that

$$s(x) f(x) + t(x) p(x) = 1.$$

Since $[p(x)] = 0$ it follows that $[s(x)][f(x)] = 1$ and we have shown $[f(x)]^{-1} = [s(x)]$ exists. Thus every non-zero element of $F[x] / \langle p(x)\rangle$ is invertible and hence $F[x] / \langle p(x)\rangle$ is a field. ■

Remark 24.17. It is interesting to see how to compute $[f(x)]^{-1}$ in the context of Corollary 24.16 explicitly. This can be done using the same technique as for finding inverses in $U_n = U(\mathbb{Z}_n)$. That is we repeatedly use the division algorithm as follows;

$$\begin{aligned} p(x) &= k_0(x) f(x) + f_1(x) \\ f(x) &= k_1(x) f_1(x) + f_2(x) \\ f_1(x) &= k_2(x) f_2(x) + f_3(x) \\ &\vdots \\ f_{l-1}(x) &= k_l(x) f_l(x) + f_{l+1} \end{aligned}$$

where $f_{l+1} \in F^\times$. We then work backwards from this set of equation to solve for f_{l+1}. Rather than explain this in general, let me show how this works in an example.

Example 24.18. Let $I := \langle x^3 + 2x + 1\rangle \subset \mathbb{Q}[x]$. Notice that $p(x) := x^3 + 2x + 1$ is irreducible by the mod 3 test. We now wish to compute $[x]^{-1}$ and $[x^2]^{-1}$. The first case is rather simple since,

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

and therefore

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

so that

$$1 = [p(x)] - [x][x^2 + 2] = -[x][x^2 + 2]$$

which shows that $[x]^{-1} = -\left[x^2+2\right].$

Now to compute $\left[x^2\right]^{-1}$. By the division algorithm we have

$$p(x) = x^3 + 2x + 1 = x \cdot x^2 + 2x + 1$$

and

$$\begin{array}{r|l} & \frac{1}{2}x - \ \ \frac{1}{4} \\ \hline 2x+1 & x^2 \\ & x^2 + \ \frac{1}{2}x \\ \hline & \quad -\frac{1}{2}x \\ & \quad -\frac{1}{2}x - \frac{1}{4} \\ \hline & \qquad\quad \frac{1}{4} \end{array}$$

so that

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

Working backwards this implies,

$$\begin{aligned} \frac{1}{4} &= x^2 - \left(\frac{1}{2}x - \frac{1}{4}\right)(2x+1) \\ &= x^2 - \left(\frac{1}{2}x - \frac{1}{4}\right)\left(p(x) - x \cdot x^2\right) \\ &= x^2\left(1 + \left(\frac{1}{2}x - \frac{1}{4}\right)x\right) - \left(\frac{1}{2}x - \frac{1}{4}\right)p(x). \end{aligned}$$

Therefore,

$$\begin{aligned} [1] = 4 \cdot \left[\frac{1}{4}\right] &= 4 \cdot \left[x^2\left(1 + \left(\frac{1}{2}x - \frac{1}{4}\right)x\right)\right] \\ &= \left[x^2\right]\left[4 + (2x-1)x\right] = \left[x^2\right]\left[2x^2 - x + 4\right] \end{aligned}$$

which shows that

$$\left[x^2\right]^{-1} = \left[2x^2 - x + 4\right].$$

As a check we should have

$$\left[x^2\right]^{-1} = \left([x]^{-1}\right)^2 = \left[\left(x^2+2\right)^2\right] = \left[x^4 + 4x^2 + 4\right]. \tag{24.3}$$

However

$$\begin{array}{r|l} & x \\ \hline x^3+2x+1 & x^4 + 0x^3 + 4x^2 + 0x + 4 \\ & x^4 + \qquad + 2x^2 + \ x \\ \hline & \qquad\qquad 2x^2 - \ x \ + 4 \end{array}$$

which is to say,

$$x^4 + 4x^2 + 4 = x \cdot p(x) + 2x^2 - x + 4$$

and therefore,

$$\left[x^2\right]^{-1} = \left[x \cdot p(x) + 2x^2 - x + 4\right] = \left[2x^2 - x + 4\right]$$

as before. Alternatively, from Eq. (24.3) we may simply use

$$\left[x^4\right] = \left[x \cdot x^3\right] = \left[x \cdot (-2x-1)\right] = -\left[2x^2 + x\right]$$

to discover,

$$\begin{aligned} \left([x]^{-1}\right)^2 &= \left[x^4 + 4x^2 + 4\right] = \left[-\left(2x^2+x\right) + 4x^2 + 4\right] \\ &= \left[2x^2 - x + 4\right]. \end{aligned}$$

Remark 24.19. If $p(x) \in F[x]$ is not irreducible, then we may still compute $U(F[x]/\langle p(x)\rangle)$ as we did for the rings $\mathbb{Z}_m$ where m was not prime. The result is,

$$U(F[x]/\langle p(x)\rangle) = \left\{[f(x)] : \deg f(x) < \deg p(x) \text{ and } \gcd(f(x), p(x)) \in F^\times\right\}.$$

The division algorithm gives us again a way to compute $[f(x)]^{-1}$ when $\gcd(f(x), p(x)) \in F^\times$.

24.2.2 Partial Fractions (not covered in class)

In this section let F be a fixed field. We are going verify in this sections standard facts about partial fraction decompositions which you were probably first introduced to when studying integral calculus.

Theorem 24.20 (Partial fractions). *Let $p(x)$, $q(x)$, and $f(x)$ be in $F[x]$ and assume that p and q are relatively prime and that $\deg f(x) < \deg p(x) + \deg q(x) = \deg(p(x)q(x))$. Then there exists unique polynomials, $a(x)$ and $b(x)$ with $\deg a(x) < \deg p(x)$ and $\deg b(x) < \deg q(x)$ such that*

$$\frac{f(x)}{p(x)q(x)} = \frac{a(x)}{p(x)} + \frac{b(x)}{q(x)}. \tag{24.4}$$

Proof. Choose $t(x)$ and $s(x) \in F[x]$ such that $1 = t(x)p(x) + s(x)q(x)$. Then

$$\begin{aligned} \frac{f(x)}{p(x)q(x)} &= \frac{f(x)}{p(x)q(x)}\left(t(x)p(x) + s(x)q(x)\right) \\ &= \frac{s(x)f(x)}{p(x)} + \frac{t(x)f(x)}{q(x)}. \end{aligned}$$

Let $a(x) := s(x) f(x) \bmod p(x)$ and $b(x) := t(x) f(x) \bmod q(x)$ in which case we have,

$$\frac{f(x)}{p(x)\, q(x)} = \frac{a(x)}{p(x)} + \frac{b(x)}{q(x)} + k(x)$$

for some $k(x) \in F[x]$. Multiplying this last equation by $p(x)\, q(x)$ then implies,

$$f(x) = a(x)\, q(x) + b(x)\, p(x) + k(x)\, q(x)\, p(x).$$

So on one hand,

$$\begin{aligned}\deg k(x) + \deg(q(x)\, p(x)) &= \deg(k(x)\, q(x)\, p(x)) \\ &= \deg(f(x) - a(x)\, q(x) - b(x)\, p(x)) < \deg(p(x)\, q(x))\end{aligned}$$

from which we conclude that $\deg k(x) = -\infty$, i.e. $k(x) = 0$. So we have proved the existence of the decomposition in Eq. (24.4).

To prove uniqueness, it suffices to show that if $f(x) = 0$ then $a(x) = 0 = b(x)$ in Eq. (24.4) or equivalently that

$$a(x)\, q(x) + b(x)\, p(x) = 0 \implies a(x) = 0 = b(x).$$

However, since $a(x)\, q(x) = -b(x)\, p(x)$ it follows that $q(x)\, | b(x)\, p(x)$ and as $q(x)$ and $p(x)$ are relatively prime, we must have $q(x)\, | b(x)$. but $\deg b(x) < \deg q(x)$ and therefore this is only possible if $b(x) = 0$ which then implies $a(x) = 0$ as well. ■

Proposition 24.21. *Suppose that $p(x) \in F[x]$ is a polynomial with $\deg p(x) \geq 1$, $n \in \mathbb{Z}_+$, and $f(x) \in F[x]$ has $\deg f(x) < \deg p(x)^n = n \deg p(x)$. Then there exists unique polynomials, $a_i(x)$ with $\deg a_i(x) < \deg p(x)$ such that*

$$\frac{f(x)}{p(x)^n} = \sum_{i=1}^{n} \frac{a_i(x)}{p(x)^i}. \tag{24.5}$$

Proof. By the division algorithm, there exists $k(x), r(x) \in \mathbb{R}[x]$ such that $f(x) = k(x)\, p(x) + a_n(x)$ where $\deg a_n(x) < \deg p(x)$. Moreover we must have $\deg k(x) < \deg p(x)^{n-1}$ since otherwise we would violate the assumption that $\deg f(x) \geq \deg p(x)^n$. Therefore,

$$\frac{f(x)}{p(x)^n} = \frac{k(x)\, p(x) + a_n(x)}{p(x)^n} = \frac{k(x)}{p(x)^{n-1}} + \frac{a_n(x)}{p(x)^n}$$

and iterating this procedure (i.e. by induction) we easily prove the existence of the decomposition in Eq. (24.5).

For uniqueness, it suffices to show that if $\sum_{i=1}^n \frac{a_i(x)}{p(x)^i} = 0$ with $\deg a_i(x) < \deg p(x)$, then $a_i(x) = 0$ for all i. Equivalently we must show if

$$0 = p(x)^n \sum_{i=1}^{n} \frac{a_i(x)}{p(x)^i} = \sum_{i=1}^{n} a_i(x)\, p(x)^{n-i} \implies a_i(x) = 0 \ \forall\, i.$$

To see this observe that

$$0 = \sum_{i=1}^{n} a_i(x)\, p(x)^{n-i} = a_1(x)\, p(x)^{n-1} + r(x)$$

where

$$r(x) = \sum_{i=2}^{n} a_i(x)\, p(x)^{n-i} \text{ and } \deg r(x) < \deg p(x)^{n-1}.$$

Therefore, by the division algorithm (or directly by degree arguments) we must have that $a_1(x) = 0$ and $r(x) = 0$. Repeating this procedure (induction again) then shows that all of the $a_i(x)$ must be zero. ■

Corollary 24.22 (Partial fractions). *Let F be a field and $\{p_i(x)\}_{i=1}^m$ be a collection of distinct monic irreducible polynomials in $F[x]$. Further let $k_i \geq 1$ for each i and $f(x) \in F[x]$ with $\deg f(x) < \sum_{i=1}^m k_i \deg p_i(x)$. Then there exists $a_{ij}(x) \in F[x]$ for $1 \leq i \leq m$ and $1 \leq j \leq k_i$ such that $\deg a_{ij}(x) < \deg p_i(x)$ for all i and j and*

$$\frac{f(x)}{\prod_{i=1}^m p_i(x)^{k_i}} = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \frac{a_{ij}(x)}{p_i(x)^j}.$$

Proof. Repeatedly make use of Theorem 24.20 and Proposition 24.21. ■

24.2.3 Factorizing Polynomials in finite time (not covered in class)

The goal of this section is to show that it is possible to factor $f(x) \in \mathbb{Z}[x]$ with a finite number of operations. The point is Mignotte's bound in Corollary 24.27 which states that if $h(x) \in \mathbb{Z}[x]$ is a factor of $f(x)$, then there is an a priori bound on the size of all of the coefficients of $h(x)$ in terms of a bound on the coefficients of $f(x)$ and the degrees of the polynomials, $f(x)$ and $h(x)$. Thus when looking for factorizations of $f(x)$ we may reduce the question to a finite number of possibilities to try out. The material presented below is taken directly from [4, p. 162 - 164].

Notation 24.23 *Given $f(x) = \sum_{k=0}^n f_k x^k \in \mathbb{C}[x]$, let*

$$\|f\|_\infty := \max_k |f_k|, \quad \|f\|_1 := \sum_k |f_k|, \text{ and } \|f\|_2 := \left(\sum_k |f_k|^2\right)^{1/2}.$$

Further we let

$$M(f) := |f_n| \prod_{i=1}^n \max(1, |z_i|)$$

where $\{z_i\}_{i=1}^n$ *are the roots of* f *counted with multiplicities, i.e.*

$$f(x) = f_n \prod_{i=1}^n (x - z_i). \tag{24.6}$$

Lemma 24.24. *If* $z \in \mathbb{C}$, *and* $f(x) \in \mathbb{C}[x]$, *then*

$$\|(x-z) f(x)\|_2 = \|(\bar{z}x - 1) f(x)\|.$$

Proof. Letting $f_{-1} = 0$ and $f_k = 0$ for $k > \deg f$, we have

$$\begin{aligned}\|(x-z) f(x)\|_2^2 &= \sum_{k=0}^\infty |f_{k-1} - z f_k|^2 \\ &= \sum_{k=0}^\infty \left(|f_{k-1}|^2 + |z f_k|^2 - 2\operatorname{Re}\left(f_{k-1}\bar{f}_k \bar{z}\right)\right) \\ &= \left(1 + |z|^2\right)\|f(x)\|_2^2 - 2\sum_{k=0}^\infty \operatorname{Re}\left(f_{k-1}\bar{f}_k\bar{z}\right).\end{aligned}$$

On the other hand,

$$\begin{aligned}\|(\bar{z}x - 1) f(x)\|_2^2 &= \sum_{k=0}^\infty |\bar{z} f_{k-1} - f_k|^2 \\ &= \sum_{k=0}^\infty \left(|\bar{z} f_{k-1}|^2 + |f_k|^2 - 2\operatorname{Re}\left(\bar{z} f_{k-1}\bar{f}_k\right)\right) \\ &= \left(1 + |z|^2\right)\|f(x)\|_2^2 - 2\sum_{k=0}^\infty \operatorname{Re}\left(f_{k-1}\bar{f}_k\bar{z}\right).\end{aligned}$$

Comparing these last two equations completes the proof. ■

Theorem 24.25 (Landau's inequality). *For all* $f(x) \in \mathbb{C}[x]$, $M(f) \le \|f\|_2$.

Proof. Let $f(x)$ be factored as in Eq. (24.6) and $S := \{1 \le i \le n : |z_i| \le 1\}$ and $L := \{1 \le i \le n : |z_i| > 1\}$ – so S represents the small roots and L the large ones. Then $M(f) = |f_n| \prod_{i\in L} |z_i|$. Now let

$$g(x) := f_n \prod_{i\in L} (\bar{z}_i x - 1) \prod_{i \in S} (x - z_i) = \left(f_n \prod_{i\in L} \bar{z}_i\right) \cdot \prod_{i\in L}\left(x - \frac{1}{\bar{z}_i}\right)\prod_{i\in S}(x - z_i).$$

Then $M(f) = |f_n| \prod_{i\in L} |z_i| = \left|f_n \prod_{i\in L} \bar{z}_i\right| = |g_n|$ and therefore,

$$M(f) \le |g_n| \le \|g\|_2.$$

Letting $h(x) := f_n \prod_{i\in S}(x - z_i)$, we have, using Lemma 24.24 repeatedly, that

$$\|g(x)\|_2 = \left\|h(x)\prod_{i\in L}(\bar{z}_i x - 1)\right\|_2 = \left\|h(x)\prod_{i\in L}(x - z_i)\right\|_2 = \|f(x)\|_2.$$

Combining the last two equations completes the proof. ■

Theorem 24.26. *Suppose that* $h(x), f(x) \in (\mathbb{C}[x])^\times$, $\deg h = m \ge 1$, $\deg f = n \ge m$, *and* $h(x) | f(x)$. *Then*

$$\|h\|_2 \le \|h\|_1 \le 2^m M(h) \le 2^m \left|\frac{h_m}{f_n}\right| \|f\|_2.$$

Proof. Write h is factored form as

$$h(x) = h_m \prod_{i=1}^m (x - u_i)$$

where each factor, $(x - u_i)$, is also a factor of $f(x)$, i.e. $\{u_i\}_{i=1}^m \subset \{z_i\}_{i=1}^n$ with multiplicities. The k^{th} coefficient of $h(x)$ is given by

$$h_k = h_m \sum_{\substack{S \subset \{1,2,\dots,m\} \\ \#S = m-k}} \prod_{i\in S}(-u_i)$$

and therefore,

$$\begin{aligned}|h_k| &\le \sum_{\substack{S \subset \{1,2,\dots,m\} \\ \#S = m-k}} |h_m| \prod_{i\in S} |u_i| \\ &\le \sum_{\substack{S \subset \{1,2,\dots,m\} \\ \#S = m-k}} M(h) = \binom{m}{k} M(h).\end{aligned}$$

Summing this equation on k (using the binomial theorem) shows,

$$\|h\|_2 \le \|h\|_1 \le 2^m M(h).$$

Finally, looking at the definition of $M(\cdot)$, we see that $M(h) \le \frac{|h_m|}{|f_m|} M(f)$. ■

Corollary 24.27 (Mignotte's bound). *Suppose that $f, g, h \in \mathbb{Z}[x]$ have degrees $\deg f = n \geq 1$, $\deg g = m$, and $\deg h = k$, and that $gh|f$ in $\mathbb{Z}[x]$. Then*

$$\begin{aligned} \|g\|_\infty \|h\|_\infty &\leq \|g\|_2 \|h\|_2 \leq \|g\|_1 \|h\|_1 \ \textit{(abstract nonsense)} \\ &\leq 2^{m+k} \|f\|_2 \leq (n+1)^{1/2} 2^{m+k} \|f\|_\infty \end{aligned} \tag{24.7}$$

and

$$\begin{aligned} &\|h\|_\infty \leq \|h\|_2 \leq 2^k \|f\|_2 \leq 2^k \|f\|_1 \ \textit{and} \\ &\|h\|_\infty \leq \|h\|_2 \leq (n+1)^{1/2} 2^k \|f\|_\infty . \end{aligned} \tag{24.8}$$

Proof. First off, notice that $M(g) M(h) = M(gh) \leq M(f)$ since $(gh)_n = g_m h_k | f_n$. By Theorems 24.26 and 24.25,

$$\begin{aligned} \|g\|_1 \|h\|_1 &\leq 2^m M(g) 2^k M(h) = 2^{m+k} M(gh) \\ &\leq 2^{m+k} M(f) \leq 2^{m+k} \|f\|_2 \leq (n+1)^{1/2} 2^{m+k} \|f\|_\infty . \end{aligned}$$

The second set of inequalities follow by taking $g = 1$ in Eq. (24.7). ■

24.2.4 Fields of fractions (not covered in class)

Suppose you met an alien from a planet where they had invented rings, but because everything on their planet involved whole units, they had never invented the idea of division. So this guy knows about $\mathbb{Z}$, but not about fractions. How do you explain in an algebraic way about $\mathbb{Q}$?

Idea. An element of $\mathbb{Q}$ is a fraction $\frac{a}{b}$ where a, b are integers, that is, elements in $\mathbb{Z}$ and $b \neq 0$. But, the word fraction is undefined here, and the notation will be difficult to justify. So instead we could do the following—we identify $\frac{a}{b}$ with an ordered pair (a, b) with $a, b \in \mathbb{Z}$ and $b \neq 0$. Secretly thinking of fractions, now we can say how to multiply and add in $\mathbb{Q}$. To multiply, we set $(a, b) \cdot (c, d) = (ac, bd)$, and to add, we set $(a, b) + (c, d) = (ad + bc, bd)$, for $b, d \neq 0$. We can also divide them: $(a, b) \div (c, d) = (ad, bc)$, when $b, c, d \neq 0$.

However, there are problems to this approach. For example

$$(1, 2)[(1, 2) + (1, 2)] = (1, 2)(4, 4) = (4, 8)$$

$$(1, 2)(1, 2) + (1, 2)(1, 2) = (1, 4) + (1, 4) = (8, 16).$$

These are not the same, so distributivity fails. So to solve this problem, we need to say when two fractions are the same. We put an equivalence relation on $\mathbb{Q}$ saying that $(a, b) \sim (c, d)$ when $ad = bc$. Then we let $\mathbb{Q}$ be the equivalence classes and **refine** our notion of $\frac{a}{b}$ to mean the equivalence class containing (a, b). This works!

Now generally we start with a ring R, which is commutative with 1. We are unhappy if we can't divide one element of R by another. So we want to replace R by a bigger ring S, with $R \subseteq S$ where S is a field. That is, in S you can divide and two elements, as long as you don't divide by zero. When can we do this? How can we do this?

Example 24.28. Let $R = \mathbb{Z}$. Then we know $\mathbb{Z} \subset \mathbb{Q}$ where $\mathbb{Q}$ is a field. $\mathbb{Q}$ is also the "smallest" field with this property in some sense.

Lemma 24.29. *Suppose S is a field and $R \subset S$ is a subring. Then R is an integral domain.*

Proof. Since S is a field, S is an integral domain. Now if $ab = 0$ for some $a, b \in R$, then $ab = 0$ in S. Since S is an integral domain, $a = 0$ or $b = 0$ in S, and so also in R. ■

So if the R we start with is not a domain, we can't hope to form a bigger ring containing R which is a field. The next theorem shows this is the only obstruction to embedding a commutative ring, R, with identity into a "field of fractions."

Theorem 24.30. *Let R be a integral domain with 1 and $R^\times := R \setminus \{0\}$. Then we may define an equivalence relation on $R \times R^\times$ by,*

$$(a, b) \sim (c, d) \iff ad = bc.$$

Let $\frac{a}{b} = [(a, b)]$ denote the equivalence class determined by $\sim$ and $S := \left\{\frac{a}{b} : a \in R \text{ and } b \in R^\times\right\}$. Further define two binary operations, $+$ and $\cdot$, on S by,

$$\frac{a}{b} + \frac{c}{d} := \frac{ad + cb}{bd} \ \textit{and} \ \frac{a}{b} \cdot \frac{c}{d} := \frac{ac}{bd}. \tag{24.9}$$

Then with these operations, S becomes a field and the map,

$$R \ni a \to a/1 \in S$$

is an injective rings isomorphism. In the future we identify $a \in R$ with $a/1 \in S$ and views R as a sub-ring of S. The new ring S is called the ***field of fractions of the integral domain,*** *R.*

Proof. We're almost already done. The proof is the same as for $\mathbb{Z}$!

1. Define $\widetilde{S} = \{(a, b) : a, b \in R, b \neq 0\}$.
2. Put an equivalence relation on $\widetilde{S}$ by saying $(a, b) \sim (c, d)$ if $ad = bc$ in R. We start by checking that "$\sim$" is an equivalence relation:
   a) $(a, b) \sim (a, b)$ since $ab = ba$ in R.
   b) $(a, b) \sim (c, d) \Rightarrow (c, d) \sim (a, d)$ since $ad = bc \Rightarrow cd = da$.

   c) If $(a,b) \sim (c,d)$ and $(c,d) \sim (e,f)$, then $(a,b) \sim (e,f)$ since $ad = bc, cf = de \Rightarrow adf = bcf = bce$ so $daf = dbe$ and $af = be$. Here we used that R is a domain in order to use cancellation.

3. We now must show that operations in Eq. (24.9) makes sense. That is if
$$\frac{a}{b} = \frac{a'}{b'} \text{ and } \frac{c}{d} = \frac{c'}{d'}, \text{ i.e. } ab' = ba' \text{ and } cd' = dc',$$
then we must show,
$$\begin{aligned} \frac{ac}{bd} &= \frac{a'c'}{b'd'}, \text{ i.e. that } acb'd' = a'c'bd \text{ and} \\ \frac{ad+cb}{bd} &= \frac{a'd'+c'b'}{b'd'}, \text{ i.e. that } (ad+cb)\,b'd' = (a'd'+c'b')\,bd. \end{aligned}$$
These are easily checked;
$$acb'd' = (ab')\,(cd') = (ba')\,(dc') = a'c'bd$$
and
$$(ad+cb)\,b'd' = ab'dd' + bb'cd' = ba'dd' + bb'dc' = (a'd' + b'c')\,bd$$
as desired.
4. We must now show that $(S, \cdot, +)$ satisfies the axioms of a ring. There is lots to do here and we leave the tedious details to the reader. However, observe that $0_S = \frac{0}{1}$ and $1_S = \frac{1}{1}$.
5. The new ring S is a field. This is almost immediate. Given $\frac{a}{b} \in S^{\times}$, then $\frac{b}{a} \in S^{\times}$ and $\frac{a}{b} \cdot \frac{b}{a} = \frac{ab}{ab} = \frac{1}{1} = 1_S$. Thus $\left(\frac{a}{b}\right)^{-1} = \left(\frac{b}{a}\right)$ as was to be expected and hence S is a field.
6. Finally we observe that $\varphi : R \to S$ defined by $\varphi(r) = \frac{r}{1}$ is a one to one ring homomorphism. It is one-to-one because $\frac{r}{1} = \frac{r'}{1}$ iff $r \cdot 1 = r' \cdot 1$, i.e. iff $r = r'$. In this way we may now identify R with $\varphi(R) \subset S$.

$\blacksquare$

Once we know that we can do this, we think of elements of the field of fractions S as fractions, and we write them that way, instead of as ordered pairs.

Example 24.31. Let $R = \mathbb{R}[x]$. Then the field of fractions S of R looks like
$$S = \left\{ \frac{a_n x^n + \cdots + a_1 x + a_0}{b_m x^m + \cdots + b_1 x + b_0} : a_i, b_j \in \mathbb{R}, \sum_{j=0}^{m} b_j x^j \neq 0 \right\}$$
and you manipulate these gadgets in the obvious way. The traditional notation for S is $\mathbb{R}(x)$.

Example 24.32.
$$\frac{x}{x+2} + \frac{x^2}{5} = \frac{5x + x^2(x+2)}{5(x+2)} = \frac{x^3 + 2x^2 + 5x}{5x + 10}$$

Example 24.33. $R = \mathbb{Z}[i] = \left\{a + bi : a, b \in \mathbb{Z}, i^2 = -1\right\}$. S consists of fractions
$$\{\frac{a+bi}{c+di} : c + di \neq 0, i^2 = -1\}$$
under the equivalence relation
$$\frac{a+bi}{c+di} \sim \frac{a+bi}{c+di}\frac{a-bi}{c-di} = \frac{ac+bd+(bc-da)i}{c^2+d^2} = \frac{ac+bd}{c^2+d^2} + \frac{(bc-ad)i}{c^2+d^2}$$
In fact, it is not too hard to show $S \cong \mathbb{Q}[i] = \{a + bi : a, b \in \mathbb{Q}\}$.

Example 24.34. An infinite field of characteristic p. Let p be prime. $\mathbb{Z}_p$ is a finite field with characteristic p. $\mathbb{Z}_p[x]$ is an infinite ring with characteristic p. Let S be the field of fractions of $\mathbb{Z}_p$, written $\mathbb{Z}_p(x)$. Then S is an infinite field of characteristic p.
$$S = \left\{ \frac{f}{g} : f, g \in \mathbb{Z}_p[x],\ g \neq 0 \right\}$$

Example 24.35. Suppose we start with a field. What is its field of fractions? For example, what is the field of fractions of $\mathbb{Q}$? $S = \{(a,b) : a, b \in \mathbb{Q}, b \neq 0\}$ under $(a,b) \sim (c,d)$ if $ad = bc$. Then $(a,b) \sim (\frac{a}{b}, 1)$. Thus the map $\mathbb{Q} \to S$ by $q \mapsto (q, 1)$ is an isomorphism.

25

Lecture 25

25.1 Vector Spaces & Review of Linear Algebra

Definition 25.1. *Let F be a field (called "scalars"). A set V (called "vectors") is called a vector space over F if*

1. *V is an Abelian group under addition (written $+$).*
2. *For each $a \in F$, $v \in V$, there is an element $av \in V$ (this is called "scalar multiplication").*
3. *$a(v+u) = av + au \ \forall a \in F, u, v, \in V$ (scalar multiplication is group homomorphism).*
4. *$(a+b)v = av + bv \ \forall a, b \in F, v \in V$.*
5. *$a(bv) = (ab)v \ \forall a, b \in F, v \in V$.*
6. *$1v = v \ \forall v \in V$, where $1 = 1_F$ is the identity of F.*

These look a lot like ring axioms. But in general, a vector space V is not a ring. The two elements participating in this multiplication come from different sets. Let us recall a couple of basic notions from Math 20F before we go onto giving examples.

Definition 25.2. *A set S of vectors linearly dependent if there are vectors $v_1, \dots, v_n \in S$ and scalars $a_1, \dots, a_N$ not all zero with $a_1 v_1 + \cdots + a_n v_N = 0$. If S is not linearly dependent it is called linearly independent.*

Definition 25.3. *A set S is a basis for the vector space V if S is independent and spans V in that sense that ever $v \in V$ is of the form $\sum a_i v_i$ for some $a_i \in F$ and $v_i \in S$.*

Theorem 25.4. *Every vector space V over F has a basis. Moreover, the number of elements in the basis is unique, and is called the* ***dimension*** *of V or more precisely the* ***dimension of*** *V* ***over*** *F which we denote by $\dim_F V$.*

Proof. See the Gallian [2, Theorem 19.1] for one proof. Alternatively go back to your Math 20F course and see how it was done there. Here is a sketch of that method. The main point is to show if $\beta := \{v_1, \dots, v_m\}$ is a basis and $\gamma := (u_1, \dots, u_n)$ is a list of any $n-$ vectors in V with $n > m$, then γ is linearly dependent. The point is, since β is a basis, there exists $a_{ij} \in F$ such that

$$u_j = \sum_{i-1}^{m} v_i a_{ij} \text{ for all } 1 \le j \le n.$$

Thus if $\{b_j\}_{j=1}^n \subset F$ we will have $\sum_{j=1}^n b_j u_j = 0$ iff

$$0 = \sum_{j=1}^{n} b_j \sum_{i-1}^{m} v_i a_{ij} = \sum_{i-1}^{m} v_i \sum_{j=1}^{n} a_{ij} b_j.$$

Thus $\sum_{j=1}^n b_j u_j = 0$ if $(b_1, \dots, b_n)$ is a solution to the system of equations,

$$\sum_{j=1}^{n} a_{ij} b_j = 0 \text{ for } 1 \le i \le m.$$

However, from basic row reduction techniques one knows that such a homogenous system of $m-$ linear equations with $n > m$ unknowns always have non-trivial solutions and $\{u_j\}_{j=1}^n$ can not be linearly independent. $\blacksquare$

Example 25.5. In Math 20F, one concentrates on examples like:

- Suppose $F = \mathbb{R}$, and
$$V = \mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}.$$
Then V is a vector space with the usual vector addition (so V is an abelian group, $\underline{0} = (0,0)$ is an additive identity) and scalar multiplication $a(x,y) = (ax, ay)$. Here $\dim_{\mathbb{R}} \mathbb{R}^2 = 2$ with basis $\{(1,0), (0,1)\}$ for example.
- Suppose $F = \mathbb{R}$, and V is Euclidean 3-space,
$$V = \mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}$$
which again is a vector space in the analogous way. In this case, $\dim_{\mathbb{R}} \mathbb{R}^3 = 3$ and $\{(1,0,0), (0,1,0), (0,0,1)\}$ is a basis.
- More generally we might have taken $F = \mathbb{R}$ and $V = \mathbb{R}^n$ for any n in which case $\dim_{\mathbb{R}} V = n$.

Now that we have some ring theory, we can note that weirder examples have already come up. There are many examples of a ring R which is also a vector space over some field F.

Example 25.6. Let $R = F[x]$ for some field F. Then in addition to R being a ring, R is a vector space over F. R is already an abelian group under addition of polynomials. We define scalar multiplication in the obvious way by,

$$a[b_n x^n + \cdots + b_1 x + b_0] := ab_n x^n + \cdots + ab_1 x + ab_0.$$

It is easy to verify the axioms of a vector space. In fact, note that what is happening here is F is a subring of $F[x]$, just the scalar polynomial functions. So we already have a way of multiplying $a \in F, f \in F[x]$ together, the usual multiplication in $F[x]$. This is really all that scalar multiplication is doing. So in fact, all of the vector space axioms reduce to some special cases of the ring axioms!

Example 25.7. Suppose that F is any field and $V = F[x]$. Then V is an infinite dimensional vector space over F with basis, $\beta := \{1, x, x^2, x^3, x^4, \dots\}$. Indeed, everything in $F[x]$ may be written as $a_n x^n + \cdots + a_0$, i.e. a linear combination of the $\{x^n\}$. Moreover, β is independent since $\sum_{i=0}^n a_i x^i = 0$ happens iff $a_i = 0$ for all i by the very definition of the polynomial ring.

Definition 25.8. *Finally recall that if V is a vector space over a field, F. Then $W \subset V$ is a* ***subspace*** *of V if W is a subgroup of V which is closed under scalar multiplication.*

Example 25.9. Suppose that F is any field, $n \in \mathbb{N}$, and

$$V = \{f \in F[x] : \deg f(x) \leq n\}.$$

The V is a subspace of $F[x]$, $\beta = \{1, x, x^2, \dots, x^n\}$ is a basis for V and hence $\dim V = n + 1$.

Example 25.10. $M_2(F)$ is a vector space over F with scalar multiplication

$$a \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} aw & ax \\ ay & az \end{bmatrix}.$$

Identify F as a subring of $M_2(F)$, via

$$F \ni a \leftrightarrow \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \in \left\{ \begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix} \middle| b \in F \right\}.$$

Then scalar multiplication is really left multiplication by

$$\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}.$$

Example 25.11. Think of $M_2(\mathbb{R})$ as a $\mathbb{R}$-vector space. Then $M_2(\mathbb{R})$ is a vector space of dimension 4 over $\mathbb{R}$. A basis is

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

These are called elementary matrices. Thus $\dim_{\mathbb{R}} M_2(\mathbb{R}) = 4$.

Example 25.12. Here is an important example. Let $F \subset E$ where F and E are fields and F is a subring of E. Then we say F is a **subfield** of E. Notice that E is a vector space over F. The scalar multiplication is defined using the multiplication in E, i.e. if $a \in F$ and $v \in E$ then $av \in E$. The vector space axioms follow from ring axioms for E.

Notation 25.13 *If F is a subfield of E we let $[E : F] := \dim_F(E)$ be the dimension of E as an F vector space.*

Example 25.14. $\mathbb{R} \subset \mathbb{C}$. Then $\mathbb{C}$ is an $\mathbb{R}$-vector space with operation for $r \in \mathbb{R}$, $a + bi \in \mathbb{C}, r(a + bi) = ra + (rb)i$ and $[\mathbb{C} : \mathbb{R}] = 2$ – a basis being $\{1, i\}$.

Proof. Every element of $\mathbb{C}$ has the form $a + bi = a(1) + b(i)$ for some $a, b \in \mathbb{R}$. Moreover if $a(1) + b(i) = 0$ then $a = 0 = b$. [This justifies thinking of $\mathbb{C}$ as the plane which we tend to identify with $\mathbb{R}^2$.] Of course only the number of basis elements is unique. $S' = \{-1, -i\}$ is also a basis for $\mathbb{C}$ over $\mathbb{R}$. Similarly, $S'' = \{1 + i, \sqrt{3} - 2i\}$ is as well.

Example 25.15. $\mathbb{Q} \subset \mathbb{Q}[\sqrt{2}]$. Then $\mathbb{Q}[\sqrt{2}]$ (which is a field, as we proved) is a $\mathbb{Q}$-vector space with operation for $q \in \mathbb{Q}$, $a + b\sqrt{2} \in \mathbb{Q}[\sqrt{2}]$, $q(a + b\sqrt{2}) = qa + (qb)\sqrt{2}$. Then $\{1, \sqrt{2}\}$ is a $\mathbb{Q}$ – basis for $\mathbb{Q}\left[\sqrt{2}\right]$ so that $\left[\mathbb{Q}[\sqrt{2}] : \mathbb{Q}\right] = 2$. The proof is as in the previous example since every element of $\mathbb{Q}(\sqrt{2})$ is of the form $a(1) + b(\sqrt{2})$ for some $a, b \in \mathbb{Q}$ and $a(1) + b(\sqrt{2}) = 0$ happens iff $a = 0 = b$. The latter fact is a consequence of $\sqrt{2}$ being irrational.

Example 25.16. We have that $\mathbb{Q}$ is a sub-field of $\mathbb{R}$. However $[\mathbb{R} : \mathbb{Q}] = \infty$. One way to see this is to notice that any finite dimensional vector space over $\mathbb{Q}$ would have to be countable. Since $\mathbb{R}$ is uncountable, $\mathbb{R}$ must be an infinite dimensional vector space over $\mathbb{Q}$.

Corollary 25.17 (of Theorem 25.4). *If F is a finite field of characteristic p, then $\#(F) = p^n$ for some $n \in \mathbb{Z}_+$.*

Proof. Identify $\mathbb{Z}_p$ with the image of the ring homomorphism,

$$\mathbb{Z}_p \ni k \to k \cdot 1 \in F.$$

Then F may be viewed as a vector space over $\mathbb{Z}_p$. Let $n := \dim_{\mathbb{Z}_p} F \in \mathbb{Z}_+$. Then $\#(F) = p^n$ by a homework problem. ■

Lecture 26

Example 26.1. Let $F = \mathbb{Z}_2$, and let $f = x^3 + x + 1$, an irreducible polynomial over $\mathbb{Z}_2$. Then $R = \mathbb{Z}_2 / \left\langle x^3 + x + 1 \right\rangle$ is a field with 8 elements. Then we can think of F as a subfield of R by identifying $a \in F$ with $a + \left\langle x^3 + x + 1 \right\rangle$ in R. As such, R is a vector space over F and we have $\dim_F R = 3$. (Notice that $8 = 2^3 = \# (F)^{\dim_F R}$.)

Theorem 26.2. *Suppose that E is a field, $G \subset E$ is a subfield of E, and $F \subset G$ is a subfield of G. If $m := [E : G] < \infty$ and $n := [G : F] < \infty$, then $[E : F] < \infty$ and*

$$[E : F] = [E : G] [G : F].$$

Moreover if $\{u_i\}_{i=1}^m$ is a basis for E over G and $\{v_j\}_{j=1}^n$ is a basis for G over F, then $\beta := \{u_i v_j : 1 \le i \le m, \ 1 \le j \le n\}$ is an F – basis for E.

Proof. For any $a \in E$ there exists $a_i \in G$ such that $a = \sum_{i=1}^m a_i u_i$ and for each i there exists $a_{ij} \in F$ such that $a_i = \sum_{j=1}^n a_{ij} v_j$. Therefore it follows that

$$a = \sum_{i=1}^m \sum_{j=1}^n a_{ij} u_i v_j$$

and hence that the F – span of β is E. So to finish the proof we need only show β is linearly independent. If $a_{ij} \in F$ satisfy, $0 = \sum_{i=1}^m \sum_{j=1}^n a_{ij} u_i v_j$, then

$$0 = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} v_j \right) u_i.$$

Since $\{u_i\}_{i=1}^m$ is a basis for E over G it follows that $\sum_{j=1}^n a_{ij} v_j = 0$ for each i and since $\{v_j\}_{j=1}^n$ is a basis for G over F it follows that $a_{ij} = 0$ for each j. This completes the proof. ■

26.1 Field Theory

Definition 26.3. *Given fields E, F, with $F \subset E$ where F is a subring of E, we say that F is a subfield of E. We call E an **extension field** of F and the entire setup, $F \subseteq E$, is called a **field extension**.*

The setup for the time being will be that E is a field and $F \subset E$ is a sub-field. Typically we might have $E = \mathbb{C}$ and $F = \mathbb{Q}$ or F is some other sub-field of $\mathbb{C}$.

Definition 26.4. *We say an element $\alpha \in E$ is **algebraic over** F is there exists $p(x) \in F[x]$ such that $p(\alpha) = 0$. Otherwise we say $\alpha \in E$ is **transcendental** over F.*

Remark 26.5. For each $n \in \mathbb{Z}_+$ there are countably many polynomials in $\mathbb{Q}[x]$ of degree n and each of these polynomials have at most n - distinct zeros. Therefore there are only countable many $\mathbb{Q}$ – algebraic elements in $\mathbb{C}$. Since $\mathbb{C}$ is uncountable, most of the element in $\mathbb{C}$ are therefore transcendental. However, proving a specific given complex number is transcendental is not so easy. We state here without proof that both e and π are transcendental numbers.

Notation 26.6 *If S is a subset of E we let $F(S)$ denote the smallest subfield of E which contains F and S. This sub-field exists. Indeed,*

$$F(S) := \cap \{K \subset E : K \textit{ is a sub-field } \ni \ S \cup F \subset K\}.$$

Lemma 26.7. *If S and T are subsets of E, then $F(S)(T) = F(S \cup T) = F(T)(S)$.*

Proof. Sine $S \cup T \subset F(S)(T)$ it follows that $F(S \cup T) \subset F(S)(T)$. Conversely it is clear that $F(S) \subset F(S \cup T)$ and that $T \subset F(S \cup T)$, therefore $F(S)(T) \subset F(S \cup T)$. ■

Example 26.8. $\mathbb{R}[i] = \mathbb{C}$ and $\mathbb{Q}(\sqrt{2}) = \left\{a + b\sqrt{2} : a, b \in \mathbb{Q}\right\}$ as we have see before.

Theorem 26.9 (Adjoining a root). *Suppose that $\alpha \in E$ is algebraic over F and $\varphi : F[x] \to E$ is the evaluation homomorphism, $\varphi(f(x)) = f(\alpha)$. Then;*

1. *$\ker \varphi = \langle p(x) \rangle$ where $p(x)$ is an irreducible polynomial which may be taken to be the unique monic polynomial $f(x) \in F[x]$ of minimal degree which has α as a root. (This polynomial is called the **minimal polynomial** of $\alpha \in E$.)*

2. *We have*

$$F(\alpha) = F[\alpha] = \{f(\alpha) : f \in F[x]\}$$
$$= \{f(\alpha) : f \in F[x] \ \text{with} \ \deg f(x) < \deg p(x)\}.$$

3. *The map, $\bar{\varphi} : F[x] / \langle p(x) \rangle \to F(\alpha)$ defined by $\bar{\varphi}([f(x)]) = f(\alpha)$ is an isomorphism of fields.*
4. $[F(\alpha) : F] = \deg(p(x)) =: n$

$$\beta := \left\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\right\} \tag{26.1}$$

is a basis for $F(\alpha)$ over F.
5. $[F(\alpha) : F] \leq \deg g(x)$ *where $g(x) \in F[x]$ is any non-zero polynomial such that $g(\alpha) = 0$.*

Proof. 1. Since $F[x]$ is a principle ideal domain and $\ker\varphi$ is an ideal, we know that $\ker\varphi = \langle p(x) \rangle$ where $p(x)$ may be taken to be the unique monic polynomial in $\ker\varphi = \{f(x) \in F[x] : f(\alpha) = 0\}$ with minimal degree. If $p(x)$ were not irreducible, it would have a non-trivial factorization, $p(x) = h(x)k(x)$ where $\deg h(x), \deg k(x) < \deg p(x)$. Moreover $p(\alpha) = 0$ would then imply $h(\alpha) = 0$ or $k(\alpha) = 0$. But then $h(x)$ or $k(x)$ would be in $\ker\varphi$ with degree less than $p(x)$ which is impossible.

2. and 3. By the first isomorphism theorem we know that

$$\bar{\varphi} : F[x] / \langle p(x) \rangle \to F[\alpha] = \{f(\alpha) : f \in F[x]\}$$

so defined above is an isomorphism of rings. Since $p(x)$ is irreducible we know $\langle p(x) \rangle$ is maximal and therefore $F[x] / \langle p(x) \rangle$ is a field. Therefore $F[\alpha]$ is a field and we are justified in writing $F(\alpha) = F[\alpha]$. Lastly if $f(x) \in F[x]$, then $f(x) = k(x)p(x) + r(x)$ with $\deg r(x) < \deg p(x)$. Therefore, $f(\alpha) = k(\alpha)p(\alpha) + r(\alpha) = r(\alpha)$ and this shows that

$$F(\alpha) = \{f(\alpha) : f \in F[x] \ \text{with} \ \deg f(x) < \deg p(x)\}.$$

4. Since $\bar{\varphi}$ is a field isomorphism it is also a vector space of $F-$isomorphism. Therefore,

$$[F(\alpha) : F] = \dim_F (F[x] / \langle p(x) \rangle)$$
$$= \dim\{f(x) \in F[x] : \deg f(x) < \deg p(x)\} = \deg p(x) = n.$$

If $f(x) = \sum_{j=0}^{n-1} a_j x^j \in F[x]$, then $f(\alpha) = \sum_{j=0}^{n-1} a_j \alpha^j$ which shows that $f(\alpha)$ is a spanning set of $F(\alpha)$. Since $[F(\alpha) : F] = n$ it follows that β must indeed be a basis.

5. This simply follows from item 4. since $\deg(p(x)) \leq \deg(g(x))$. ■

Example 26.10. $\sqrt{-3}$ has minimal polynomial $x^2 + 3$ over $\mathbb{Q}$. Why? It is monic, $\sqrt{-3}$ is a root, and it is irreducible over $\mathbb{Q}$ since it has no real and hence no rational roots.

Example 26.11. What is the minimal polynomial over $\mathbb{R}$ for $i \in \mathbb{C}$? Answer; $x^2 + 1$.

27

Lecture 27

Example 27.1. Since x^2-3 is irreducible over $\mathbb{Q}$, it must be the minimal polynomial for $\sqrt{3}$. Therefore,

$$\left[\mathbb{Q}\left(\sqrt{3}\right):\mathbb{Q}\right]=\deg\left(x^2-3\right)=2,$$

and the general element of $\mathbb{Q}\left(\sqrt{3}\right)$ may be written as $a+b\sqrt{3}$ with $a,b\in\mathbb{Q}$.

Example 27.2. Let us now consider $\mathbb{Q}\left(\sqrt{3},\sqrt{5}\right)=\mathbb{Q}\left(\sqrt{3}\right)\left(\sqrt{5}\right)$. Notice that $\sqrt{5}$ is a root of x^2-5. Moreover we claim that x^2-5 is irreducible over $\mathbb{Q}\left(\sqrt{3}\right)$. The point is that $\pm\sqrt{5}\notin\mathbb{Q}\left(\sqrt{3}\right)$ and so x^2-5 has no roots in $\mathbb{Q}\left(\sqrt{3}\right)$. Indeed, if $\sqrt{5}=a+b\sqrt{3}\in\mathbb{Q}\left(\sqrt{3}\right)$ then it would follow that

$$5=\left(a+b\sqrt{3}\right)^2=a^2+3b^2+2ab\sqrt{3}.$$

Since $\sqrt{3}\notin\mathbb{Q}$, we would have to have $ab=0$, i.e. $a^2=5$ or $3b^2=5$ for some $a,b\in\mathbb{Q}$ which is not possible since x^2-5 and $3x^3-5$ are both irreducible over $\mathbb{Q}$. Thus we may conclude that, x^2-5 is the minimal polynomial over $\mathbb{Q}\left(\sqrt{3}\right)$ for $\sqrt{5}$ and

$$\left[\mathbb{Q}\left(\sqrt{3},\sqrt{5}\right):\mathbb{Q}\left(\sqrt{3}\right)\right]=\deg\left(x^2-5\right)=2.$$

Consequently, by Theorem 26.2,

$$\left[\mathbb{Q}\left(\sqrt{3},\sqrt{5}\right):Q\right]=\left[\mathbb{Q}\left(\sqrt{3},\sqrt{5}\right):\mathbb{Q}\left(\sqrt{3}\right)\right]\left[\mathbb{Q}\left(\sqrt{3}\right):\mathbb{Q}\right]=2\cdot 2=4 \quad (27.1)$$

and moreover $\left\{1,\sqrt{3},\sqrt{5},\sqrt{3}\sqrt{5}=\sqrt{15}\right\}$ is a basis for $\mathbb{Q}\left(\sqrt{3},\sqrt{5}\right)$.

Proposition 27.3. *If $\alpha\in\mathbb{C}$ is transcendental, then $\varphi:\mathbb{Q}[x]\to\mathbb{C}$ defined by $\varphi(f(x))=f(\alpha)$ has trivial kernel and therefore $\mathbb{Q}[x]\cong\mathbb{Q}[\alpha]$. Moreover, if $\mathbb{Q}(x)$ denotes the field of rational function with rational coefficients, then $\mathbb{Q}(\alpha)\cong\mathbb{Q}(x)$.*

From now on we will be sticking with the algebraic numbers inside of $\mathbb{C}$.

27.1 Ruler and Compass Constructions

Let us identify the plane with the complex numbers. The type of question we would like to address, given $0,1\in\mathbb{C}$ (i.e. fix an origin in the plan and a unit length) what complex numbers in the plane are constructible just using a straight edge and a compass. Let us begin with some (hopefully) familiar constructions from high school geometry.

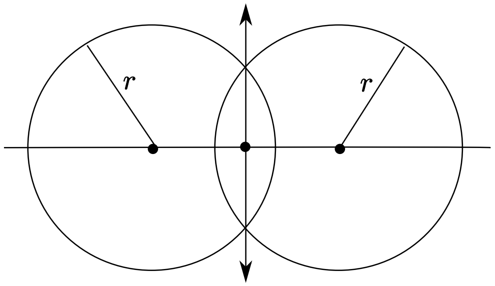


Fig. 27.1. This figure should remind you of how to bisect a line segment $\overline{ab}$ to get point c. Alternatively if you start with point c on the line l, this picture indicates how to construct the line perpendicular to l which goes through c.

Using the abilities of this construction it now follows that given a line l and a point p not in l we can construct the line l' going through p which intersects l at a right angle, see Figure 27.2.

Making use of dropping perpendicular lines, given a line l and a point p not in l we we can construct the line l' which goes through p that is parallel to l as in Figure 27.3.

The other standard construction from high school geometry is the ability to bisect an angle which is demonstrated in Figure 27.4.

We these constructions in hand we may now work out the numbers which are constructible. We do this in the next few results.

Lemma 27.4. *Suppose that $r,s\in(0,\infty)$ are given are representable by line segments of these lengths where we are given a unit length segment as reference. Then we can also represent; $1/r$, rs, and $\sqrt{s}$.*

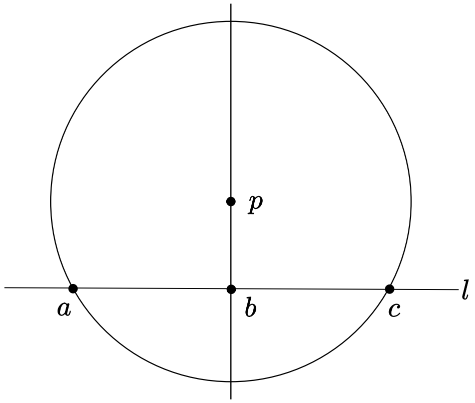


Fig. 27.2. A perpendicular bisector.

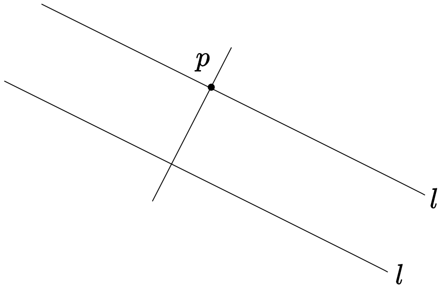


Fig. 27.3. Constructing parallel lines using an intermediary perpendicular line.

Proof. The proofs of these assertions are encoded in the following pictures.

To construct the square root of r the reader is referred to Figure 27.7 below. Using the notation in this figure we learn by Pythagorean's theorem that

$$(s + 1)^2 = |ib - 1|^2 + |ib - \ s|^2 .$$

Expanding out both sides of this equation gives,

$$s^2 + 2\,s + 1 = b^2 + 1 + b^2 + \ s^2 = \ s^2 + 2b^2 + 1,$$

which implies that $b^2 = s$, i.e. $b = \sqrt{s}$.

■

It should be obvious that given ways to measure $r, s \in \mathbb{R}$ we can construct the point $r + is \in \mathbb{C}$ with a straight edge and compass, see Figure 27.8.

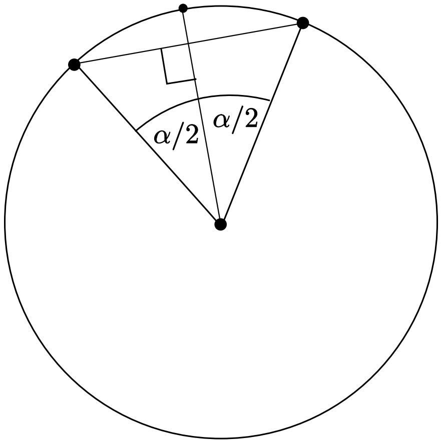


Fig. 27.4. Bisecting a given angle.

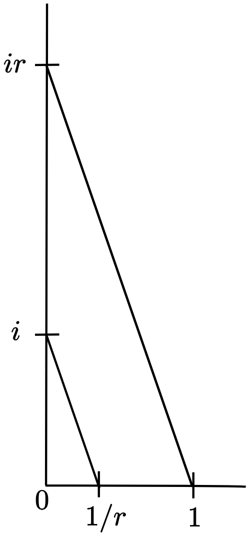


Fig. 27.5. Given segments of length 1 and r here is how to construct a segment of length $1/r$. This may be viewed as a special case the the construction in Figure 27.6 as well.

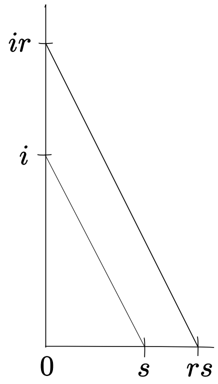


Fig. 27.6. Given segments of length 1, r, and s here is how to construct a segment of length rs. On the other hand if we turn this around and suppose that $\alpha := rs$ and r are the given lengths, then this picture shows how to construct $s = \alpha/r$ as well.

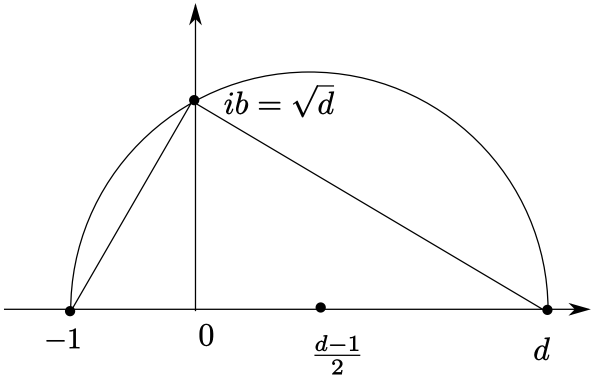


Fig. 27.7. This is a semi-circle of radius $(d+1)/2$.

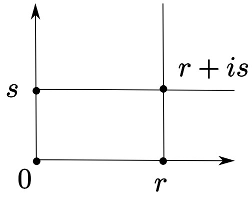


Fig. 27.8. Constructing a complex number from its real and imaginary parts.

Remark 27.5. Suppose $E \subset \mathbb{C}$ is a sub-field of $\mathbb{C}$ which is closed under complex conjugation. Let us now look for a square root, $x + iy$, of $a + ib \in E$. Thus we must have,

$$x^2 - y^2 = a \text{ and } 2xy = b.$$

Solving the last equation for y and substituting the result into the first gives an equation for x, namely,

$$x^2 - \left(\frac{b}{2x}\right)^2 = a \iff 4x^4 - 4ax^2 - b^2 = 0.$$

We are looking for real roots of this equation. By the quadratic formula it follows that

$$x^2 = \frac{4a \pm \sqrt{16a^2 + 16b^2}}{2 \cdot 4} = \frac{a \pm \sqrt{a^2 + b^2}}{2}$$

and we must choose the positive root here to learn,

$$x^2 = \frac{a + \sqrt{a^2 + b^2}}{2}.$$

Thus we find that

$$x = \pm\sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}}$$

and therefore that

$$y = \pm\frac{b}{2\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}}} = \pm\frac{b}{\sqrt{2a + 2\sqrt{a^2 + b^2}}}.$$

Thus we see that $x^2 \in E\left[\sqrt{a^2 + b^2}\right]$ and that $x, y \in E\left[\sqrt{a^2 + b^2}, \sqrt{\frac{a+\sqrt{a^2+b^2}}{2}}\right]$.

As an example if $a + ib = \frac{3}{5} + \frac{4}{5}i$ then

$$x^2 = \frac{\frac{3}{5} + \sqrt{\left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2}}{2} = \frac{4}{5} \text{ and } x = \frac{2}{\sqrt{5}}.$$

Lemma 27.6. *Suppose that $z, w \in \mathbb{C}$ are given complex numbers which have put on the complex plane along with 1. Then;*

1. *the real and imaginary parts of z are representable.*
2. *$z + w$ and $z - w$ are representable.*
3. *$\bar{z}$ is representable.*
4. *zw is representable.*
5. *If $z \neq 0$, then $1/z$ is representable.*

6. *If $z \neq 0$, the two square-roots of z are representable.*

Proof. Most of these are simple to realize.

1. The real and imaginary parts are found by right angle projections of z onto that x and y – axis which we know how to do.
2. It should be clear how to represent $z + w$ and $z - w$ when z and w are real. Thus we may do the complex case by considering the real and imaginary parts. (Better yet, use the usual graphical methods for vector addition which only need the ability to construct parallel lines going through given points.)
3. Similarly if $z = x + iy$ then we can represent x and y and hence x and $-y$ and therefore $\bar{z} = x - iy$. (Alternatively you may easily do this graphically as well.)
4. Again this can be done by working with the real and imaginary parts of zw and using that we know how to work with real lengths. Alternatively write $z = re^{i\theta}$ and $w = \rho e^{i\alpha}$. Then the line going through $e^{i(\theta+\alpha)}$ as in Figure 27.9 and mark the point on this line which is distance $r\rho$ from the origin.
5. For this simply observe that $\bar{z}z = |z|^2$ from which it follows that
$$\frac{1}{z} = \bar{z}\frac{1}{|z|^2}$$
which is representable by what we have already proved.
6. Case 1. Suppose that $z = re^{i\theta} = r\cos\theta + ir\sin\theta$ with $0 \leq \theta \leq \pi$. To construct the square roots of z, draw the line through $0 \in \mathbb{C}$ which bisects the angle θ between $\overline{0\,1}$ and $\overline{0\,z}$. Then mark the two points on this line which are $\sqrt{|z|}$ from the origin. These are the graphical representations of the two square roots of z.
Case 2. Suppose that $z = re^{i\theta} = r\cos\theta + ir\sin\theta$ with $0 \leq \theta \leq \pi$. In this case we may write $z = re^{-i(2\pi-\theta)}$ and then work as in the previous example to bisect the angle between $\overline{0\,1}$ and $\overline{0\,z}$ and mark the points as before. In doing this we will construct,
$$\pm\sqrt{r}e^{-i(2\pi-\theta)/2} = \pm\sqrt{r}e^{-i\pi+\theta/2} = \mp\sqrt{r}e^{i\theta/2}$$
which again represents the two square roots of z.

■

Our results so far easily give the following propositions.

Proposition 27.7. *If $\{0,1\}$ are plotted in the complex plane, we may construct $\mathbb{Q}[i] = \mathbb{Q}+i\mathbb{Q} \subset \mathbb{C}$ using only a compass and a straight edge. Moreover we may construct $\sqrt{z}$ for all $z \in \mathbb{Q}[i]$.*

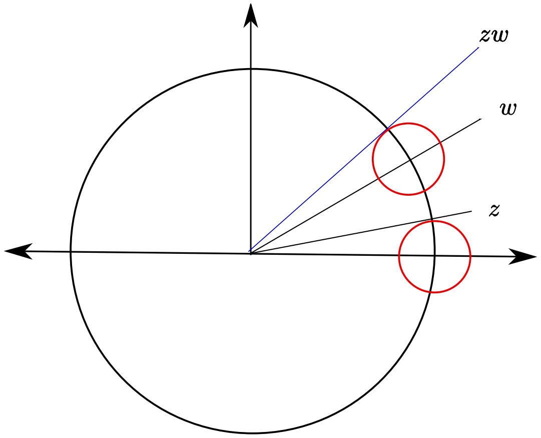


Fig. 27.9. Multiplying complex numbers with a straight edge and a compass.

Proposition 27.8. *Suppose that $S \subset \mathbb{C}$ is a set containing $\{0,1\}$ and $F(S)$ are the complex numbers which are constructible from S using compass and a straight edge. Then S is a sub-field of $\mathbb{C}$ which contains $\mathbb{Q}$. Moreover if $z \in F(S)$ then $\operatorname{Re} z$, $\operatorname{Im} z$, and $\bar{z}$ are all back in $F(S)$.*

Proof. From Lemma 27.6, $F(S)$ is closed under subtraction, multiplication and complex conjugation. From this it follows that $F(S)$ *is* a field which is closed under complex conjugations. Since
$$\operatorname{Re} z = \frac{z+\bar{z}}{2} \text{ and } \operatorname{Im} z = \frac{z-\bar{z}}{2i}$$
it now follows that $\operatorname{Re} z, \operatorname{Im} z \in F(S)$. ■

Theorem 27.9 (Theorem 9.24 of Knapp, p. 466). *The set, $\mathcal{C}$, of x coordinates that can be constructed from $x = 0$ and $x = 1$ by straightedge and compass forms a subfield of $\mathbb{R}$ such that the square root of any positive element of the field lies in the field. Conversely the members of $\mathcal{C}$ are those real numbers lying in some subfield F_n of $\mathbb{R}$ of the form,*
$$F_1 = \mathbb{Q}(\sqrt{a_0}),\ F_2 = F_1(\sqrt{a_1}),\ \ldots F_n = F_{n-1}(\sqrt{a_{n-1}}),$$
where $a_j \in F_j$ and $a_0, \ldots, a_{n-1} \geq 0$.

Proof. We have already seen that we can construct the elements described in F_n by a ruler and straightedge so it only remains to see that we can not construct anything else by these methods. To see this, observe that the only points we may construct are gotten by intersection lines going through two points already constructed with other such lines or with circles centered at points with radii from points we have already constructed.

So suppose that we have constructed a sub-field, $E \subset \mathbb{Q}$ using a ruler and straightedge. Then a line joining (a_1, b_1) and (a_2, b_2) with $a_i, b_j \in E$ is described by

$$\frac{y - b_1}{x - a_1} = \frac{b_2 - b_1}{a_2 - a_1}$$

or equivalently of the form $y = mx + b$ or $x = my + b$ with $m, b \in E$. An allowed circle is of the form,

$$(y - u)^2 + (x - v)^2 = c^2 \text{ with } u, v, c \in E.$$

1. Then the intersection of two lines;

   $$y = mx + b \text{ and } y = m'x + b'$$

   gives a linear equation for x of the form $(m - m')\, x + b - b' = 0$ which only has solutions from E. So no new points are found this way.
2. If we intersect a line and a circle, we find an equation for x of the form,

   $$(mx + b - u)^2 + (x - v)^2 = c^2$$

   which is quadratic in x. If there is (real) solution to this equation it will lie in $E\left(\sqrt{\gamma}\right)$ for some $\gamma \in E$ with $\gamma > 0$.
3. Now suppose we intersect two circles say,

   $$(y - u_1)^2 + (x - v_1)^2 = c_1^2 \text{ and } (y - u_2)^2 + (x - v_2)^2 = c_2^2$$

   with all of the u's, v's and c's coming from E. By translating the whole picture by (u_2, v_2) we may assume that $u_2 = v_2 = 0$. Thus it suffices to consider the intersection points of

   $$(y - u)^2 + (x - v)^2 = c^2 \text{ and } y^2 + x^2 = d^2 \text{ with } u, v, c, d \in E.$$

   Since

   $$c^2 = (y - u)^2 + (x - v)^2 = u^2 - 2vx - 2uy + v^2 + x^2 + y^2$$

   at the intersection points we must have

   $$c^2 = u^2 - 2vx - 2uy + v^2 + d^2$$

   which is a linear equation for x. Thus we are back to case 2. namely the intersection of a line and a circle and the proof is complete.

■

Corollary 27.10. *Letting F_n be as in Theorem 27.9 we have $[F_n : \mathbb{Q}] = 2^k$ for some $0 \le k \le n$.*

28

Lecture 28

As an application of this corollary, let us show that $\sqrt[3]{p} \notin F_n$ for any prime p. The reason is that if $\sqrt[3]{p} \in F_n$ then $\mathbb{Q}\left(\sqrt[3]{p}\right) \subset F_n$ and we must have

$$[F_n : \mathbb{Q}] = [F_n : \mathbb{Q}\left(\sqrt[3]{p}\right)]\left[\mathbb{Q}\left(\sqrt[3]{p}\right) : \mathbb{Q}\right].$$

However, $x^3 - p$ is irreducible over $\mathbb{Q}$ by Eisenstein and therefore $\left[\mathbb{Q}\left(\sqrt[3]{p}\right) : \mathbb{Q}\right] = \deg\left(x^3 - p\right) = 3$. But $3 \nmid 2^k$ and so it is impossible for $\sqrt[3]{p}$ to be in F_n.

28.1 Geometric Consequences for the Greeks

Example 28.1 (Doubling the cube). In this example we show that it is impossible to "double the cube" with a compass and straight edge. That is we can not use a compass and straight edge only to construct a cube with volume 2. To do this we would need to construct a cube with side length equal to $\sqrt[3]{2}$ which have just seen is impossible.

Example 28.2 (Trisecting an angle). In this example we show that it is impossible to trisect a 60° angle. First off notice that

$$\cos 60 = 1/2 \text{ and } \sin 60 = \sqrt{3}/2$$

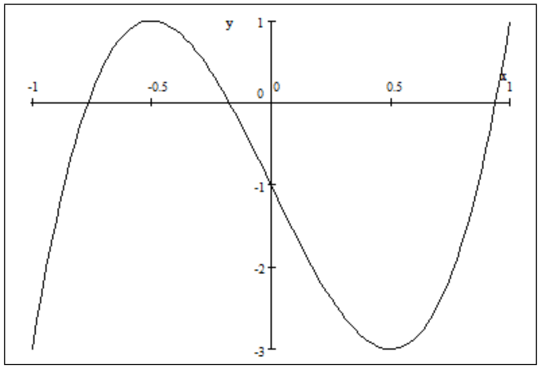
which are constructible and therefore so is 60°. In order to trisect this angle it would have to be possible to construct $a := \cos\left(20°\right)$. To see this is not possible, observe that

$$\begin{aligned}\cos 3\theta &= \operatorname{Re}\left(e^{i3\theta}\right) = \operatorname{Re}\left(\cos\theta + i\sin\theta\right)^3 \\ &= \cos^3\theta - 3\cos\theta\sin^2\theta = \cos^3\theta - 3\cos\theta\left(1 - \cos^2\theta\right) \\ &= 4\cos^3\theta - 3\cos\theta. \end{aligned} \tag{28.1}$$

Thus taking $\theta = 20°$ in Eq. (28.1) we learn that

$$4\alpha^3 - 3\alpha = \cos 60 = \frac{1}{2},$$

i.e. α is a root of $f(x) = 8x^3 - 6x - 1$.



The graphs of $f(x)$.

But the mod 5 test shows this polynomial is irreducible since,

x	0	1	2	3	4
$f(x) \bmod 5$	4	1	1	2	2

.

Alternatively one may check that $\{\pm 1, \pm 1/2, \pm 1/4, \pm 1/8\}$ are not[1] roots of $f(x)$ so that $f(x)$ has no rational roots. No matter how this is done we learn that $[\mathbb{Q}(\alpha) : \mathbb{Q}] = \deg(f(x)) = 3$ which does not divide 2^k for any k and hence is not constructible. This same argument applies to all three roots of $f(x)$.

Lemma 28.3. *Suppose that D is a regular n – gon in $\mathbb{C}$, then the center of D may be determined by summing the coordinates of the vertices of the n – gon D and in particular the center of D may be found using only a ruler and compass.*

Proof. Suppose that the vertices of D are described by $x_k := z + \alpha\omega^k$ for $k = 0, 1, 2 \ldots, n-1$ where ω is a primitive n^{th} – root of unity, $\alpha \neq 0$ in $\mathbb{C}$ and z is the center we are trying to find. Then making use of the relation,

$$\sum_{k=0}^{n-1} \omega^k = \frac{\omega^n - 1}{\omega - 1} = 0 \tag{28.2}$$

[1] From the graph in Figure 28.2, we see that $-1/4$ and $-1/8$ are the most likely candidates to be a rational roots of $f(x)$ – however $f(-1/4) = \frac{3}{8} \neq 0$ and $f(-1/8) = -\frac{17}{64} \neq 0$.

we see that

$$\frac{1}{n}\sum_{k=0}^{n-1} x_k = \frac{1}{n}\sum_{k=0}^{n-1}\left(z+\alpha\omega^k\right) = z.$$

Better yet if n is even draw lines through ω^k and $\bar{\omega}^k$ to find the center as the common intersection point while if n is odd draw lines through ω^k and the midpoint of the opposite side to find the center, see Figure 28.1.

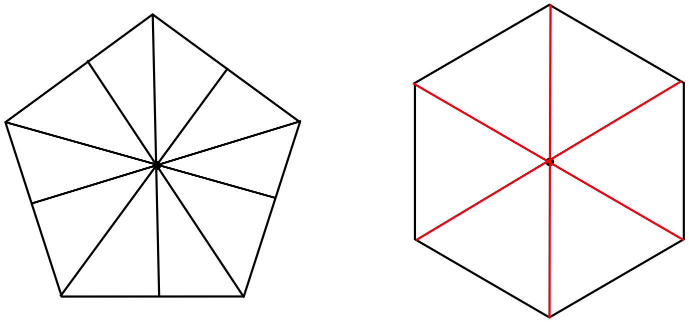


Fig. 28.1. Finding the center of a pertagon and a hexagon.

■

Example 28.4 (9 - gons are not constructible). If a regular 9 -gon were constructible then we could also construct the angle, $40° = 360/9$, i.e. we could construct $\alpha := \cos 40$. But by virtue of Eq. (28.1) we know that

$$4\alpha^3 - 3\alpha = \cos(3\cdot 40) = \cos(120) = \frac{1}{2}.$$

As in Example 28.2, none of the roots of $4x^3 - 3x - 1/2$ are constructible and therefore $\alpha = \cos 40°$ is not constructible.

Example 28.5 (Pentagons are constructible). In order to construct a pentagon we must be able to construct $\theta := 2\pi/5$ or equivalently $\alpha := \cos(2\pi/5)$. However, recall from Eq. (28.2) that

$$\begin{aligned} 0 &= 1 + e^{i\theta} + e^{i2\theta} + e^{i3\theta} + e^{i4\theta} \\ &= 1 + e^{i\theta} + e^{i2\theta} + e^{-i2\theta} + e^{-i\theta} \\ &= 1 + e^{i\theta} + e^{-i\theta} + e^{i2\theta} + e^{-i2\theta}. \end{aligned}$$

Taking the real part of this identity gives,

$$0 = 1 + 2\cos\theta + 2\cos 2\theta.$$

Then using the double angle formula,

$$\cos 2\theta = \cos^2\theta - \sin^2\theta = 2\cos^2\theta - 1$$

allows us to conclude that

$$2\left(2\cos^2\theta - 1\right) + 2\cos\theta + 1 = 0,$$

i.e. α is a root of

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

This polynomial is irreducible since

x	0	1	2
$f(x) \bmod 3$	2	2	1

.

Alternatively use the quadratic formula to see the roots of $f(x)$ are given by

$$\frac{-2\pm\sqrt{16+4}}{8} = \frac{-2\pm 2\sqrt{5}}{8} = \frac{-1\pm\sqrt{5}}{4}$$

so that $f(x)$ has no rational roots. Thus we conclude that $\mathbb{Q}(\alpha) = \mathbb{Q}\left(\sqrt{5}\right)$ so that α is constructible. Actually we have won the game when we found that α was a root of some $f(x) \in \mathbb{Q}[x]$ with $\deg f(x) = 2$ – we need not have gone any further (why?).

The general result along these lines (see Gallian [2, Theorem 33.5, p. 570]) is the following theorem due to Gauss, 1796.

Theorem 28.6. *It is possible to construct the regular n -gon with a straightedge and compass iff $n = 2^k p_1 \dots p_l$, where $k \geq 0$ and $p_1, \dots, p_l$ are and distinct primes each of which is of the form, $2^m + 1$ for some $m \geq 2$.*

For example when $5 = 2^2 + 1$ so pentagon's are constructible as we have seen. On the other hand $7 \neq 2^m + 1$ and therefore the 7 - gon is not constructible. Also $9 = 3 \cdot 3$ so the 9 -gon is not constructible since the odd prime, 3, appears twice in the decomposition of 9. On the other hand, $17 = 2^4 + 1$ is prime so the 17 - gon is constructible.

28.2 Splitting fields over $\mathbb{Q}$

Definition 28.7. *Given a field $F \subseteq E$ and a polynomial $f \in F[x]$, we say that f* ***splits over*** *E if f factors into irreducibles as*

$$f(x) = c(x-a_1)(x-a_2)\dots(x-a_n)$$

for some $a_i \in E$ and $c \in F^\times$.

Example 28.8. $f(x) = x^2 - 2$ does not split over $\mathbb{Q}$ but does split over $\mathbb{R}$;

$$x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2}).$$

Theorem 28.9. *Given any polynomial $f \in F[x]$ for some field F, there exists a field extension $F \subseteq E$ such that f splits over E.*

We'll mostly stick to the case $F = \mathbb{Q}$. In this case to find a field over which a polynomial splits we may make use Theorem 19.12 – the fundamental theorem of algebra which we restate here.

Theorem 28.10 (Fundamental theorem of algebra). *Every polynomial $f \in \mathbb{C}[x]$ splits over $\mathbb{C}$.*

Now suppose $f(x) \in \mathbb{Q}[x]$. We know $f\left(x\right)$ splits over $\mathbb{C}$, but there is a big distance between $\mathbb{Q}$ and $\mathbb{C}$, and sometimes we want to change our fields only a little to a find a field in which $f\left(x\right)$ splits.

Definition 28.11. *Let $F \subset E$ be a field extension, and $f\left(x\right) \in F[x]$. Suppose $f\left(x\right)$ splits over E. The splitting field for $f\left(x\right)$ is the smallest subfield K with $F \subseteq K \subseteq E$ such that $f\left(x\right)$ splits over K. Explicitly, if $f(x) = (x - a_1) \dots (x - a_n)$ with $a_1, \dots, a_n \in E$, then $K = F\left(a_1, \dots, a_n\right)$ is the smallest subfield of E containing F and $a_1, \dots a_n$.*

Mostly we will use this definition in the case $F = \mathbb{Q}$, $E = \mathbb{C}$. Then for any polynomial $f\left(x\right) \in \mathbb{Q}[x]$, $f\left(x\right)$ splits over $\mathbb{C}$, so $f\left(x\right)$ has a splitting field inside $\mathbb{C}$.

Example 28.12. According to Theorem 28.10, $f\left(x\right) = x^3 + 1 \in \mathbb{Q}\left[x\right]$ must split over $\mathbb{C}$ – but how? It is not necessarily easy to find the factors in general! However in this case we see that -1 is a root of f and by the division algorithm that $x^3 + 1 = (x + 1)(x^2 - x + 1)$. The roots of $x^2 - x + 1$ are

$$\frac{1 \pm \sqrt{1 - 4}}{2} = \frac{1}{2} \pm \frac{\sqrt{3}i}{2}$$

and thus we may conclude that (over $\mathbb{C}$) $f\left(x\right)$ factorizes as;

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

The splitting field of $f\left(x\right)$ over $\mathbb{Q}$ is then

$$K = \mathbb{Q}\left(-1, \frac{1}{2} - \frac{\sqrt{3}i}{2}, \frac{1}{2} + \frac{\sqrt{3}i}{2}\right).$$

We can simplify our expression for K and in the process get a better idea what is in it. Since $-1 \in \mathbb{Q}$ already, the inclusion in the list defining K is irrelevant. Also observe that $\left(\frac{1}{2} + \frac{\sqrt{3}i}{2}\right) \in K$ and $\frac{1}{2} \in \mathbb{Q} \subset K$ allows us to conclude that

$$\sqrt{-3} = 2\frac{\sqrt{3}}{2}i = 2\left(\frac{1}{2} + \frac{\sqrt{3}i}{2} - \frac{1}{2}\right) \in K$$

and therefore, $\mathbb{Q}(\sqrt{-3}) \subseteq K$. Conversely,

$$-1,\ \frac{1}{2} + \frac{\sqrt{3}i}{2},\ \frac{1}{2} - \frac{\sqrt{3}i}{2} \in \mathbb{Q}(\sqrt{-3}) \implies K \subset \mathbb{Q}(\sqrt{-3})$$

and hence we may conclude that

$$K = \mathbb{Q}(\sqrt{-3}) = \mathbb{Q}\left[\sqrt{-3}\right] = \{a + b\sqrt{-3} : a, b \in \mathbb{Q}\}.$$

Let us recall how to compute $\left(a + b\sqrt{-3}\right)^{-1}$;

$$\begin{aligned}\left(a + b\sqrt{-3}\right)^{-1} &= \frac{1}{a + b\sqrt{-3}} \cdot \frac{a - b\sqrt{-3}}{a - b\sqrt{-3}} \\ &= \frac{a - b\sqrt{-3}}{a^2 + 3b^2} = \frac{a}{a^2 + 3b^2} - \frac{b}{a^2 + 3b^2}\sqrt{-3}\end{aligned}$$

which is well defined provided that $a + b\sqrt{-3} \neq 0$. Let us summarize what we have proved.

Lemma 28.13. *The splitting field for $x^3 + 1 \in \mathbb{Q}\left[x\right]$ is $\mathbb{Q}\left(\sqrt{-3}\right) = \mathbb{Q}\left[\sqrt{-3}\right]$. This is also the splitting field for $x^2 - x + 1$ and for $x^2 + 3$.*

Example 28.14. In this example we wish to find the splitting field for $x^4 - 3 \in \mathbb{Q}\left[x\right]$. Let $\omega := e^{i2\pi/4} = e^{i\pi/2} = i$ be a primitive 4th root of unity. Then the roots of $x^4 - 3$ are

$$\sqrt[4]{3} \cdot \left\{1, \omega, \omega^2, \omega^3\right\} = \sqrt[4]{3} \cdot \{\pm 1, \pm i\} = \cdot \left\{\pm\sqrt[4]{3}, \pm i\sqrt[4]{3}\right\}.$$

Thus the splitting field for $f\left(x\right)$ is

$$\mathbb{Q}\left(\pm\sqrt[4]{3}, \pm i\sqrt[4]{3}\right) = \mathbb{Q}\left(\sqrt[4]{3}, i\right).$$

Since $x^2 + 1$ is irreducible[2] over $\mathbb{Q}\left(\sqrt[4]{3}\right)$ and $x^4 - 3$ is irreducible (Eisenstein's criteria for example) over $\mathbb{Q}$, it follows that

[2] It is irreducible over $\mathbb{R}$ and hence over $\mathbb{Q}\left(\sqrt[4]{3}\right)$ since $x^2 + 1$ has not real roots.

$$\begin{aligned}\left[\mathbb{Q}\left(\sqrt[4]{3},i\right):\mathbb{Q}\right] &= \left[\mathbb{Q}\left(\sqrt[4]{3},i\right):\mathbb{Q}\left(\sqrt[4]{3}\right)\right]\left[\mathbb{Q}\left(\sqrt[4]{3}\right):\mathbb{Q}\right]\\ &= \deg\left(x^2+1\right)\cdot\deg\left(x^4-3\right) = 2\cdot 4 = 8.\end{aligned}$$

A basis for $\mathbb{Q}\left(\sqrt[4]{3},i\right)$ over $\mathbb{Q}$ is

$$\{1,i\}\cdot\left\{1,3^{1/4},3^{1/2},3^{3/4}\right\} = \left\{1,\sqrt{3},\sqrt[4]{3},3^{\frac{3}{4}},i,i\sqrt{3},i\sqrt[4]{3},i3^{\frac{3}{4}}\right\}.$$

28.3 More practice on understanding field extensions of $\mathbb{Q}$

Theorem 28.15 (Primitive element theorem. Steinitz, 1910). *If $f(x)\in\mathbb{Q}[x]$ and $\mathbb{Q}\subseteq K\subseteq\mathbb{C}$ is the splitting field for $f(x)$, then $K=\mathbb{Q}(a)$ for some single element $a\in\mathbb{C}$.*

We will not prove this theorem here (but see Gallian[2, Theorem 21.6 on p.375]). In certain examples, we can see how to find a single element by guessing.

Lemma 28.16. *Suppose that p and q are distinct primes. Then $\sqrt{p}\notin\mathbb{Q}\left(\sqrt{q}\right)$.*

Proof. If $\sqrt{p}\in\mathbb{Q}\left(\sqrt{q}\right)$ then $\sqrt{p}=a+b\sqrt{q}$ for some $a,b\in\mathbb{Q}$. Squaring this equation would then imply that

$$p=a^2+b^2q+2ab\sqrt{q}. \tag{28.3}$$

Since $\sqrt{q}$ is irrational, we must have $ab=0$, i.e. $a=0$ or $b=0$. So in order to solve Eq. (28.3), we must solve one of the equations, $qb^2=p$ or $a^2=p$ with $a,b\in\mathbb{Q}$. But it is easy to check that both qx^2-p and x^2-p are irreducible of $\mathbb{Q}$ so that these equations have not solutions and hence there is no solution to Eq. (28.3), i.e. $\sqrt{p}\notin\mathbb{Q}\left(\sqrt{q}\right)$. To see that qx^2-p and x^2-p are irreducible of $\mathbb{Q}$ observe that polynomials are primitive and therefore we may apply Eisenstein criterion using the prime, p. ■

Example 28.17. Consider $\mathbb{Q}(\sqrt{3},\sqrt{5})$. This is the smallest subfield of $\mathbb{C}$ containing $\mathbb{Q}$, $\sqrt{3}$, and $\sqrt{5}$. We claim that $\mathbb{Q}(\sqrt{3},\sqrt{5})=\mathbb{Q}(\sqrt{3}+\sqrt{5})$

Proof. Since $\sqrt{3}+\sqrt{5}$ is in $\mathbb{Q}(\sqrt{3},\sqrt{5})$, and $\mathbb{Q}\subset\mathbb{Q}(\sqrt{3},\sqrt{5})$, it follows that $\mathbb{Q}(\sqrt{3}+\sqrt{5})\subseteq\mathbb{Q}(\sqrt{3},\sqrt{5})$. The harder part is to prove the reverse inclusion.

Since $\mathbb{Q}(\sqrt{3}+\sqrt{5})$ is a field, we know $(\sqrt{3}+\sqrt{5})^{-1}\in\mathbb{Q}(\sqrt{3}+\sqrt{5})$. Since

$$\frac{1}{\sqrt{3}+\sqrt{5}}=\frac{1}{\sqrt{3}+\sqrt{5}}\frac{\sqrt{3}-\sqrt{5}}{\sqrt{3}-\sqrt{5}}=\frac{\sqrt{3}-\sqrt{5}}{-2}=-\frac{1}{2}\sqrt{3}+\frac{1}{2}\sqrt{5}$$

it follows that $-\frac{1}{2}\sqrt{3}+\frac{1}{2}\sqrt{5}\in\mathbb{Q}(\sqrt{3}+\sqrt{5})$. Therefore,

$$\sqrt{3}-\sqrt{5}=-2\left(-\frac{1}{2}\sqrt{3}+\frac{1}{2}\sqrt{5}\right)\in\mathbb{Q}(\sqrt{3}+\sqrt{5})$$

and hence

$$\sqrt{3}+\sqrt{5}\pm\left(\sqrt{3}-\sqrt{5}\right)\in\mathbb{Q}(\sqrt{3}+\sqrt{5})$$

from which it follows that $\sqrt{3},\sqrt{5}\in\mathbb{Q}(\sqrt{3}+\sqrt{5})$. Thus we have shown $\mathbb{Q}(\sqrt{3},\sqrt{5})\subseteq\mathbb{Q}(\sqrt{3}+\sqrt{5})$ which completes the proof. ■

In general finding the minimal polynomial for some algebraic element, $\alpha\in\mathbb{C}$, can be challenging.

Example 28.18. Let us find the minimal polynomial for $\alpha:=\sqrt{3}+\sqrt{5}$ over $\mathbb{Q}$. We first look for a monic polynomial, $p(x)\in\mathbb{Q}[x]$ such that $p(\alpha)=0$. We begin by observing that

$$\alpha^2=(\sqrt{3}+\sqrt{5})^2=8+2\sqrt{15}$$

and therefore,

$$\left(\alpha^2-8\right)^2=\left(2\sqrt{15}\right)^2=4\cdot 15=60. \tag{28.4}$$

Thus we see that

$$p(x):=\left(x^2-8\right)^2-60=x^4-16x^2+4 \tag{28.5}$$

is a polynomial having α as a root.

We claim this polynomial is the minimal polynomial for α. One way to see this is to use Example 27.2 where we showed,

$$4=[\mathbb{Q}[\alpha]:Q]=\deg(\text{the minimal polynomial for }\alpha).$$

Since $\deg p(x)=4$ and $p(\alpha)=0$, it must in fact be the minimal polynomial for α!

Remark 28.19. Another way to show that $p(x)$ in Eq. (28.5) is the minimal polynomial such that $p(\alpha)=0$ would be to show that $p(x)$ is irreducible over $\mathbb{Q}$. One way to do this is to first find all of the roots of $p(x)$. From Eq. (28.5), $p(a)=0$ for some $a\in\mathbb{C}$ then $\left(a^2-8\right)^2=60$ and therefore,

$$a^2=8\pm\sqrt{60}=8\pm 2\sqrt{15}.$$

So we may conclude that $\left\{a(\varepsilon,\delta)=\varepsilon\sqrt{8+\delta 2\sqrt{15}}\right\}_{\varepsilon,\delta=\pm1}$ are all the roots of $p(x)$. To see that $p(x)$ is irreducible it suffices to observe that $p(x)$ can not be factored into a product of quadratic polynomials since $a(\varepsilon,\delta)\,a(\varepsilon_0,\delta_0)\notin\mathbb{Q}$ whenever $(\varepsilon,\delta)\neq(\varepsilon_0,\delta_0)$.

Note: We could also try to use the $\bmod p$ test here. However the test fails for $p=2$ and $p=3$. For example, if $\bar{p}(x)=p(x) \bmod 3$, then

$$\bar{p}(x)=x^4-x^2+1=x^4+2x^2+1=\left(x^2+1\right)^2$$

is reducible. So using a $\bmod p$ test is becoming rather painful.

Example 28.20. Let us now consider $\mathbb{Q}\left(\sqrt{3},\sqrt{5},\sqrt{7}\right)$. I claim that $\sqrt{7} \notin \mathbb{Q}\left(\sqrt{3},\sqrt{5}\right)$. To see this let $x,y \in \mathbb{Q}\left(\sqrt{3}\right)$ so that $x+y\sqrt{5}$ is the general element of $\mathbb{Q}\left(\sqrt{3},\sqrt{5}\right)$. Thus we would have to have $\sqrt{7}=x+y\sqrt{5}$ for some $x,y \in \mathbb{Q}\left(\sqrt{3}\right)$. Squaring this equation then implies that

$$7=x^2+5y^2+2xy\sqrt{5}.$$

Since we have already see that $\sqrt{5} \notin \mathbb{Q}\left(\sqrt{3}\right)$, it will only be possible to solve this equation if $xy=0$, i.e. either $x=0$ or $y=0$. Thus we must either solve $x^2=7$ or $5y^2=7$ for some x or y in $\mathbb{Q}\left(\sqrt{3}\right)$. or equivalently. But working as in the previous example, these equations have no solutions over $\mathbb{Q}\left(\sqrt{3}\right)$ and we are done. From these observations we may conclude that

$$\left[\mathbb{Q}\left(\sqrt{3},\sqrt{5},\sqrt{7}\right):\mathbb{Q}\right]=2^3=8.$$

Example 28.21. In this example we wish to compute $n:=\left[\mathbb{Q}\left(\sqrt[3]{2},\sqrt[4]{3}\right):\mathbb{Q}\right]$. Using

$$\begin{aligned}\left[\mathbb{Q}\left(\sqrt[3]{2},\sqrt[4]{3}\right):\mathbb{Q}\right] &= \left[\mathbb{Q}\left(\sqrt[3]{2},\sqrt[4]{3}\right):\mathbb{Q}\left(\sqrt[3]{2}\right)\right]\left[\mathbb{Q}\left(\sqrt[3]{2}\right):\mathbb{Q}\right]\\ &= \left[\mathbb{Q}\left(\sqrt[3]{2},\sqrt[4]{3}\right):\mathbb{Q}\left(\sqrt[3]{2}\right)\right]\cdot 3\\ &\text{and}\\ \left[\mathbb{Q}\left(\sqrt[3]{2},\sqrt[4]{3}\right):\mathbb{Q}\right] &= \left[\mathbb{Q}\left(\sqrt[3]{2},\sqrt[4]{3}\right):\mathbb{Q}\left(\sqrt[4]{3}\right)\right]\left[\mathbb{Q}\left(\sqrt[4]{3}\right):\mathbb{Q}\right]\\ &= \left[\mathbb{Q}\left(\sqrt[3]{2},\sqrt[4]{3}\right):\mathbb{Q}\left(\sqrt[4]{3}\right)\right]\cdot 4\end{aligned}$$

it follows that $3|n$ and $4|n$ and therefore that $12|n$. Since $\left[\mathbb{Q}\left(\sqrt[3]{2},\sqrt[4]{3}\right):\mathbb{Q}\left(\sqrt[3]{2}\right)\right] \leq 4$ as $\sqrt[4]{3}$ is a root of $x^4-4 \in \mathbb{Q}\left(\sqrt[3]{2}\right)[x]$ we also know that $n \leq 12$. Therefore $n=12$.

29

Lecture 29 (Review Day)

The Final Exam Questions will be similar to Quiz Questions.

29.1 Definitions you should know:

- Ring (but I will not ask you to check that something is or is not a ring directly from the definition).
- Identity element of a ring.
- Commutative ring.
- Unit.
- Subring.
- Zero-divisor and the cancellation property.
- Integral domain
- Field.
- Characteristic of a ring. (I only care about characteristic for rings R with identity, and I define it to be the smallest positive integer n such that $n \cdot 1 = 0$, or if no such n exists the characteristic is defined to be 0.)
- Ideal.
- Factor ring.
- Prime ideal.
- Maximal ideal.
- Homomorphism and isomorphism.
- Kernel and image of a homomorphism.
- Polynomial rings, $R[x]$ and most importantly $F[x]$ where F is a field.
- Principle ideal domain.
- Unique factorization domain.
- $F(\alpha)$ where F is a sub-field of a field E and $\alpha \in E$, e.g. $\mathbb{Q}\left(\sqrt[4]{5}\right)$.
- $[E : F] = \dim_F(E)$.
- Algebraic and transcendental elements of an extension field.
- Minimal polynomial.

29.2 Examples of Rings

We have only studied a few classes of rings. You should know all of these and their basic properties. How do you multiply and add in each one? Which are integral domains and which aren't? Which are fields? Which are commutative and which are noncommutative? Which have an identity element and what is it? What is the characteristic of each ring?

- Rings of numbers: $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.
- $\mathbb{Z}_m$, the integers modulo m, for any $m \geq 2$.
- Matrix rings: $M_2(F)$,which is 2×2 - matrices with entries from F. Here F could be any of the rings of numbers above, or even $\mathbb{Z}_m$ for some m.
- The Gaussian integers $\mathbb{Z}[i] = \{a+b : a, b \in \mathbb{Z}\}$, where $i = \sqrt{-1}$. Also studied $\mathbb{Z}_m[i]$ a bit.
- The ring of polynomials $F[x]$, which consists of all elements of the form

  $$ax^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0,$$

  where the coefficients ai all come from F. Here F could be any of the rings of numbers above, or even $\mathbb{Z}_m$ for some m.
- The ring $\mathbb{Z}[\sqrt{m}]$, where m is a positive integer which is not a square. This ring consists of all elements

  $$\{a + b\sqrt{m} : a, v \in \mathbb{Z}\}.$$

- Given any two rings R and S, the direct sum of R and S is a new ring

  $$R \oplus S = \{(r, s) : r \in R \text{ and } s \in S\},$$

  with component-wise addition and multiplication.
- $F[\alpha] := \{p(\alpha) : p \in F[x]\}$ where $\alpha \in E$ and E is some field extension of F.

29.3 Important theorems and techniques

- Know how to check if a subset of a ring is a subring.
- Know how to check if a subset of a ring is an ideal of the ring.
- Know the theorem that a finite integral domain with identity is a field.
- Know that $\mathbb{Z}_m$ is a field precisely when m is prime, and understand why this fails when m is not prime.
- Understand the example $\mathbb{Q}\left[\sqrt{2}\right]$ and understand the proof that it is a field.

- Know the theorem that the characteristic of a domain is a prime number (or 0).
- Given a commutative ring R with element a, know the definition of the **principle ideal** generated by a, written as $\langle a\rangle$.
- Understand the definition of a factor ring and how to do addition and multiplication in such a ring.
- Understand some important examples where factor rings can be shown to be the same as other familiar rings. $\mathbb{Z}/\langle m\rangle \cong \mathbb{Z}_m$. $\mathbb{R}[x]/\langle x\rangle \cong \mathbb{R}$. There are also various problems where one looks at factor rings of $\mathbb{Z}[i]$. For example, we showed $\mathbb{Z}[i]/\langle 2-i\rangle \cong \mathbb{Z}_5$.
- Know the theorem that a ideal I of a commutative ring R is prime if and only if R/I is a domain, and that I is maximal if and only if R/I is a field. As a corollary, know that a commutative ring R with identity is a field if and only if R and $\{0\}$ are the only ideals of R.
- Be able to check if a function between two rings is a homomorphism, and if it is a isomorphism.
- Know that the kernel of a homomorphism $\varphi : R \to S$ is always an ideal of R , and the image of a homomorphism is always a subring of S. Know the statement of the 1^{st} – isomorphism theorem: $R/\ker(\varphi) \cong \operatorname{Im}\varphi$ and how to use it.
- Know that given any ring R with identity, there is a homomorphism $\mathbb{Z} \to R$ sending a to $a \cdot 1$. The kernel of this homomorphism is exactly $\langle m\rangle$, where m is the characteristic of R.
- Know how to determine the homomorphisms, $\varphi : \mathbb{Z}[i] \to R$ such that $\varphi(1) = 1$ – recall i must be sent to $t \in R$ such that $t^2 = -1_R$.
- Understand how to manipulate polynomials and how to carry out long division in $F[x]$ where F is a field like $\mathbb{Q}$ or $\mathbb{Z}_p$ where p is prime.
- Understand how to find the roots of polynomials and their relationship to linear factors.
- Know how to check if a polynomial in $\mathbb{Q}[x]$ is irreducible or not. See Section 22.2 for a summary of this point.
- Know the relationships between prime (maximal) principle ideals and prime and maximal elements of an integral domain. See Theorem 23.6 for a summary.
- If $F \subset E \subset G$ are fields then $[G : F] = [G : E][E : F]$.
- $F(\alpha) \subset E$ is the smallest subfield of E containing F and α.
- Recall that if α is algebraic over F, then $F(\alpha) = F[\alpha]$ and $[F(\alpha) : F] = \deg f(x)$ where $f(x) \in F[x]$ is the minimal polynomial over F of α. See Theorem 26.9 for details.

30

Error Correcting Codes

30.1 Algebraic Coding Theory.

Applications to telecommunications, internet, etc., not cryptography.

Idea: we want to send out a message over a cable line, fiber optic line, satellite connection, etc. but somehow avoid errors which are introduced by noise in the communication channel.

Our messages usually consists of a sequence of words in some alphabet and this alphabet is often $\{0, 1\}$ coming from the bits in a computer. "Noise" in the system can cause errors by flipping the value of a given bit. We want to know if the received message has errors and if so we would like to correct them. An easy method for doing this is to send the message more than once!

For example suppose the message is sent in triplicate. Then if there is one error in the entire process, it can be detected and corrected. An alternate version is to send each digit three times in a row.

Example 30.1. Code words are all strings of four 0's and 1's. Say the message we would like to transmit is 0101 1100 0001. We would then send

$$000\ \ 111\ \ 000\ \ 111\ \ 111111\ \ 000000\ \ 000\ \ 000\ \ 000\ \ 111.$$

If there were a single transmission error, for example maybe we received, 000 111000 111 **101** 111000 000..., then we know that 101 is obviously wrong since it is not a string of three like digits. Also, assuming there was at most one error, it is clear that we should interpret 101 to be 111. So this method can correct any single error. It can also detect two errors, but it is not able to correct them. For example if we only know that at most two errors could have occurred in the above transmission we may still conclude that 101 was an error but now we don't know if it should have been 111 (one error) or 000 (two errors). So up to two errors may be detected but not corrected and knowing at most one error occurred we can detect and correct that error. Three errors in general cannot even be detected by this scheme.

This is the general idea, and it's all fine, but it's wasteful. You send three times the number of digits in the actual message, and there may be some expense involved in the extra digits (satellite time, mostly time issues). So the topic of this lecture and the next is to discuss one way (there are many more complicated ones) to create error correcting / detecting codes which are more efficient (require fewer digits to be transmitted).

Definition 30.2. *An (n, k)-**linear binary code** is a kdimensional subspace V of the vector space*

$$\mathbb{Z}_2^n = \{(a_1, a_2, \dots, a_n : a_i \in \mathbb{Z}_2\}.$$

*Members of V are called **code words.***

Out of laziness (and easier to read) we write tuples without parentheses or commas. So the 4-tuple $(0, 1, 1, 0) \in \mathbb{Z}_2^4$ is written 0110, that is, vectors are identified with strings of 0's and 1's.

Example 30.3 (The main one we will study). Let G be the 4×7 matrix

$$G = \begin{pmatrix} 1\,0\,0\,0 \mid 1\,1\,0 \\ 0\,1\,0\,0 \mid 1\,0\,1 \\ 0\,0\,1\,0 \mid 1\,1\,1 \\ 0\,0\,0\,1 \mid 0\,1\,1 \end{pmatrix} \tag{30.1}$$

with entries from $\mathbb{Z}_2 = \{0, 1\}$. Consider the linear transformation

$$\mathbb{Z}_2^4 \to \mathbb{Z}_2^7 \text{ defined by}$$
$$\varphi(a_1, a_2, a_3, a_4) := (a_1, a_2, a_3, a_4)G. \tag{30.2}$$

The image of φ is a four-dimensional subspace of $\mathbb{Z}_2^7$. So $V = \operatorname{Ran}(\varphi)$ is a (7,4) binary code.

Similarly, if

$$G = \begin{pmatrix} 1\,0\,0 \mid 1\,1\,0 \\ 0\,1\,0 \mid 1\,0\,1 \\ 0\,0\,1 \mid 1\,1\,1 \end{pmatrix} \tag{30.3}$$

we may construct $(6, 3)$ binary code,

$$V := \{(a_1, a_2, a_3)\, G : a_1, a_2, a_3 \in \mathbb{Z}_2\} \subset \mathbb{Z}_2^6.$$

Example 30.4. Let G and φ be as in Eqs. (30.1) and (30.2), then (for example),

$$\varphi(1, 0, 1, 0) = \overset{\text{row 1 of } G}{(1000110)} + \overset{\text{row 3 of } G}{(0010111)} = (1010001).$$

In general, the code words are all possible sums of some collection of rows of G!

$$\begin{array}{lll}
0000 \mapsto 0000000 & 1001 \mapsto 1001101 & 0111 \mapsto 0111011 \\
1000 \mapsto 1000110 & 0110 \mapsto 0110010 & 1111 \mapsto 1111111 \\
0100 \mapsto 0100101 & 0101 \mapsto 0101110 & \\
0010 \mapsto 0010111 & 0011 \mapsto 0011100 & \\
0001 \mapsto 0001011 & 1110 \mapsto 1110100 & \\
1100 \mapsto 1100011 & 1101 \mapsto 1101000 & \\
1010 \mapsto 1010001 & 1011 \mapsto 1011010 &
\end{array}$$

V has 16 vectors out of the $2^7 = 128$ vectors in $\mathbb{Z}_2^7$. V is a subspace of $\mathbb{Z}_2^7$ automatically since it is the image of a linear transformation So the sum of any two vectors in the code is another code vector – very important!

(Also V is closed under scalar multiplication. This doesn't mean much since $1 \cdot v = v$ and $0 \cdot v = 0$ for any v.)

Notice. Any 4-tuple $a_1a_2a_3a_4$ is **encoded** by φ as a 7-triple $a_1a_2a_3a_4b_1b_2b_3$. So the first four digits are the same and there are three extra digits. These are called **check digits**. Now suppose we want to send a message. The message is a sequence of 4-tuples in $\mathbb{Z}_2$. We encode these as 7-tuples using φ so we send a sequence of length 7 code words instead.

Claim. This code can correct any error! We are only sending less than fifty percent extra digits (versus two-hundred percent!)

Example 30.5. The message is

$$\begin{array}{l}
0001 \;\; 1010 \;\; 1110 \text{ encodes to} \\
0001011 \;\; 1010001 \;\; 1110100.
\end{array}$$

If a single digit error occurs and the message is received as 0001011 **1110001** 1110100, we see that 1110001 is not a code word.

This is all set up so that ;

Example 30.6. 1. If 1 digit in a code word is altered, the result is not a code word. So the error is detected.
2. There is a **unique** code word that differs from the altered one by a single digit. So the error can be corrected.
3. Above, 1010001 is the only word that differs from 1110001 by a single digit, so we know that 1010001 was the intended second word in the message.

Could just check that the code has properties 1 and 2. But we want to be more systematic and so we will be able to produce other codes with properties like this.

Definition 30.7. *Let v, w be vectors in $\mathbb{Z}_2^n$. The* ***distance,*** *$d(v, w)$, between v and w is the number of coordinates in which the two vectors differ. The* ***weight of a vector,*** *$wt(v)$, is $d(v, 0)$ which is the number of 1's appearing in v. If $V \subset \mathbb{Z}_2^n$ is a code, the* ***weight of the code,*** *$wt(V)$, is the smallest of the weights of the non-zero vectors in the code, V, i.e.*

$$wt\,(V) := \min\{wt\,(v) : v \in V \setminus \{0\}\}\,.$$

Example 30.8. In the (7,4) code V above, $d(0100101, 01111001) = 3$, and $wt(0100101) = 3$. By inspection every vector in the code has at least 3 ones (except the zero vector), so the weight of the code, V, is 3.

Theorem 30.9. *Let u, v, w be vectors in a code. Then;*

1. *$d(u, v) = wt(u - v)$ and*
2. *$d(u, v) \le d(u, w) + d(w, v)$ (**triangle inequality**).*

Proof. 1. $d(u, v)$ is the number of coordinates where u, v differ. The i^{th} coordinate of $u - v$ is 0 if u and v have the same i^{th} coordinate, and 1 otherwise. So the number of 1's in $u - v$ is the number of differences between u and v. So $d(u, v) = wt(u - v)$.

2. Note that for any vectors x, y

$$wt(x + y) \le wt(x) + wt(y)$$

This is clear: the total number of ones in $x + y$ is at most the number of ones in x, plus the number of ones in y, but usually smaller. Then

$$d(u, v) = wt(u - v) \le wt(u - w) + wt(w - v) = d(u, w) + d(w, v).$$

■

We should think of $d(\cdot, \cdot)$ as a distance function. The smaller $d(u, w)$ is, the "closer" u and w are to each other.

Theorem 30.10. *If V is a (n, k) code with weight greater than or equal to $2t + 1$, then the code can correct any t or fewer errors. Alternatively, the code can detect (but not correct) any $2t$ or fewer errors.*

Proof. Think of the elements of V as a subset of $\mathbb{Z}_2^n$. Then since V is closed under subtraction, for any $v_1, v_2 \in V$, $d(v_1, v_2) = wt(v_1 - v_2)$. So either $v_1 = v_2$ so $v_1 - v_2 = 0$, or $d(v_1, v_2) \ge wt(v_1 - v_2) \ge 2t + 1$. In other words, any two distinct code words are at least $2t + 1$ apart from one another. This means that if we consider the "sphere" of radius t about each code word, these sphere do not intersect!, see Figure 30.1. Here is the formal argument.

Error correction. Let $v \in V$ and suppose $v' \in \mathbb{Z}_2^n$ is the result of t or fewer errors to v, i.e. $d(v', v) \le t$. Let w be another code word (i.e. a vector in V), with $w \ne v$. Then

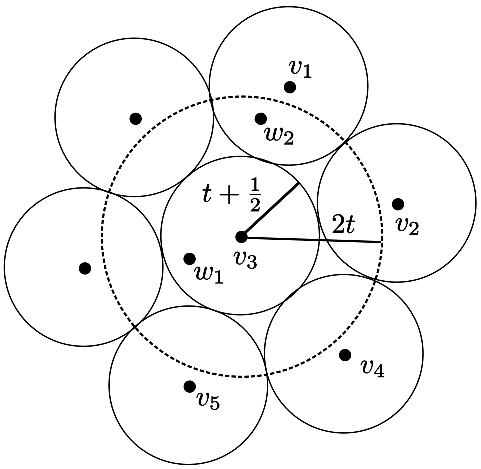


Fig. 30.1. A schematic form of the geometry of the hamming code. In this figure v_3 is the only code word within distance t from w_1. Also w_2 is a message with distance $2t$ of v_3 but not equal to v_3. As we see from the picture, w_2 can not be equal to any code word so an error has been detected. On the othe rhand we do now know how to correct this error since the message w_2 could be v_1 or v_3 or possibly even some other code word.

$$\begin{aligned} 2t+1 &\leq d\left(v, w\right) \leq d\left(v, v'\right) + d\left(v', w\right) \\ &\leq t + d\left(v', w\right) \end{aligned} \tag{30.4}$$

which shows, $d\left(v', w\right) \geq t+1$ and therefore v is the unique code within distance t from w. So v' can be corrected to v, by simply finding the code word of minimum distance to v (fewest differences in position).

Error detection. Next suppose instead we alter a code word v to some $v' \in \mathbb{Z}_2^n$ with $v' \neq v$ but differing from v by changing at most $2t$ digits, i.e. $1 \leq d(v, v') \leq 2t$. We claim that v' is not a code word (i.e. $v' \notin V$) and therefore we can detect $2t$ or less errors to a code word. Indeed, if $v' \in V$, then $0 \neq v - v' \in V$ (as V is a vector space!) and therefore,

$$2t \geq d\left(v, v'\right) = wt\left(v - v'\right) \geq 2t+1.$$

The last inequality is absurd and hence we may conclude that $v' \notin V$ and we have detected that an error has occurred. ■

Corollary 30.11. *The (7,4) code above (with weight* $3 = 2 \cdot 1 + 1$*) can correct any* 1 *error or alternatively detect any two errors.*

Next time we will answer the question; how do we efficiently detect and correct errors, without just reading through the list of code words one by one?

30.2 Hamming Codes

Recall that a binary (n, k) code is a $k-$ dimensional subspace V of $\mathbb{Z}_2^n$. A **Hamming code** is one given by a $k \times n$ matrix $G = \left[I \mid *\right]$ where I is the identity matrix and $*$ is any matrix. The Hamming code associated to G is the subspace or row vectors;

$$V := \{vG : v \in \mathbb{Z}_2^n\} \subset \mathbb{Z}_2^n.$$

Equivalently, V consists of all linear combinations of the rows of G.

Example 30.12. The matrix,

$$G = \begin{bmatrix} 1\,0\,0 \mid 1\,1\,0 \\ 0\,1\,0 \mid 1\,0\,1 \\ 0\,0\,1 \mid 1\,1\,1 \end{bmatrix}$$

determines a binary Hamming code

$$V = \{100110, 010101, 001111, 000000, 110011, 011010, 101001, 111100\}.$$

Remember that the weight, $wt(v)$, of a vector v is the number of 1's appearing in v and the distance between v and w, $d(v, w)$, is the number of positions or coordinates in which the vectors v and w differ. The weight of a code is the weight of the smallest nonzero vector in the code. The weight of V for the (6,3) code in Example 30.12 above is 3. So we can take $t = 1$ in Theorem 30.10 to learn this code can correct any single digit error or detect any two errors.

Example 30.13. In this example we use the (6,3) code of Example 30.12.

1. If 111100 is sent with one error as 011100. Then 1111000 is the unique code word at distance 1 from 011100, so 011100 can be corrected to 111100.
2. If 100110 is sent out with two errors as 100011. The error can be detected, because this is not a code word. The error can not be corrected, since 101001 is also distance 2 away.
3. If 110011 is sent with three errors as 011010. The error cannot be detected, since 011010 is also a code word.

The actual decoding for the (6,3) code of Example 30.12 can be done by going down the list, and picking the “closest” code word to the received word. For more complicated codes, there are too many code words do this, so “decoding” becomes more difficult. However, there is a more ”mechanical” way to do decoding for a Hamming code.

Let the binary code (n, k) be determined by the matrix

$$G = \left[I \mid A\right]$$

where I is the $k \times k$ identity matrix, and A is $k \times (n-k)$. We then define the corresponding **parity check matrix** by,

$$H = \begin{bmatrix} -A \\ -- \\ I \end{bmatrix}$$

where $-A$ is $k \times (n-k)$ and I is the $(n-k) \times (n-k)$ identity matrix. This is a general definition referring to the case where $\mathbb{Z}_2$ is replaced by some finite field. In our case of binary codes we always have $A = -A$ and therefore

$$H = \begin{bmatrix} A \\ -- \\ I \end{bmatrix}$$

Example 30.14. For the (6,3) code of Example 30.12,

$$H = \begin{bmatrix} -1 & -1 & 0 \\ -1 & 0 & -1 \\ -1 & -1 & -1 \\ - & - & - \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ - & - & - \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Theorem 30.15. *Suppose the code given by G corrects any single error, and suppose the parity check matrix has nonzero distinct rows. Then H can be used to correct any single error as follows:*

Suppose at most one error is made to the code word $\overline{v} \in V$, where V is the code given by G, and v' is the received word. Then if $v'H = 0$, then v' is a code word and $v' = v$. If $v'H \neq 0$, then $v'H$ is equal to the i^{th} row of H for some i. Then v' has a single error, and the error occurred in the i^{th} coordinate.

Proof. Suppose $v \in V$. If v is a code word, (no error), then $v = wG$ for some row vector $w \in \mathbb{Z}_2^4$. Then $vH = wGH$. But

$$GH = \begin{bmatrix} I & A \end{bmatrix} \begin{bmatrix} -A \\ I \end{bmatrix} = I(-A) + AI = 0$$

by block multiplication. So $GH = 0$, and $wGH = 0$.

If $v' = v$ modulo one error, then $v' = v + e_i$ for some i where

$$e_i = (0, 0, \ldots, 0, 1, 0, \ldots 0, 0) \qquad (1 \text{ in the } i^{\text{th}} \text{ slot}).$$

Then

$$v'H = (v + e_i)H = vH + e_iH = 0 + e_iH.$$

By above e_iH is the i^{th} row of H. Since the rows of H are distinct we can determine this row from v' simply by computing $v'H$. So $v'H$ is the i^{th} -row of H we know that the error occurred at position i and we may correct it by flipping the bit at this position. ■

Of course there are more complicated versions of this theorem to detect and correct more than one error but we will not cover this here.

Example 30.16. Consider the code word, $v = 110011$. Suppose the word is transmitted with at most one error—say the recipient gets $v' = 110011$. The recipient calculates

$$v'H = [1\ 1\ 0\ 0\ 1\ 1] \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [0\ 0\ 0]$$

So Theorem 30.15 says that $v' = v$ and there was no error.

Suppose instead that the receiver gets $v' = 110111$ for which,

$$v'H = [1\ 1\ 0\ 1\ 1\ 1] \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [1\ 0\ 0].$$

The resulting vector is row four of H and therefore the error occurred in position 4. The message was so supposed to be $v = 110011$.

References

1. Richard A. Dean, *Classical abstract algebra*, Harper & Row, Publishers, Inc., New York, 1990.
2. Joseph A. Gallian, *Contemporary abstract algebra*, Houghton Mifflin, Boston, MA, 2002.
3. Anthony W. Knapp, *Basic algebra*, Cornerstones, Birkhäuser Boston Inc., Boston, MA, 2006, Along with a companion volume ıt Advanced algebra. MR MR2257570 (2007e:00001)
4. Joachim von zur Gathen and Jürgen Gerhard, *Modern computer algebra*, second ed., Cambridge University Press, Cambridge, 2003. MR MR2001757 (2004g:68202)