CS 5602 Introduction to Cryptography
Lecture 08
Mathematical Preliminaries 06

George Markowsky
Computer Science Department
Missouri University of Science & Technology

1

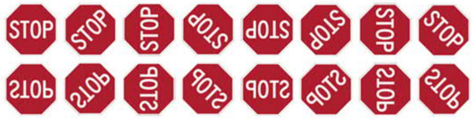
GCD for Polynomials

- The GCD problem that I gave you needs more conditions if you use ℤ[x]
- I will give everyone full credit on the HW and reassign the problem using ℝ[x] instead
- You will find that this makes a lot more sense
- Can also do this for ℂ[x]
- Will come back to discuss the ℤ[x] case if needed

2

Dihedral Groups – Wikipedia

In mathematics, a dihedral group is the group of symmetries of a regular polygon, which includes rotations and reflections. Dihedral groups are among the simplest examples of finite groups, and they play an important role in group theory, geometry, and chemistry.

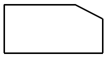


Non-Abelian!

3

Application to Cellphone SIMs

- What is the group of automorphisms of the following shape?



- The group of automorphisms is just {e}
- Why is this significant?
- You can only insert a SIM in one way!

4

Every Group is a Subgroup of a Permutation Group

- Let G be a group, and g ϵ G and consider $f_g(h) = g*h$
- Note that f_g is a bijection because $f^{-1}_g(h)$ can be defined by $g^{-1}*h$
- In particular, $f_g(f^{-1}_g(h)) = f^{-1}_g(f_g(h)) = h$ for all h ϵ G
- It is easy to see that $f_g(f_q(h)) = f_{g*q}$ and that the function $\Gamma: G \rightarrow Bij(G)$ is an injection that preserves group multiplication
- The image, Γ(G) is a subgroup of Bij(G)
- For practical purposes, Bij(G) is often too large to handle
- If S is a set with n elements, Bij(S) is usually denoted by S_n

5

Groups and Reality

- The divisibility properties of the orders of groups often make it easy to prove that there are not any groups of a certain type
- It turns out that groups describe reality and we can often make profound statements about reality based on the properties of groups
- Groups are an essential tool in quantum mechanics
- Obviously, beyond the scope of this course!
- Will give you a little taste of this

6

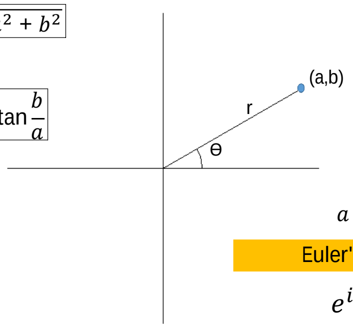
Complex numbers

- The complex numbers, ℂ, form a group under addition and subtraction
- (a + bi) + (c + di) = (a+b) + (c+d)i
- A more interesting structure is the complex numbers minus 0 considered under multiplication
- Note that (a + bi)*(c+di) = (ac – bd) + (ad + bc)i
- If you use the polar representation of a complex number, multiplication has a very interesting and useful representation
- Recall, that |a + bi| = $\sqrt{a^2 + b^2}$, also called abs(a+bi)

7

$r = \sqrt{a^2 + b^2}$

$\theta = \arctan\frac{b}{a}$



$a = r\cos\theta$

$b = r\sin\theta$

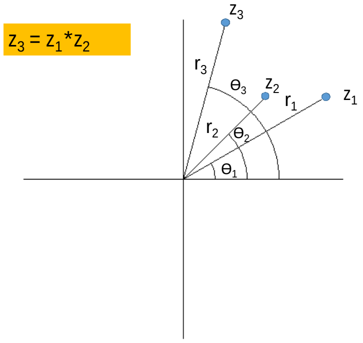
$a + bi = re^{i\theta}$

Euler's Formula

$e^{i\pi}$= -1

8

$z_3 = z_1 * z_2$



$z_1 = r_1e^{i\theta_1}$

$z_2 = r_2e^{i\theta_2}$

$z_3 = r_1r_2e^{i(\theta_1+\theta_2)}$

9

Complex Multiplication

- Thus, multiplying by a + bi is like multiplying by the length of a +bi and rotating by the angle of a + bi
- We are actually interested in complex numbers with magnitude 1 so that nothing is shrunk or expanded by multiplying
- Let U = { a+bi | where |a+bi| = 1 }
- It is easy to see that U is closed under multiplication
- It turns out that the only subgroups of finite order of U all look like { 1, $e^{\pi i/k}$, $e^{2\pi i/k}$, ..., $e^{(2k-1)\pi i/k}$}
- These values are all called roots of unity
- The situation with complex numbers is more complicated

10

2D Traformations that Preserve 0 and Length

- If you ask about bijections f: ℂ → ℂ such that f(0) = 0 and |f(a-b)| = |a-b|, it turns out that there are two types of functions

1. The rotations by a member of U
2. The conjugate mapping $z \rightarrow \bar{z}$, where (a+bi) → (a-bi)
3. The combination of rotations by members of U and the conjugate mapping
4. Note that the conjugate operator does not commute with the conjugate function
5. Let's denote the conjugate map by Conj

11

Non-Commuting Operators

- Conj((i)×(2+i)) = Conj(-1+2i) = -1 – 2i
- i×Conj(2+i) = i(2-i) = 1 + 2i

12

Sylow Theorems

- Theorem 1: For every prime factor p with multiplicity n of the order of a finite group G, there exists a Sylow p-subgroup of G, of order p^n.
- The following weaker version of theorem 1 was first proved by Cauchy, and is known as Cauchy's theorem.
- Corollary: Given a finite group G and a prime number p dividing the order of G, then there exists an element (and hence a subgroup) of order p in G.[1]
- Theorem 2: Given a finite group G and a prime number p, all Sylow p-subgroups of G of the same size are conjugate to each other, i.e. if H and K are Sylow p-subgroups of G, then there exists an element g in G with $g^{-1}Hg = K$.

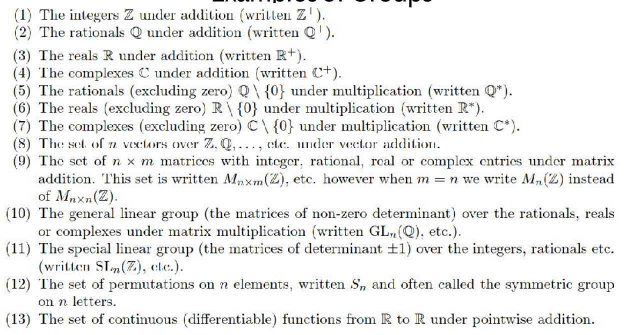
13

Sylow Theorems

- Theorem 3: Let p be a prime factor with multiplicity n of the order of a finite group G, so that the order of G can be written as $p^n m$, where n > 0 and p does not divide m. Let n_p be the number of Sylow p-subgroups of G. Then the following hold:
- n_p divides m, which is the index of the Sylow p-subgroup in G.
- $n_p \equiv 1 \pmod{p}$.
- $n_p = |G : N_G(P)|$, where P is any Sylow p-subgroup of G and N_G denotes the normalizer.

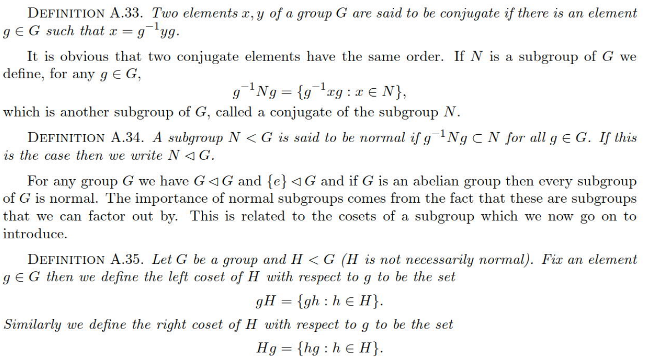
14

Examples of Groups



(1) The integers $\mathbb{Z}$ under addition (written $\mathbb{Z}^{+}$).
(2) The rationals $\mathbb{Q}$ under addition (written $\mathbb{Q}^{+}$).
(3) The reals $\mathbb{R}$ under addition (written $\mathbb{R}^{+}$).
(4) The complexes $\mathbb{C}$ under addition (written $\mathbb{C}^{+}$).
(5) The rationals (excluding zero) $\mathbb{Q} \setminus \{0\}$ under multiplication (written $\mathbb{Q}^{*}$).
(6) The reals (excluding zero) $\mathbb{R} \setminus \{0\}$ under multiplication (written $\mathbb{R}^{*}$).
(7) The complexes (excluding zero) $\mathbb{C} \setminus \{0\}$ under multiplication (written $\mathbb{C}^{*}$).
(8) The set of n vectors over $\mathbb{Z}, \mathbb{Q}, \ldots$, etc. under vector addition.
(9) The set of $n \times m$ matrices with integer, rational, real or complex entries under matrix addition. This set is written $M_{n \times m}(\mathbb{Z})$, etc. however when $m = n$ we write $M_n(\mathbb{Z})$ instead of $M_{n \times n}(\mathbb{Z})$.
(10) The general linear group (the matrices of non-zero determinant) over the rationals, reals or complexes under matrix multiplication (written $\mathrm{GL}_n(\mathbb{Q})$, etc.).
(11) The special linear group (the matrices of determinant ±1) over the integers, rationals etc. (written $\mathrm{SL}_m(\mathbb{Z})$, etc.).
(12) The set of permutations on n elements, written S_n and often called the symmetric group on n letters.
(13) The set of continuous (differentiable) functions from $\mathbb{R}$ to $\mathbb{R}$ under pointwise addition.

15



DEFINITION A.33. *Two elements x, y of a group G are said to be conjugate if there is an element $g \in G$ such that $x = g^{-1}yg$.*

It is obvious that two conjugate elements have the same order. If N is a subgroup of G we define, for any $g \in G$,

$$g^{-1}Ng = \{g^{-1}xg : x \in N\},$$

which is another subgroup of G, called a conjugate of the subgroup N.

DEFINITION A.34. *A subgroup $N < G$ is said to be normal if $g^{-1}Ng \subset N$ for all $g \in G$. If this is the case then we write $N \lhd G$.*

For any group G we have $G \lhd G$ and $\{e\} \lhd G$ and if G is an abelian group then every subgroup of G is normal. The importance of normal subgroups comes from the fact that these are subgroups that we can factor out by. This is related to the cosets of a subgroup which we now go on to introduce.

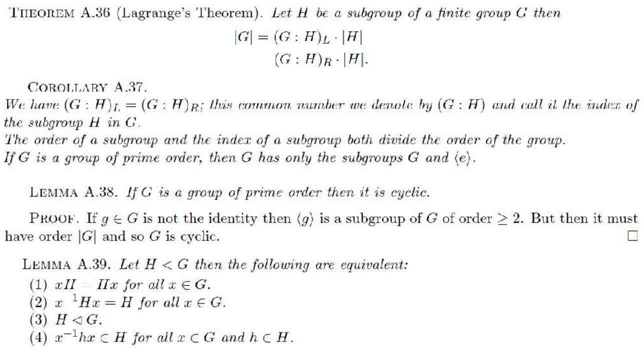
DEFINITION A.35. *Let G be a group and $H < G$ (H is not necessarily normal). Fix an element $g \in G$ then we define the left coset of H with respect to g to be the set*

$$gH = \{gh : h \in H\}.$$

Similarly we define the right coset of H with respect to g to be the set

$$Hg = \{hg : h \in H\}.$$

16



THEOREM A.36 (Lagrange's Theorem). *Let H be a subgroup of a finite group G then*

$$|G| = (G : H)_L \cdot |H|$$
$$(G : H)_R \cdot |H|.$$

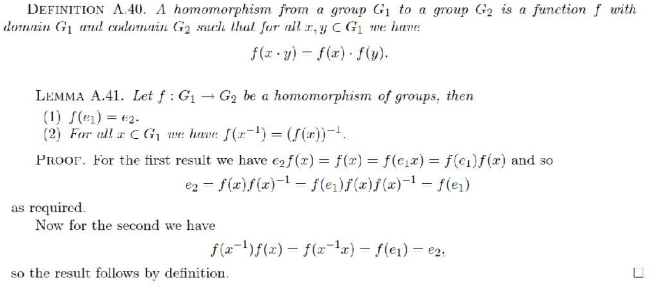
COROLLARY A.37.
We have $(G : H)_L = (G : H)_R$; this common number we denote by $(G : H)$ and call it the index of the subgroup H in G.
The order of a subgroup and the index of a subgroup both divide the order of the group.
If G is a group of prime order, then G has only the subgroups G and $\langle e \rangle$.

LEMMA A.38. *If G is a group of prime order then it is cyclic.*

PROOF. If $g \in G$ is not the identity then $\langle g \rangle$ is a subgroup of G of order ≥ 2. But then it must have order $|G|$ and so G is cyclic. □

LEMMA A.39. *Let $H < G$ then the following are equivalent:*
(1) *$xH = Hx$ for all $x \in G$.*
(2) *$x^{-1}Hx = H$ for all $x \in G$.*
(3) *$H \lhd G$.*
(4) *$x^{-1}hx \subset H$ for all $x \subset G$ and $h \subset H$.*

17



DEFINITION A.40. *A homomorphism from a group G_1 to a group G_2 is a function f with domain G_1 and codomain G_2 such that for all $x, y \in G_1$ we have*

$$f(x \cdot y) = f(x) \cdot f(y).$$

LEMMA A.41. *Let $f : G_1 \to G_2$ be a homomorphism of groups, then*
(1) *$f(e_1) = e_2$.*
(2) *For all $x \in G_1$ we have $f(x^{-1}) = (f(x))^{-1}$.*

PROOF. For the first result we have $e_2 f(x) = f(x) = f(e_1 x) = f(e_1)f(x)$ and so

$$e_2 = f(x)f(x)^{-1} = f(e_1)f(x)f(x)^{-1} = f(e_1)$$

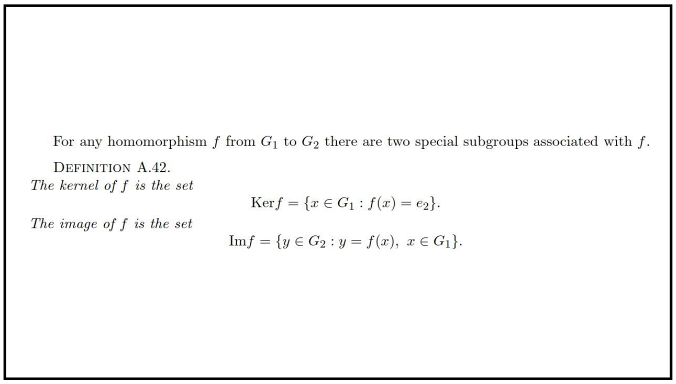
as required.

Now for the second we have

$$f(x^{-1})f(x) = f(x^{-1}x) = f(e_1) = e_2,$$

so the result follows by definition. □

18



For any homomorphism f from G_1 to G_2 there are two special subgroups associated with f.

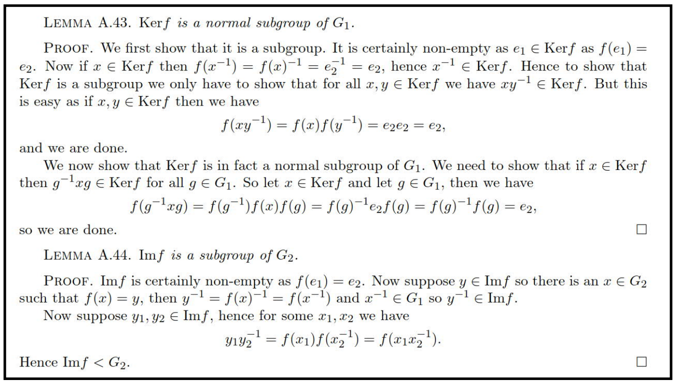
Definition A.42.
The kernel of f is the set

$$\mathrm{Ker} f = \{x \in G_1 : f(x) = e_2\}.$$

The image of f is the set

$$\mathrm{Im} f = \{y \in G_2 : y = f(x),\ x \in G_1\}.$$

19



Lemma A.43. Ker$f$ *is a normal subgroup of G_1.*

Proof. We first show that it is a subgroup. It is certainly non-empty as $e_1 \in \mathrm{Ker} f$ as $f(e_1) = e_2$. Now if $x \in \mathrm{Ker} f$ then $f(x^{-1}) = f(x)^{-1} = e_2^{-1} = e_2$, hence $x^{-1} \in \mathrm{Ker} f$. Hence to show that Ker$f$ is a subgroup we only have to show that for all $x, y \in \mathrm{Ker} f$ we have $xy^{-1} \in \mathrm{Ker} f$. But this is easy as if $x, y \in \mathrm{Ker} f$ then we have

$$f(xy^{-1}) = f(x)f(y^{-1}) = e_2 e_2 = e_2,$$

and we are done.

We now show that Ker$f$ is in fact a normal subgroup of G_1. We need to show that if $x \in \mathrm{Ker} f$ then $g^{-1}xg \in \mathrm{Ker} f$ for all $g \in G_1$. So let $x \in \mathrm{Ker} f$ and let $g \in G_1$, then we have

$$f(g^{-1}xg) = f(g^{-1})f(x)f(g) = f(g)^{-1}e_2 f(g) = f(g)^{-1}f(g) = e_2,$$

so we are done. □

Lemma A.44. Imf *is a subgroup of G_2.*

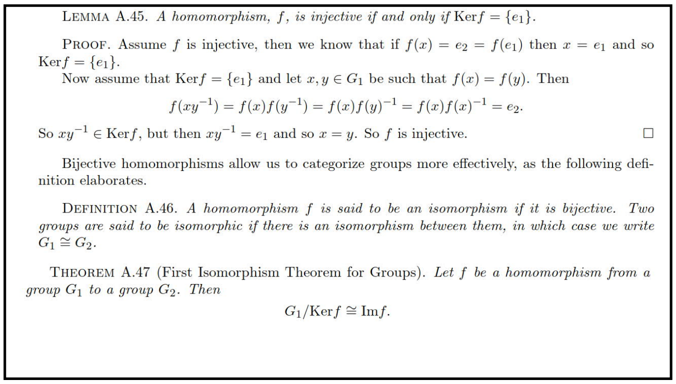
Proof. Imf is certainly non-empty as $f(e_1) = e_2$. Now suppose $y \in \mathrm{Im} f$ so there is an $x \in G_2$ such that $f(x) = y$, then $y^{-1} = f(x)^{-1} = f(x^{-1})$ and $x^{-1} \in G_1$ so $y^{-1} \in \mathrm{Im} f$.

Now suppose $y_1, y_2 \in \mathrm{Im} f$, hence for some x_1, x_2 we have

$$y_1 y_2^{-1} = f(x_1)f(x_2^{-1}) = f(x_1 x_2^{-1}).$$

Hence Im$f < G_2$. □

20



Lemma A.45. *A homomorphism, f, is injective if and only if* Ker$f = \{e_1\}$.

Proof. Assume f is injective, then we know that if $f(x) = e_2 = f(e_1)$ then $x = e_1$ and so Ker$f = \{e_1\}$.

Now assume that Ker$f = \{e_1\}$ and let $x, y \in G_1$ be such that $f(x) = f(y)$. Then

$$f(xy^{-1}) = f(x)f(y^{-1}) = f(x)f(y)^{-1} = f(x)f(x)^{-1} = e_2.$$

So $xy^{-1} \in \mathrm{Ker} f$, but then $xy^{-1} = e_1$ and so $x = y$. So f is injective. □

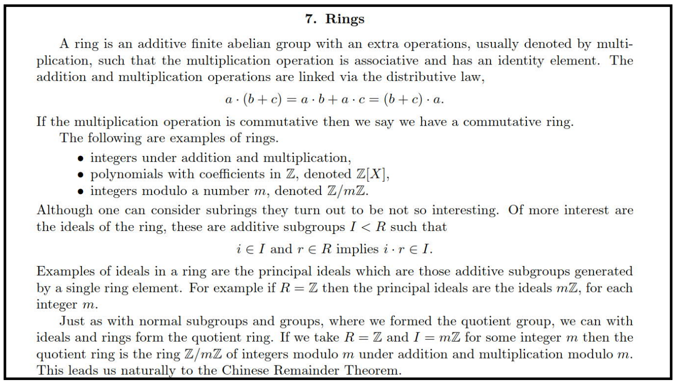
Bijective homomorphisms allow us to categorize groups more effectively, as the following definition elaborates.

Definition A.46. *A homomorphism f is said to be an isomorphism if it is bijective. Two groups are said to be isomorphic if there is an isomorphism between them, in which case we write $G_1 \cong G_2$.*

Theorem A.47 (First Isomorphism Theorem for Groups). *Let f be a homomorphism from a group G_1 to a group G_2. Then*

$$G_1/\mathrm{Ker} f \cong \mathrm{Im} f.$$

21



7. Rings

A ring is an additive finite abelian group with an extra operations, usually denoted by multiplication, such that the multiplication operation is associative and has an identity element. The addition and multiplication operations are linked via the distributive law,

$$a \cdot (b + c) = a \cdot b + a \cdot c = (b + c) \cdot a.$$

If the multiplication operation is commutative then we say we have a commutative ring.

The following are examples of rings.

- integers under addition and multiplication,
- polynomials with coefficients in $\mathbb{Z}$, denoted $\mathbb{Z}[X]$,
- integers modulo a number m, denoted $\mathbb{Z}/m\mathbb{Z}$.

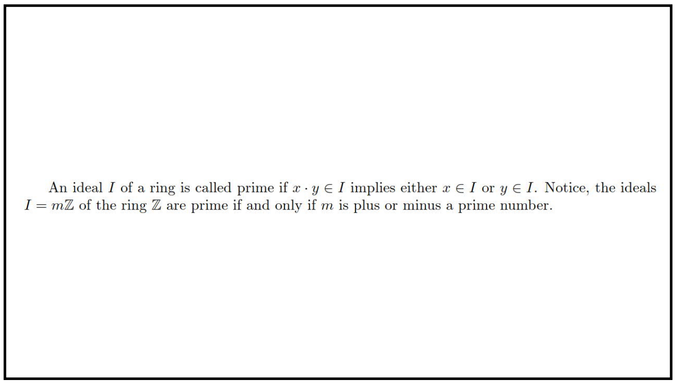
Although one can consider subrings they turn out to be not so interesting. Of more interest are the ideals of the ring, these are additive subgroups $I < R$ such that

$$i \in I \text{ and } r \in R \text{ implies } i \cdot r \in I.$$

Examples of ideals in a ring are the principal ideals which are those additive subgroups generated by a single ring element. For example if $R = \mathbb{Z}$ then the principal ideals are the ideals $m\mathbb{Z}$, for each integer m.

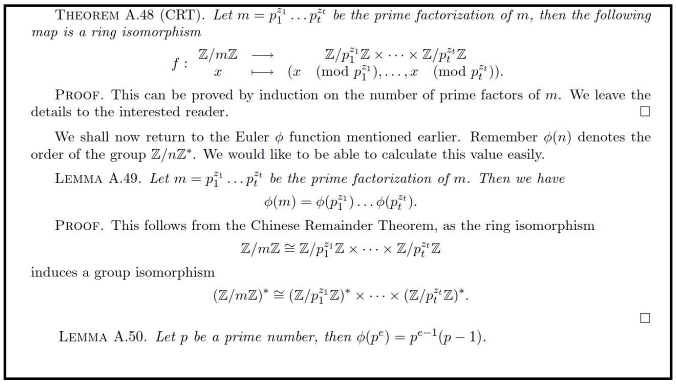
Just as with normal subgroups and groups, where we formed the quotient group, we can with ideals and rings form the quotient ring. If we take $R = \mathbb{Z}$ and $I = m\mathbb{Z}$ for some integer m then the quotient ring is the ring $\mathbb{Z}/m\mathbb{Z}$ of integers modulo m under addition and multiplication modulo m. This leads us naturally to the Chinese Remainder Theorem.

22



An ideal I of a ring is called prime if $x \cdot y \in I$ implies either $x \in I$ or $y \in I$. Notice, the ideals $I = m\mathbb{Z}$ of the ring $\mathbb{Z}$ are prime if and only if m is plus or minus a prime number.

23



Theorem A.48 (CRT). *Let $m = p_1^{z_1} \dots p_t^{z_t}$ be the prime factorization of m, then the following map is a ring isomorphism*

$$f: \begin{array}{ccc} \mathbb{Z}/m\mathbb{Z} & \longrightarrow & \mathbb{Z}/p_1^{z_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_t^{z_t}\mathbb{Z} \\ x & \longmapsto & (x \pmod{p_1^{z_1}}, \dots, x \pmod{p_t^{z_t}}). \end{array}$$

Proof. This can be proved by induction on the number of prime factors of m. We leave the details to the interested reader. □

We shall now return to the Euler ϕ function mentioned earlier. Remember $\phi(n)$ denotes the order of the group $\mathbb{Z}/n\mathbb{Z}^*$. We would like to be able to calculate this value easily.

Lemma A.49. *Let $m = p_1^{z_1} \dots p_t^{z_t}$ be the prime factorization of m. Then we have*

$$\phi(m) = \phi(p_1^{z_1}) \dots \phi(p_t^{z_t}).$$

Proof. This follows from the Chinese Remainder Theorem, as the ring isomorphism

$$\mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}/p_1^{z_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_t^{z_t}\mathbb{Z}$$

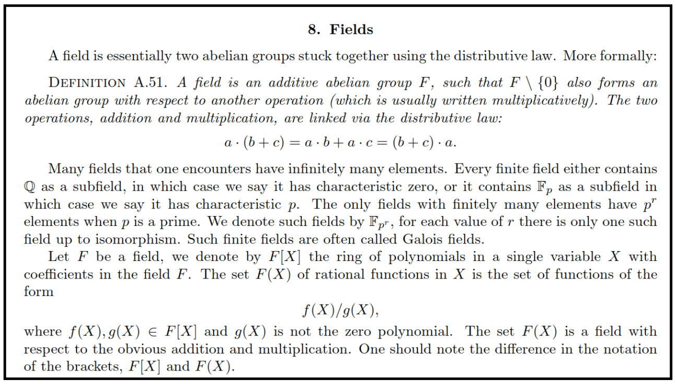
induces a group isomorphism

$$(\mathbb{Z}/m\mathbb{Z})^* \cong (\mathbb{Z}/p_1^{z_1}\mathbb{Z})^* \times \cdots \times (\mathbb{Z}/p_t^{z_t}\mathbb{Z})^*.$$

□

Lemma A.50. *Let p be a prime number, then $\phi(p^e) = p^{e-1}(p-1)$.*

24



8. Fields

A field is essentially two abelian groups stuck together using the distributive law. More formally:

Definition A.51. *A field is an additive abelian group F, such that $F \setminus \{0\}$ also forms an abelian group with respect to another operation (which is usually written multiplicatively). The two operations, addition and multiplication, are linked via the distributive law:*

$$a \cdot (b + c) = a \cdot b + a \cdot c = (b + c) \cdot a.$$

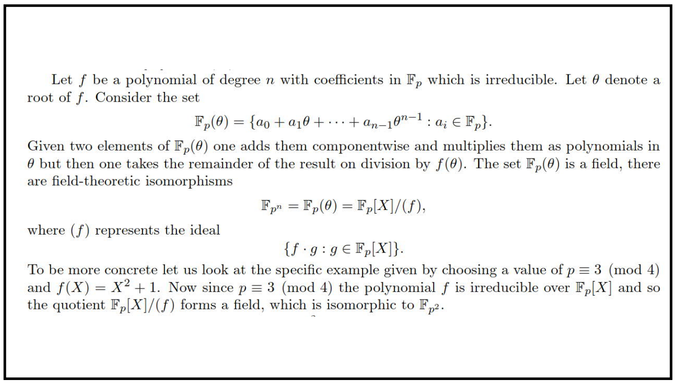
Many fields that one encounters have infinitely many elements. Every finite field either contains $\mathbb{Q}$ as a subfield, in which case we say it has characteristic zero, or it contains $\mathbb{F}_p$ as a subfield in which case we say it has characteristic p. The only fields with finitely many elements have p^r elements when p is a prime. We denote such fields by $\mathbb{F}_{p^r}$, for each value of r there is only one such field up to isomorphism. Such finite fields are often called Galois fields.

Let F be a field, we denote by $F[X]$ the ring of polynomials in a single variable X with coefficients in the field F. The set $F(X)$ of rational functions in X is the set of functions of the form

$$f(X)/g(X),$$

where $f(X), g(X) \in F[X]$ and $g(X)$ is not the zero polynomial. The set $F(X)$ is a field with respect to the obvious addition and multiplication. One should note the difference in the notation of the brackets, $F[X]$ and $F(X)$.

25



Let f be a polynomial of degree n with coefficients in $\mathbb{F}_p$ which is irreducible. Let θ denote a root of f. Consider the set

$$\mathbb{F}_p(\theta) = \{a_0 + a_1\theta + \cdots + a_{n-1}\theta^{n-1} : a_i \in \mathbb{F}_p\}.$$

Given two elements of $\mathbb{F}_p(\theta)$ one adds them componentwise and multiplies them as polynomials in θ but then one takes the remainder of the result on division by $f(\theta)$. The set $\mathbb{F}_p(\theta)$ is a field, there are field-theoretic isomorphisms

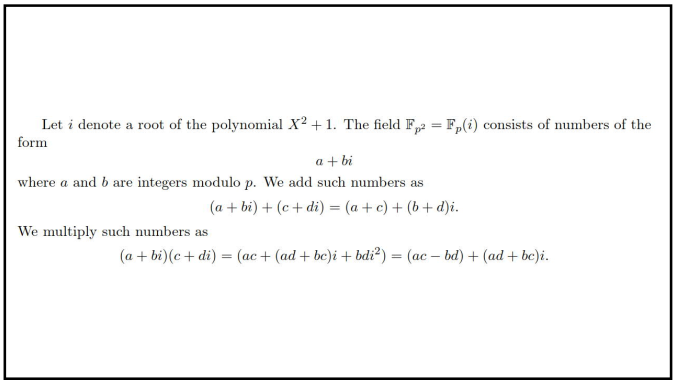
$$\mathbb{F}_{p^n} = \mathbb{F}_p(\theta) = \mathbb{F}_p[X]/(f),$$

where (f) represents the ideal

$$\{f \cdot g : g \in \mathbb{F}_p[X]\}.$$

To be more concrete let us look at the specific example given by choosing a value of $p \equiv 3 \pmod 4$ and $f(X) = X^2 + 1$. Now since $p \equiv 3 \pmod 4$ the polynomial f is irreducible over $\mathbb{F}_p[X]$ and so the quotient $\mathbb{F}_p[X]/(f)$ forms a field, which is isomorphic to $\mathbb{F}_{p^2}$.

26



Let i denote a root of the polynomial $X^2 + 1$. The field $\mathbb{F}_{p^2} = \mathbb{F}_p(i)$ consists of numbers of the form

$$a + bi$$

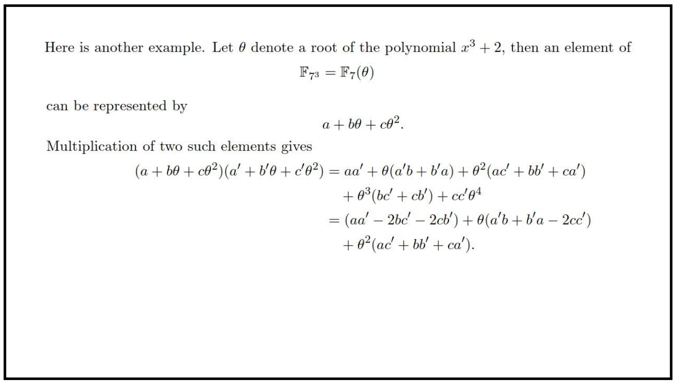
where a and b are integers modulo p. We add such numbers as

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

We multiply such numbers as

$$(a + bi)(c + di) = (ac + (ad + bc)i + bdi^2) = (ac - bd) + (ad + bc)i.$$

27



Here is another example. Let θ denote a root of the polynomial $x^3 + 2$, then an element of

$$\mathbb{F}_{7^3} = \mathbb{F}_7(\theta)$$

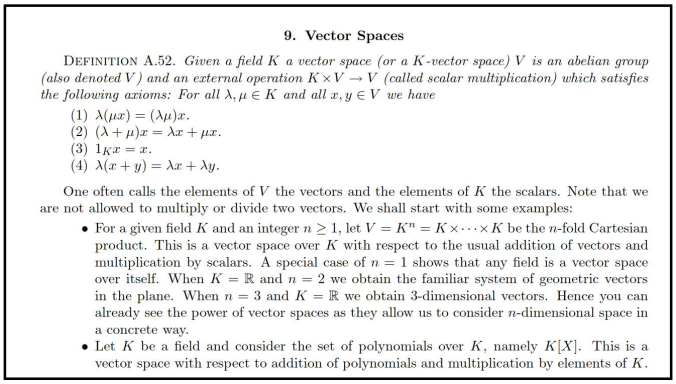
can be represented by

$$a + b\theta + c\theta^2.$$

Multiplication of two such elements gives

$$\begin{aligned}(a + b\theta + c\theta^2)(a' + b'\theta + c'\theta^2) &= aa' + \theta(a'b + b'a) + \theta^2(ac' + bb' + ca') \\ &\quad + \theta^3(bc' + cb') + cc'\theta^4 \\ &= (aa' - 2bc' - 2cb') + \theta(a'b + b'a - 2cc') \\ &\quad + \theta^2(ac' + bb' + ca').\end{aligned}$$

28



9. Vector Spaces

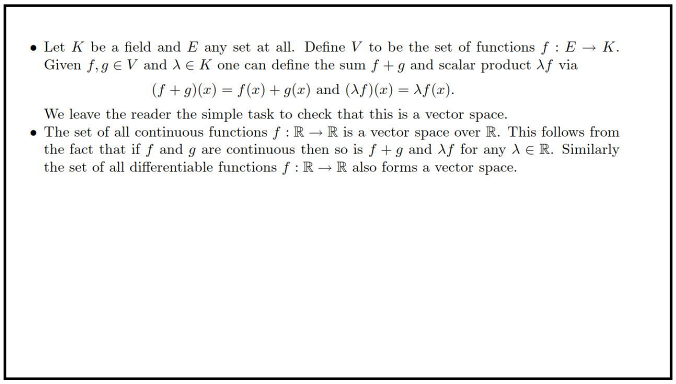
Definition A.52. *Given a field K a vector space (or a K-vector space) V is an abelian group (also denoted V) and an external operation $K \times V \to V$ (called scalar multiplication) which satisfies the following axioms: For all $\lambda, \mu \in K$ and all $x, y \in V$ we have*

(1) $\lambda(\mu x) = (\lambda\mu)x$.
(2) $(\lambda + \mu)x = \lambda x + \mu x$.
(3) $1_K x = x$.
(4) $\lambda(x + y) = \lambda x + \lambda y$.

One often calls the elements of V the vectors and the elements of K the scalars. Note that we are not allowed to multiply or divide two vectors. We shall start with some examples:

- For a given field K and an integer $n \geq 1$, let $V = K^n = K \times \cdots \times K$ be the n-fold Cartesian product. This is a vector space over K with respect to the usual addition of vectors and multiplication by scalars. A special case of $n = 1$ shows that any field is a vector space over itself. When $K = \mathbb{R}$ and $n = 2$ we obtain the familiar system of geometric vectors in the plane. When $n = 3$ and $K = \mathbb{R}$ we obtain 3-dimensional vectors. Hence you can already see the power of vector spaces as they allow us to consider n-dimensional space in a concrete way.
- Let K be a field and consider the set of polynomials over K, namely $K[X]$. This is a vector space with respect to addition of polynomials and multiplication by elements of K.

29



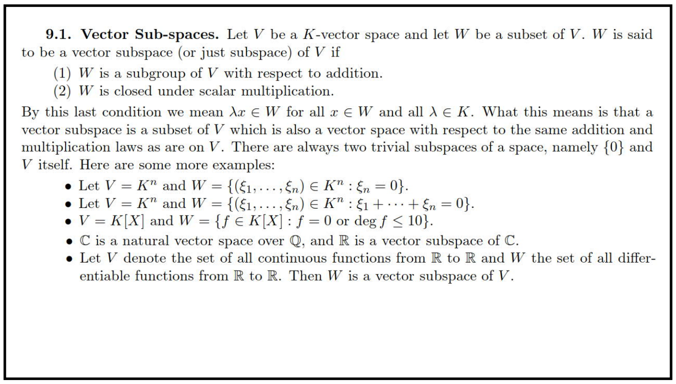
- Let K be a field and E any set at all. Define V to be the set of functions $f : E \to K$. Given $f, g \in V$ and $\lambda \in K$ one can define the sum $f + g$ and scalar product λf via

$$(f + g)(x) = f(x) + g(x) \text{ and } (\lambda f)(x) = \lambda f(x).$$

We leave the reader the simple task to check that this is a vector space.

- The set of all continuous functions $f : \mathbb{R} \to \mathbb{R}$ is a vector space over $\mathbb{R}$. This follows from the fact that if f and g are continuous then so is $f + g$ and λf for any $\lambda \in \mathbb{R}$. Similarly the set of all differentiable functions $f : \mathbb{R} \to \mathbb{R}$ also forms a vector space.

30



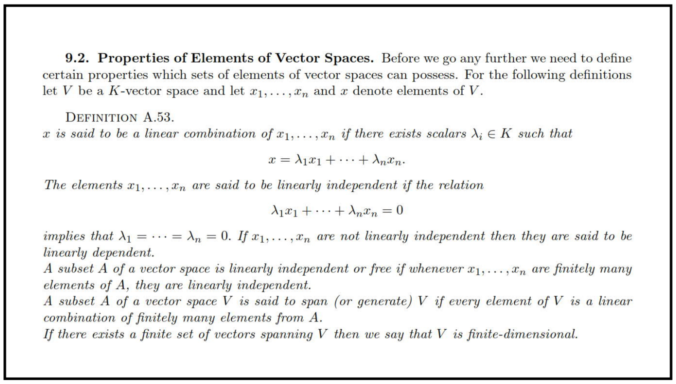
9.1. Vector Sub-spaces. Let V be a K-vector space and let W be a subset of V. W is said to be a vector subspace (or just subspace) of V if

(1) W is a subgroup of V with respect to addition.
(2) W is closed under scalar multiplication.

By this last condition we mean $\lambda x \in W$ for all $x \in W$ and all $\lambda \in K$. What this means is that a vector subspace is a subset of V which is also a vector space with respect to the same addition and multiplication laws as are on V. There are always two trivial subspaces of a space, namely $\{0\}$ and V itself. Here are some more examples:

- Let $V = K^n$ and $W = \{(\xi_1, \ldots, \xi_n) \in K^n : \xi_n = 0\}$.
- Let $V = K^n$ and $W = \{(\xi_1, \ldots, \xi_n) \in K^n : \xi_1 + \cdots + \xi_n = 0\}$.
- $V = K[X]$ and $W = \{f \in K[X] : f = 0 \text{ or } \deg f \leq 10\}$.
- $\mathbb{C}$ is a natural vector space over $\mathbb{Q}$, and $\mathbb{R}$ is a vector subspace of $\mathbb{C}$.
- Let V denote the set of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$ and W the set of all differentiable functions from $\mathbb{R}$ to $\mathbb{R}$. Then W is a vector subspace of V.

31



9.2. Properties of Elements of Vector Spaces. Before we go any further we need to define certain properties which sets of elements of vector spaces can possess. For the following definitions let V be a K-vector space and let $x_1, \ldots, x_n$ and x denote elements of V.

DEFINITION A.53.
x is said to be a linear combination of $x_1, \ldots, x_n$ if there exists scalars $\lambda_i \in K$ such that

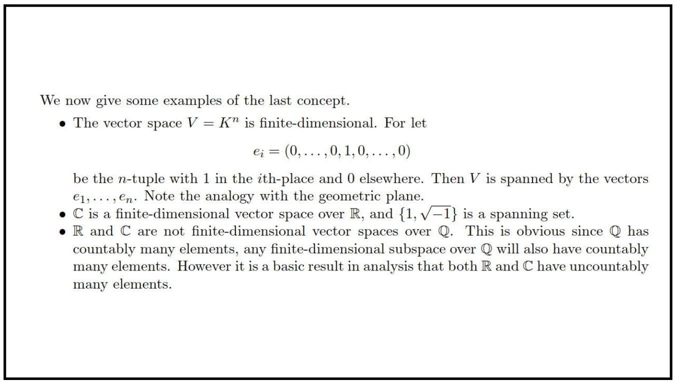
$$x = \lambda_1 x_1 + \cdots + \lambda_n x_n.$$

The elements $x_1, \ldots, x_n$ are said to be linearly independent if the relation

$$\lambda_1 x_1 + \cdots + \lambda_n x_n = 0$$

implies that $\lambda_1 = \cdots = \lambda_n = 0$. If $x_1, \ldots, x_n$ are not linearly independent then they are said to be linearly dependent.
A subset A of a vector space is linearly independent or free if whenever $x_1, \ldots, x_n$ are finitely many elements of A, they are linearly independent.
A subset A of a vector space V is said to span (or generate) V if every element of V is a linear combination of finitely many elements from A.
If there exists a finite set of vectors spanning V then we say that V is finite-dimensional.

32



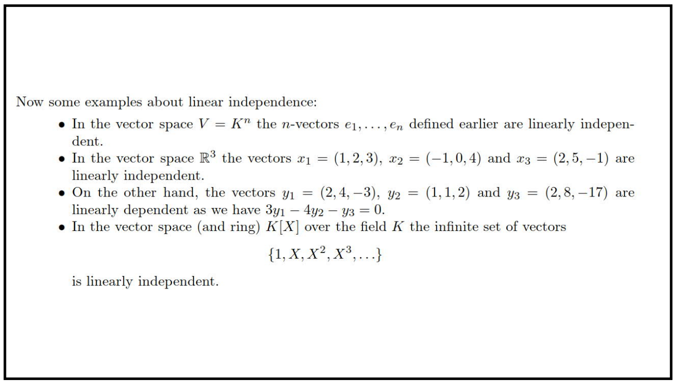
We now give some examples of the last concept.

- The vector space $V = K^n$ is finite-dimensional. For let

$$e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$$

  be the n-tuple with 1 in the ith-place and 0 elsewhere. Then V is spanned by the vectors $e_1, \ldots, e_n$. Note the analogy with the geometric plane.
- $\mathbb{C}$ is a finite-dimensional vector space over $\mathbb{R}$, and $\{1, \sqrt{-1}\}$ is a spanning set.
- $\mathbb{R}$ and $\mathbb{C}$ are not finite-dimensional vector spaces over $\mathbb{Q}$. This is obvious since $\mathbb{Q}$ has countably many elements, any finite-dimensional subspace over $\mathbb{Q}$ will also have countably many elements. However it is a basic result in analysis that both $\mathbb{R}$ and $\mathbb{C}$ have uncountably many elements.

33



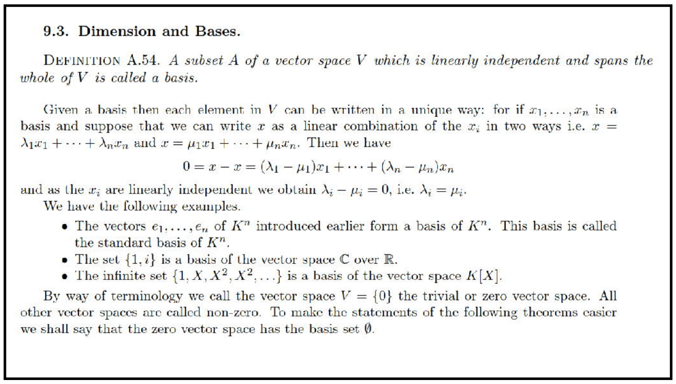
Now some examples about linear independence:

- In the vector space $V = K^n$ the n-vectors $e_1, \ldots, e_n$ defined earlier are linearly independent.
- In the vector space $\mathbb{R}^3$ the vectors $x_1 = (1, 2, 3)$, $x_2 = (-1, 0, 4)$ and $x_3 = (2, 5, -1)$ are linearly independent.
- On the other hand, the vectors $y_1 = (2, 4, -3)$, $y_2 = (1, 1, 2)$ and $y_3 = (2, 8, -17)$ are linearly dependent as we have $3y_1 - 4y_2 - y_3 = 0$.
- In the vector space (and ring) $K[X]$ over the field K the infinite set of vectors

$$\{1, X, X^2, X^3, \ldots\}$$

  is linearly independent.

34



9.3. Dimension and Bases.

DEFINITION A.54. *A subset A of a vector space V which is linearly independent and spans the whole of V is called a basis.*

Given a basis then each element in V can be written in a unique way: for if $x_1, \ldots, x_n$ is a basis and suppose that we can write x as a linear combination of the x_i in two ways i.e. $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$ and $x = \mu_1 x_1 + \cdots + \mu_n x_n$. Then we have

$$0 = x - x = (\lambda_1 - \mu_1)x_1 + \cdots + (\lambda_n - \mu_n)x_n$$

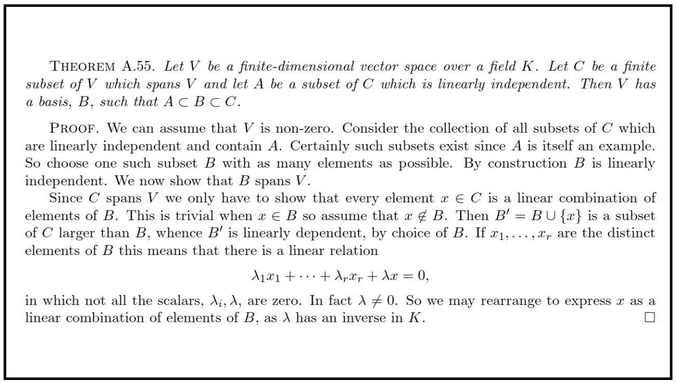
and as the x_i are linearly independent we obtain $\lambda_i - \mu_i = 0$, i.e. $\lambda_i = \mu_i$.

We have the following examples.

- The vectors $e_1, \ldots, e_n$ of K^n introduced earlier form a basis of K^n. This basis is called the standard basis of K^n.
- The set $\{1, i\}$ is a basis of the vector space $\mathbb{C}$ over $\mathbb{R}$.
- The infinite set $\{1, X, X^2, X^2, \ldots\}$ is a basis of the vector space $K[X]$.

By way of terminology we call the vector space $V = \{0\}$ the trivial or zero vector space. All other vector spaces are called non-zero. To make the statements of the following theorems easier we shall say that the zero vector space has the basis set $\emptyset$.

35



THEOREM A.55. *Let V be a finite-dimensional vector space over a field K. Let C be a finite subset of V which spans V and let A be a subset of C which is linearly independent. Then V has a basis, B, such that $A \subset B \subset C$.*

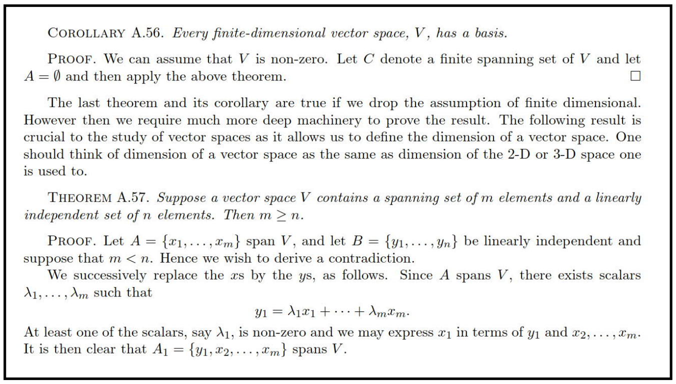
PROOF. We can assume that V is non-zero. Consider the collection of all subsets of C which are linearly independent and contain A. Certainly such subsets exist since A is itself an example. So choose one such subset B with as many elements as possible. By construction B is linearly independent. We now show that B spans V.

Since C spans V we only have to show that every element $x \in C$ is a linear combination of elements of B. This is trivial when $x \in B$ so assume that $x \notin B$. Then $B' = B \cup \{x\}$ is a subset of C larger than B, whence B' is linearly dependent, by choice of B. If $x_1, \ldots, x_r$ are the distinct elements of B this means that there is a linear relation

$$\lambda_1 x_1 + \cdots + \lambda_r x_r + \lambda x = 0,$$

in which not all the scalars, λ_i, λ, are zero. In fact $\lambda \neq 0$. So we may rearrange to express x as a linear combination of elements of B, as λ has an inverse in K. □

36



Corollary A.56. *Every finite-dimensional vector space, V, has a basis.*

Proof. We can assume that V is non-zero. Let C denote a finite spanning set of V and let $A = \emptyset$ and then apply the above theorem. □

The last theorem and its corollary are true if we drop the assumption of finite dimensional. However then we require much more deep machinery to prove the result. The following result is crucial to the study of vector spaces as it allows us to define the dimension of a vector space. One should think of dimension of a vector space as the same as dimension of the 2-D or 3-D space one is used to.

Theorem A.57. *Suppose a vector space V contains a spanning set of m elements and a linearly independent set of n elements. Then $m \geq n$.*

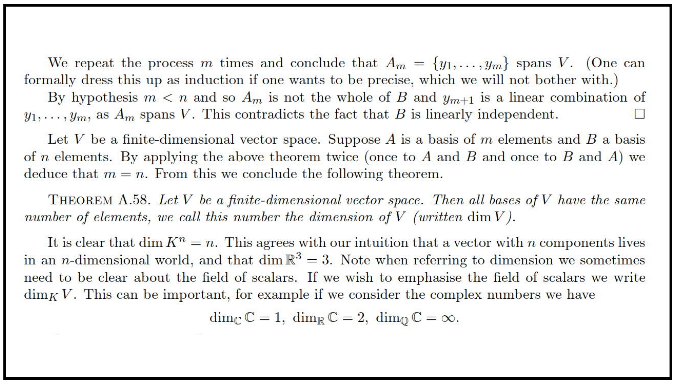
Proof. Let $A = \{x_1, \dots, x_m\}$ span V, and let $B = \{y_1, \dots, y_n\}$ be linearly independent and suppose that $m < n$. Hence we wish to derive a contradiction.

We successively replace the xs by the ys, as follows. Since A spans V, there exists scalars $\lambda_1, \dots, \lambda_m$ such that

$$y_1 = \lambda_1 x_1 + \cdots + \lambda_m x_m.$$

At least one of the scalars, say λ_1, is non-zero and we may express x_1 in terms of y_1 and $x_2, \dots, x_m$. It is then clear that $A_1 = \{y_1, x_2, \dots, x_m\}$ spans V.

37



We repeat the process m times and conclude that $A_m = \{y_1, \dots, y_m\}$ spans V. (One can formally dress this up as induction if one wants to be precise, which we will not bother with.)

By hypothesis $m < n$ and so A_m is not the whole of B and y_{m+1} is a linear combination of $y_1, \dots, y_m$, as A_m spans V. This contradicts the fact that B is linearly independent. □

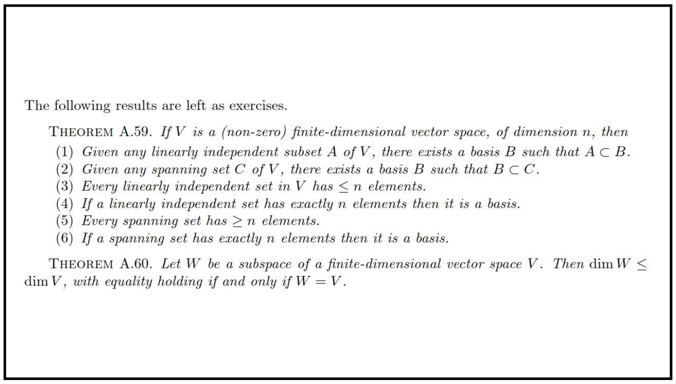
Let V be a finite-dimensional vector space. Suppose A is a basis of m elements and B a basis of n elements. By applying the above theorem twice (once to A and B and once to B and A) we deduce that $m = n$. From this we conclude the following theorem.

Theorem A.58. *Let V be a finite-dimensional vector space. Then all bases of V have the same number of elements, we call this number the dimension of V (written $\dim V$).*

It is clear that $\dim K^n = n$. This agrees with our intuition that a vector with n components lives in an n-dimensional world, and that $\dim \mathbb{R}^3 = 3$. Note when referring to dimension we sometimes need to be clear about the field of scalars. If we wish to emphasise the field of scalars we write $\dim_K V$. This can be important, for example if we consider the complex numbers we have

$$\dim_{\mathbb{C}} \mathbb{C} = 1, \ \dim_{\mathbb{R}} \mathbb{C} = 2, \ \dim_{\mathbb{Q}} \mathbb{C} = \infty.$$

38



The following results are left as exercises.

Theorem A.59. *If V is a (non-zero) finite-dimensional vector space, of dimension n, then*

(1) *Given any linearly independent subset A of V, there exists a basis B such that $A \subset B$.*
(2) *Given any spanning set C of V, there exists a basis B such that $B \subset C$.*
(3) *Every linearly independent set in V has $\leq n$ elements.*
(4) *If a linearly independent set has exactly n elements then it is a basis.*
(5) *Every spanning set has $\geq n$ elements.*
(6) *If a spanning set has exactly n elements then it is a basis.*

Theorem A.60. *Let W be a subspace of a finite-dimensional vector space V. Then $\dim W \leq \dim V$, with equality holding if and only if $W = V$.*

39