Revision Sheet - Groups and Vector Spaces

January 15, 2020

Contents

1 Definitions.

1.1 Groups and Subgroups

1.1 Definition. Fix an integer $n \geq 1$. We let $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$, and we add and multiply members of $\mathbb{Z}_n$ 'modulo' n. That is, we add or multiply two given members of $\mathbb{Z}_n$ as usual, and then find the remainder of the answer on division by n. This is called the *ring of integers modulo* n.

1.2 Definition. A *group* is a non-empty set G on which is defined an associative binary operation $\circ$ such that there is an identity e ($e \circ x = x$ and $x \circ e = x$ for all $x \in G$), and each $x \in G$ has an inverse in G (an element y such that $x \circ y = e$ and $y \circ x = e$).

1.3 Definition. We say that a group $(G, \circ)$ is *abelian* if the operation $\circ$ is commutative, that is, $x \circ y = y \circ x$ for all $x, y \in G$.

1.4 Definition. Given an element x of a group G, and a positive integer n, we define the power $x^n \in G$ by
$$x^n = \underbrace{xx\dots x}_{n \text{ copies}} \in G.$$
We also define $x^0 = 1$ and negative powers by $x^{-n} = (x^n)^{-1}$. For an additive group we use the alternative notation $nx = x + x + \cdots + x$, $0x = 0$, $(-n)x = -(nx)$.

1.5 Definition. We say that elements x, y in a group G *commute* if $xy = yx$.

1.6 Definition. Let $(G, \circ)$ be a group. A *subgroup* of $(G, \circ)$ is a subset H of G such that H becomes a group with the same operation $\circ$.

1.7 Definition. The *order* of a group G, denoted by $|G|$, is the number of elements in the set G, either a positive integer or infinity.

1.8 Definition. The *order* of an element x of a group G is the smallest integer $n > 0$ such that $x^n = 1$. If no such n exists we say that x has infinite order. (In an additive group the condition is $nx = 0$.)

1.9 Definition. If x is an element of a group G we let
$$\langle x \rangle = \{x^n : n \in \mathbb{Z}\}.$$
(or in additive notation $\langle x \rangle = \{nx : n \in \mathbb{Z}\}$). It is a subgroup of G. We call it the *subgroup of G generated by x*. We say that G is *generated by x*, or that x is a *generator* for G if $G = \langle x \rangle$. We say that G is a *cyclic* group if it has a generator.

1.10 Definition. If G and H are groups, then we consider the cartesian product
$$G \times H = \{(g, h) : g \in G, h \in H\}$$
with the operation $\circ$ defined by
$$(g, h) \circ (g', h') = (gg', hh').$$
It is easy to see that it is a group. We call it the *direct product* of G and H. The identity element is $1 = (1_G, 1_H)$. The inverse of (g, h) is (g^{-1}, h^{-1}). (If G and H are additive groups we use the notation $(g, h) + (g', h') = (g + g', h + h')$.)

1.2 Homomorphisms, Isomorphisms, and Permutations

1.11 Definition. Let $(G, \circ)$ and $(H, \circ)$ be groups. A mapping $\theta : G \to H$ is a *homomorphism* if $\theta(g \circ g') = \theta(g) \circ \theta(g')$ for all $g, g' \in G$. It is an *isomorphism* if in addition it is a bijection. We say that groups G and H are *isomorphic*, and write $G \cong H$, if there is an isomorphism $\theta : G \to H$.

1.12 Definition. Let H be a subgroup of a group G. A *(right) coset* of H in G is a subset of the form

$$Hx = \{hx : h \in H\}$$

for some $x \in G$. If G is an additive group we use the notation $H + x = \{h + x : h \in H\}$ instead. Note that even if G is infinite, we still have the notion of 'right coset'. Finiteness is just used in the final part of the proof of Lagrange's Theorem.

1.13 Definition. If H is a subgroup of a finite group G, the *index* of H in G is the number of different cosets of H in G. We denote it by $|G : H|$.

1.14 Definition. A *permutation* of a set A is a bijective mapping from A to itself, $\pi : A \to A$. The set of all permutations of A forms a group under composition of mappings $\pi \circ \sigma$, where

$$(\pi \circ \sigma)(a) = \pi(\sigma(a))$$

for $a \in A$. The identity element is the identity map *id*. Since π is bijective, it has an inverse mapping π^{-1}, and that is the inverse to π in this group. We shall only be interested in permutations of the set $A = \{1, 2, \ldots, n\}$ for n a positive integer. The set of all such permutations is called the *symmetric group of degree* n and denoted by S_n.

1.15 Definition. Let k, n be a positive integers with $k \leq n$ and let $a_1, a_2, \ldots, a_k$ be distinct elements in the set $\{1, 2, \ldots, n\}$. We denote by $(a_1\ a_2\ \ldots\ a_k)$ the permutation in S_n sending

$$a_1 \mapsto a_2 \mapsto a_3 \mapsto \cdots \mapsto a_k \mapsto a_1$$

and with $a \mapsto a$ for all a not in the list. It is called a *cycle of length* k or a k*-cycle*. A 2-cycle is also called a *transposition*.

1.16 Definition. Given a permutation $\pi \in S_n$, the corresponding *permutation matrix* is the $n \times n$ matrix A_π whose jth column is $\mathbf{e}_{\pi(j)}$, for all j. Equivalently $A_\pi \mathbf{e}_j = \mathbf{e}_{\pi(j)}$. Explicitly $A_\pi = (a_{ij})$ where

$$a_{ij} = \begin{cases} 1 & (\text{if } i = \pi(j)) \\ 0 & (\text{otherwise}) \end{cases}$$

1.17 Definition. The *sign* or *signature* of a permutation π is $\epsilon(\pi) = \det(A_\pi)$.

1.18 Definition. A permutation which can be written as a product of an odd/even number of transpositions is called an *odd/even permutation*.

1.19 Definition. The set of even permutations in S_n (which forms a subgroup of S_n) is called the *alternating group* A_n *of degree* n.

1.20 Definition. The *kernel* of a homomorphism $\theta : G \to H$ is the set $\ker \theta = \{g \in G : \theta(g) = 1\}$. It is a subset of G. The *image* of a homomorphism $\theta : G \to H$ is the set $\operatorname{im} \theta = \{\theta(g) : g \in G\}$. It is a subset of H.

1.3 Conjugacy, Normal Subgroups

1.21 Definition. Elements x, y of a group G are said to be *conjugate* in G if there is $g \in G$ with $y = g^{-1}xg$. The set of all elements conjugate to a given element x is called a *conjugacy class*. The conjugacy class containing x is

$$\mathrm{conj}_G(x) = \{g^{-1}xg : g \in G\}.$$

1.22 Definition. If x is an element of a group G, the *centralizer* of x in G is the set $C_G(x) = \{g \in G : gx = xg\}$. It is easy to see that it is a subgroup of G.

1.23 Definition. A subgroup H of a group G is said to be *normal* if $g^{-1}hg \in H$ for all $h \in H$ and $g \in G$. It is equivalent that H is a union of conjugacy classes. We denote this by $H \lhd G$.

1.24 Definition. If H is a normal subgroup of G, then we denote by G/H the set of cosets of H in G, and we equip it with the multiplication defined by $(Hg)(Hg') = H(gg')$. The lemma shows that this is well-defined. It turns G/H into a group, called the *quotient group* of G by H. The map $\theta : G \to G/H$, $\theta(g) = Hg$ is a homomorphism.

1.25 Definition. A group G is *simple* if it has no non-trivial proper normal subgroups. That is, if the only normal subgroups are $\{1\}$ and G.

1.4 Fields and Vector Spaces

1.26 Definition. A *field* consists of a set F with binary operations $+$ and $\cdot$ satisfying (i) The operation $+$ turns F into an additive group. The identity element is denoted by 0. (ii) The product $a \cdot b$ is defined and in F for all $a, b \in F$, it is associative and commutative, and it turns $F^* = \{x \in F : x \neq 0\}$ into an abelian group. (iii) The product $\cdot$ is distributive over $+$, that is, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

1.27 Definition. Let F be a field. A *vector space over* F, or *an* F*-vector space* consists of a set V, whose elements are called *vectors*, together with operations of addition of vectors, $+$, and scalar multiplication satisfying the following axioms. (addition) The set V of vectors is an additive group under $+$. (closure) Scalar multiplication $a\mathbf{v}$ is defined and in V for all scalars $a \in F$ and $\mathbf{v} \in V$. (compatibility of multiplication) $(ab)\mathbf{v} = a(b\mathbf{v})$ for all $a, b \in F$ and $\mathbf{v} \in V$. (identity) $1\mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in V$. (distributivity) $a(\mathbf{v} + \mathbf{w}) = (a\mathbf{v}) + (a\mathbf{w})$ for all $a \in F$ and $\mathbf{v}, \mathbf{w} \in V$. $(a + b)\mathbf{v} = (a\mathbf{v}) + (b\mathbf{v})$ for all $a + b \in F$ and $\mathbf{v} \in V$. We denote by $\mathbf{0}$ the identity element for V under $+$. The zero vector. We can define subtraction for vectors by defining $\mathbf{u} - \mathbf{v}$ to be equal to $\mathbf{u} + (-\mathbf{v})$.

1.28 Definition. Let V be a vector space over a field F. By a *subspace* of V we mean a subset U of V such that U becomes a vector space with the same operations of addition of vectors and scalar multiplication in V.

1.5 Linear Mappings, Basis Vectors,

1.29 Definition. Let V, W be vector spaces over a field F. A mapping $\theta : V \to W$ is called a *linear mapping* (or *linear transformation*, *linear operator*, or *homomorphism of vector spaces*) if (i) $\theta(\mathbf{v}+\mathbf{v}') = \theta(\mathbf{v})+\theta(\mathbf{v}')$ for all $\mathbf{v}, \mathbf{v}' \in V$, and (ii) $\theta(a\mathbf{v}) = a\theta(\mathbf{v})$ for all $a \in F$ and $\mathbf{v} \in V$. (It follows that $\theta(a\mathbf{v}+b\mathbf{v}') = a\theta(\mathbf{v})+b\theta(\mathbf{v}')$ for all $a, b \in F$ and $\mathbf{v}, \mathbf{v}' \in V$. In fact this can be used as a characterization of linear mappings.) An *isomorphism of vector spaces* is a linear map which is a bijection. If so, we write $V \cong W$.

1.30 Definition. The *span* of a finite set of vectors $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ in a vector space V is the set of all linear combinations of them,

$$\operatorname{span} S = \{a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n : a_1, \dots, a_n \in F\}.$$

1.31 Definition. Let V be a vector space and let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a finite subset of V. We say that S is *linearly independent* if there is no linear relation between the elements of S of the form

$$a_1\mathbf{v}_1 + \ldots + a_n\mathbf{v}_n = \mathbf{0}$$

with $a_1, \dots, a_n \in F$, other than the trivial one with $a_1 = \ldots = a_n = 0$. Otherwise S is said to be *linearly dependent*.

1.32 Definition. Let V be a vector space. We say that a finite set of vectors $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a *basis* of V if it is linearly independent and it spans V (i.e. $\operatorname{span} S = V$).

1.33 Definition. If a vector space V has a basis with n elements, then we say that V has *dimension* n. We call V *finite-dimensional* in this case, and write $\dim V = n$. If V does not have a (finite) basis, then it is said to be *infinite-dimensional*.

1.34 Definition. Let V be a vector space over F. If U is a subspace of V, then the *quotient vector space* V/U is the quotient group under addition, with scalar multiplication defined by $a(U+\mathbf{v}) = U + a\mathbf{v}$. It is easy to see that the natural map $V \to V/U$, $\mathbf{v} \mapsto U + \mathbf{v}$ is a linear map.

1.35 Definition. If $\theta : V \to W$ is a linear map, then the *rank* of θ is $r(\theta) = \dim \operatorname{im} \theta$ and the *nullity* of θ is $n(\theta) = \dim \ker \theta$.

1.36 Definition. Suppose that $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis of a vector space V over F. In this case the map $\phi_S : F^n \to V$ is an isomorphism. Thus for each $\mathbf{v} \in V$ there is a unique vector $\mathbf{x} = (x_1, \dots, x_n)^T \in F^n$ such that $\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$. We call it the *coordinates of* $\mathbf{v}$ *with respect to* S, and denote it by $[\mathbf{v}]_S$.

1.6 Matrices of Linear Mappings

1.37 Definition. Let $\theta : V \to W$ be a linear map, let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V and let $R = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ be a basis of W. The *matrix of* θ *with respect to the* *basis S of V and the basis T of W* is the matrix $A = (a_{ij})$ whose jth column is the coordinates of $\theta(\mathbf{v}_j)$ with respect to R.

Thus

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

where

$$\begin{array}{rcl} \theta(\mathbf{v}_1) & = & a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m \\ \theta(\mathbf{v}_2) & = & a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{m2}\mathbf{w}_m \\ \dots & & \\ \theta(\mathbf{v}_n) & = & a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m. \end{array}$$

or $\theta(\mathbf{v}_j) = \sum_{i=1}^n a_{ij}\mathbf{w}_i$.

Special case. If $\theta : V \to V$ is a linear map from a vector space to itself, and we use the same basis for both the source and target copies of V, then we speak of the *matrix of θ with respect to S*.

1.38 Definition. If $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $S' = \{\mathbf{v}'_1, \dots, \mathbf{v}'_n\}$ are bases of V then the *transition matrix from S to S'* is the matrix $P = (p_{ij})$ whose jth column is the coordinates of $\mathbf{v}'_j$ with respect to S. Thus $\mathbf{v}'_j = \sum_{i=1}^n p_{ij}\mathbf{v}_i$.

We have $[\mathbf{v}]_S = P[\mathbf{v}]_{S'}$ for $\mathbf{v} \in V$ since if $\mathbf{x} = [\mathbf{v}]_{S'}$, then

$$\mathbf{v} = \sum_{j=1}^n x_j\mathbf{v}'_j = \sum_{j=1}^n x_j \sum_{i=1}^n p_{ij}\mathbf{v}_i = \sum_{i=1}^n (\sum_{j=1}^n p_{ij}x_j)\mathbf{v}_i = \sum_{i=1}^n (P\mathbf{x})_i\mathbf{v}_i.$$

Note that P is invertible; its inverse is the transition matrix in the opposite direction.

1.39 Definition. We shall most often be interested in linear maps $\theta : V \to V$ from a vector space to iteslf. We call them *endomorphisms*. In this case we shall use the same basis S for both the source and target vector spaces. We speak about the matrix for θ with respect to the basis S used for both source and target copies of V.

1.40 Definition. Two $n \times n$ matrices A, A' are *similar* if there is an invertible matrix P with $A' = P^{-1}AP$.

1.41 Definition. Suppose A is an $n \times n$ matrix and $\lambda \in F$.
Geometric multiplicity of λ = dimension of the λ-eigenspace $Esp(\lambda)$ for A.
Algebraic multiplicity of λ = multiplicity of λ as a root of the characteristic poly $\chi_A(t)$.

1.42 Definition. A set of vectors $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is *orthogonal* if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i \neq j$. It is *orthonormal* if also $|\mathbf{v}_i| = 1$ for all i, so

$$\mathbf{v}_i \cdot \mathbf{v}_i = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j). \end{cases}$$

1.43 Definition. A real $n \times n$ matrix P is said to be *orthogonal* if it is invertible and $P^{-1} = P^T$, that is, $P^T P = I = PP^T$.

(In fact you only need to check that $P^T P = I$. It follows that $\det P \neq 0$, so P is invertible, so $P^{-1} = P^T$.)

The set of orthogonal matrices forms a subgroup $O_n(\mathbb{R})$ of $\mathrm{GL}_n(\mathbb{R})$, the *orthogonal group*. The set of orthogonal matrices of determinant 1 forms a subgroup $SO_n(\mathbb{R})$, the *special orthogonal group.*

All the above are taken directly from the notes, these definitions are all required knowledge (to the best of my memory).

2 Theorems, Lemmas, Corollaries, etc.

2.1 Groups, Subgroups, and Order

2.1 Proposition. *The identity element of a group is unique.*

Proof. Suppose that e and f are identity elements for the group G. Consider ef. Since e is an identity element $ef = f$. Since f is an identity element $ef = e$. Thus $e = f$. $\square$

2.2 Proposition. *Given a group G and an element $x \in G$, there is only one inverse for x, that is, there is only one element y with $xy = 1$ and $yx = 1$.*

Proof. Say $xy = yx = 1$ and $xz = zx = 1$. Then $(yx)z = 1z = z$. But also $(yx)z = y(xz) = y1 = y$. Thus $y = z$. $\square$

2.3 Proposition. *For elements x, y in a group G, the following properties hold. (1) If $xy = 1$, then $x = y^{-1}$ and $y = x^{-1}$. (2) $(xy)^{-1} = y^{-1}x^{-1}$. (3) $(x^{-1})^{-1} = x$.*

2.4 Proposition. *[Cancellation] For an elements x, y, z of a group G. (i) If $xy = xz$ then $y = z$. (ii) If $yx = zx$ then $y = z$.*

2.5 Proposition. *The multiplication table for a group is a 'Latin square', that is, each row and each column contains all group elements, once each.*

Proof. e.g. for rows. Cancellation shows that each element only occurs once, for if $xy = xy'$ then $y = y'$. Now the element z occurs in the row for x, at column y if $xy = z$, so we may take $y = x^{-1}z$. $\square$

2.6 Lemma. *If H is a subgroup of G, then (i) they have the same identity element (in particular H contains the identity of G), and (ii) the inverse of any element of H is the same whether you use the group structure of H or that of G.*

Proof. (i) Denote them by 1_H and 1_G. Then $1_G 1_H = 1_H$ and $1_H 1_H = 1_H$ so $1_G 1_H = 1_H 1_H$, and hence $1_G = 1_H$ by cancellation. (ii) Say $h \in H$ has inverse y in H. Then $hy = 1 = yh$. Then y is the inverse of h in G. $\square$

2.7 Theorem (Subgroup criterion)**.** *Let* $(G, \circ)$ *be a group. A subset* H *of* G *is a subgroup if and only if it satisfies the following properties (i)* $1 \in H$*, (ii)* $xy \in H$ *for all* $x, y \in H$*, and (iii)* $x^{-1} \in H$ *for all* $x \in H$*.*

Proof. First suppose that (i), (ii) and (iii) hold. Then (ii) says that H is closed under $\circ$, and it inherits associativity from G. Then $1 \in H$ by (i), and it is an identity for H. Also each element $x \in H$ has an inverse in $x^{-1} \in H$ by (iii). Thus H is a subgroup.

Conversely suppose that H is a subgroup. Then since H is closed under $\circ$, (ii) holds. Now (i) and (iii) follow from the lemma. □

2.8 Theorem. *Suppose* $x \in G$*. (i) If* x *has infinite order, then all powers* x^k $(k \in \mathbb{Z})$ *are distinct. In particular* $x^k = 1$ *if and only if* $k = 0$*. (ii) If* x *has finite order* n*, then as* k *increases, the powers* x^k *repeat in cycles of length* n*. In particular* $x^k = 1$ *if and only if* k *is a multiple of* n *(even for negative* k*).*

Proof. (i) Suppose that $x^j = x^k$ where $j < k$. Then $x^{k-j} = 1$, contrary to x having infinite order. (ii) The first n powers $x^0, x^1, x^2, \ldots, x^{n-1}$ are all distinct, for if $x^j = x^k$ where $0 \leq j < k \leq n-1$ then $x^{k-j} = 1$ and $0 < k - j < n$, contrary to x having order n. Now for any integer N we can divide by n giving an integer quotient q and remainder r with $0 \leq r \leq n-1$. Then $N = nq + r$, and $x^N = x^{nq+r} = (x^n)^q x^r = 1^q x^r = 1x^r = x^r$. □

2.2 Structure of Groups

2.9 Lemma. *The order of the group* $\langle x \rangle$ *is equal to the order of the element* x*.*

Proof. Follows from Theorem 2.8. □

2.10 Theorem. *A finite group of order* n *is cyclic if and only if it contains an element of order* n*.*

Proof. If G is cyclic, then any generator is of order n. Conversely if G contains an element of order n, then $\langle x \rangle$ is a subgroup of G with the same number of elements, so we must have $\langle x \rangle = G$. □

2.11 Proposition. *Any cyclic group is abelian.*

Proof. $(x^n)(x^m) = x^{n+m} = (x^m)(x^n)$. □

2.12 Theorem. *Any subgroup of* $\mathbb{Z}$ *is of the form* $k\mathbb{Z}$ *for some* k *(which is the same as* $\langle k \rangle$ *in this case), so is cyclic.*

Proof. Suppose H is a subgroup of $\mathbb{Z}$. If $H = \{0\}$ then $H = 0\mathbb{Z}$, so we may suppose that H contains a non-zero element. Then H contains a positive element. Let $k > 0$ be minimal with $k \in H$. Then $k\mathbb{Z} \subseteq H$. Suppose that $h \in H$ and $h \notin k\mathbb{Z}$. Dividing by k gives integer quotient q and remainder r with $0 < r < k$. Then by the choice of k we have $r \notin H$. But then $h = qk + r$, so $r = h - qk \in H$. Contradiction. □

2.13 Theorem. *Any subgroup of a cyclic group is cyclic.*

Proof. Suppose $G = \langle x \rangle$ and $H \leq G$. Define

$$K = \{k \in \mathbb{Z} : x^k \in H\}.$$

This is a subgroup of $\mathbb{Z}$, for $x^0 = 1 \in H$, so $0 \in K$. Also if $k, j \in K$ then $x^k, x^j \in H$, so $x^{k+j} = x^k x^j \in H$, so $k + j \in K$, and $x^{-k} = (x^k)^{-1} \in H$, so $-k \in K$. Thus by the previous theorem $K = n\mathbb{Z}$ for some n. Then $H = \{x^k : k \in K\} = \{x^{nj} : j \in \mathbb{Z}\} = \{((x^n)^j : j \in \mathbb{Z}\} = \langle x^n \rangle$, so H is cyclic. $\square$

2.14 Lemma. *The order of $(g, h) \in G \times H$ is the least common multiple of the orders of g and h (or ∞ if g or h has infinite order).*

Proof. Suppose that the orders n and m of g and h are finite. The order of (g, h) is the least positive N with $(g, h)^N = 1 = (1_{G,H})$, so with $g^N = 1_G$ and $h^N = 1_H$. This holds if and only if N is a common multiple of n and m. $\square$

2.15 Theorem. *If G and H are finite cyclic groups, then $G \times H$ is cyclic if and only the orders of G and H are coprime.*

Proof. Let $G = \langle x \rangle$ and $H = \langle y \rangle$ have order n and m. Suppose n and m are coprime.

Then (x, y) has order the least common multiple of n and m, but since they are coprime this is nm. This $G \times H$ is cyclic. Conversely suppose that n and m are not coprime.

Then their least common multiple ℓ is $< nm$. Then for any integers j, k we have $(x^j, y^k)^\ell = (x^{j\ell}, y^{k\ell}) = (1_G, 1_H)$ since $j\ell$ is a multiple of the order of n and $k\ell$ is a multiple of m. Thus every element of $G \times H$ has order $\leq \ell$. Thus no element has order equal to the order of $G \times H$, so $G \times H$ is not cyclic. $\square$

2.16 Lemma. *If $\theta : G \to H$ is a homomorphism, then $\theta(1_G) = 1_H$ and $\theta(g^{-1}) = (\theta(g))^{-1}$ for all $g \in G$.*

Proof. $\theta(1_G)\theta(1_G) = \theta(1_G^2) = \theta(1_G) = \theta(1_G)1_H$, so $\theta(1_G) = 1_H$ by cancellation. Now $gg^{-1} = 1_G$, so $\theta(g)\theta(g^{-1}) = \theta(1_G) = 1_H$, giving $\theta(g^{-1}) = (\theta(g))^{-1}$. $\square$

2.17 Proposition. *Suppose that $\theta : G \to H$ is an isomorphism. Then: (i) $|G| = |H|$. (ii) $\theta(1_G) = 1_H$. (iii) $\theta(g^{-1}) = (\theta(g))^{-1}$ for all $g \in G$. (iv) For all $g \in G$ the elements g and $\theta(g)$ have the same order. (v) For each n, the groups G and H have the same number of elements of order n. (v) G is abelian if and only if H is abelian. (vi) G is cyclic if and only if H is cyclic.*

Proof. Straightforward, since G and H have the same multiplication table, and all of these properties can be read off from the multiplication table. $\square$

2.18 Theorem. *Two cyclic groups are isomorphic if and only if they have the same order.*

Proof. If they are isomorphic they must have the same order. For the converse, suppose the groups are $G = \langle x \rangle$ and $H = \langle y \rangle$ and that they have the same order. If the order is infinite, then the elements x^k are all distinct, so we can define $\theta : G \to H$ by $\theta(x^k) = y^k$. It is a homomorphism since $\theta(x^j x^k) = \theta(x^{j+k}) = y^{j+k} = y^j y^k = \theta(x^j)\theta(x^k)$. Clearly it is bijective, so it is an isomorphism. Thus suppose the order is finite, say n. The powers x^k repeat with period n, and similarly the powers y^k. Thus we can again define $\theta : G \to H$ with $\theta(x^k) = y^k$, and again get an isomorphism. □

Examined: Q2(i) 2017.

2.19 Corollary. *[Chinese Remainder Theorem]* $\mathbb{Z}_n \times \mathbb{Z}_m \cong \mathbb{Z}_{nm}$ *if and only if* n *and* m *are coprime.*

Proof. If n and m are coprime, then $\mathbb{Z}_n \times \mathbb{Z}_m$ is cyclic. Also $\mathbb{Z}_{nm}$ is cyclic of the same order, so they must be isomorphic. If n and m are not coprime then $\mathbb{Z}_n \times \mathbb{Z}_m$ is not cyclic, so not isomorphic to $\mathbb{Z}_{nm}$. □

2.20 Theorem. *(Lagrange): If* H *is a subgroup of the finite group* G*, then* $|H|$ *divides* $|G|$*. [This is the longest proof thus far, I have chosen not to omit it as this appears to be* **THE MOST IMPORTANT** *theorem on this half of the module].*

Proof. We define $\sim$ on G by letting $x \sim y$ if $xy^{-1} \in H$. We verify that this is an equivalence relation on G: reflexivity: $x \sim x$ since $xx^{-1} = 1 \in H$ (as H contains the identity) symmetry: if $x \sim y$ then $xy^{-1} \in H$. Hence $yx^{-1} = ((yx^{-1})^{-1})^{-1} = (xy^{-1})^{-1} \in H$ so $y \sim x$ (as H is closed under inverses) transitivity: if $x \sim y$ and $y \sim z$ then $xy^{-1}, yz^{-1} \in H$, so $xz^{-1} = (xy^{-1})(yz^{-1}) \in H$ so $x \sim z$ (as H is closed under the operation). Since $\sim$ is an equivalence relation, it partitions G into $\sim$-classes. The $\sim$-class containing x is $\{y : y \sim x\} = \{y : yx^{-1} \in H\} = \{y : yx^{-1} = h$, some $h \in H\} = \{y : y = hx$, some $h \in H\}$. This set is written Hx, and is called the *right coset* of H in G containing x. To sum up, every member of G lies in some right coset of H, and the right cosets form a partition of G. Finally we see that each right coset Hx has $|H|$ members (noting that $H.1 = H$, so that H is itself a right coset). Map H to Hx by f where $f(h) = hx$. By definition of Hx this maps H onto Hx, and f is 1–1, since if $f(h_1) = f(h_2)$, then $h_1 x = h_2 x$, so by cancellation, $h_1 = h_2$. This f is a 1–1 map from H onto Hx, so Hx has $|H|$ members. □

Examined 2018, Q2(i) [State], and also 2019 2(ii) [Prove]

2.21 Theorem. *If* H *is a subgroup of a finite group* G*, then* $|G| = |H|.|G : H|$*.*

Proof. Follows from 2.20 □

2.22 Corollary. *The order of an element of a finite group divides the order of the group.*

Proof. The order of x is the same as the order of the subgroup $\langle x \rangle$ of G. □

2.23 Corollary. *Any group of prime order is cyclic.*

Proof. If G has order p, then any non-identity element has order $\neq 1$, so must have order p. Thus it generates the group. $\square$

2.24 Corollary. *If G is a group of order n, then $x^n = 1$ for all $x \in G$.*

Proof. The order d of x divides n, so $n = dk$ for some k. Then $x^n = x^{dk} = (x^d)^k = 1^k = 1$. $\square$

2.25 Corollary. *[Fermat's little theorem] If p is a prime number and a is coprime to p, then $a^{p-1} \equiv 1 \pmod{p}$ (which means that $a^{p-1} - 1$ is a multiple of p).*

Proof. Consider $a \in \mathbb{Z}_p^*$. We have $(a)^{p-1} = 1$, so $a^{p-1} = 1$. $\square$

2.26 Theorem. *Up to isomorphism the groups of order 4 are the cyclic group C_4 and the Klein four group $V \cong C_2 \times C_2$.*

Proof. If it is not cyclic, then every element has order 1 or 2, so every element has $x^2 = 1$. This determines the multiplication table to be that of the Klein four group. $\square$

Examined 2018 Q2(ii)

2.27 Theorem. *Up to isomorphism the groups of order 6 are the cyclic group C_6 and the dihedral group D_3. Proof omitted for length, see 3.1 for proof.*

2.28 Theorem. *Up to isomorphism the groups of order 8 are $\mathbb{Z}_8$, $\mathbb{Z}_4 \times \mathbb{Z}_2$, $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$, the dihedral group D_4 and the quaternion group Q. Given without proof.*

2.3 Permutations, Homomorphism, Symmetric and Alternating groups

2.29 Proposition. *The group S_n has order $n!$.*

2.30 Remarks. (i) Cycle notation doesn't tell you which S_n you are working in. For example the cycle $(2\ 5\ 4)$ could be a permutation in S_n for any $n \geq 5$. (ii) A k-cycle can be written in k different ways. For example $(2\ 5\ 4) = (5\ 4\ 2) = (4\ 2\ 5)$. A 1-cycle is the identity. (iii) A k-cycle has order k. (iv) We say a collection if cycles is *disjoint* if there is no number a occurring in two of them. For example $(2\ 5\ 4)$ and $(1\ 3)$ are disjoint. Disjoint cycles commute, $(2\ 5\ 4)(1\ 3) = (1\ 3)(2\ 5\ 4)$.

2.31 Theorem. *Every permutation can be written as a product of disjoint cycles. The decomposition is essentially unique, apart from the order of the cycles and the different ways of writing a cycle.*

2.32 Corollary. *To find the order of a permutation, write it as a product of disjoint cycles and take the least common multiple of their lengths.*

Proof. Write π as a product of disjoint cycles, say $\pi = c_1 c_2 \dots c_k$. The order is the least $d > 0$ with $\pi^d = e$. Since disjoint cycles commute we get $\pi^d = c_1^d c_2^d \dots c_k^d = e$. Now the permutations $c_1^d, \dots, c_k^d$ act on disjoint subsets of $\{1, \dots, n\}$, so the only way that their product can be the identity is if each of them is the identity, so d must be a multiple of the orders of the cycles. □

2.33 Corollary. *Every permutation can be written as a product of transpositions.*

Proof. We have $(a_1\ a_2\ a_3\ \dots\ a_k) = (a_1\ a_k) \dots (a_1\ a_3)(a_1\ a_2)$. □

2.34 Lemma. *For permutations $\pi, \sigma \in S_n$ we have $A_{\pi\sigma} = A_\pi A_\sigma$ and $\epsilon(\pi\sigma) = \epsilon(\pi)\epsilon(\sigma)$.*

Proof. $A_\pi A_\sigma \mathbf{e}_j = A_\pi \mathbf{e}_{\sigma(j)} = \mathbf{e}_{\pi(\sigma(j))} = A_{\pi\sigma}\mathbf{e}_j$. Then $\epsilon(\pi\sigma) = \det(A_{\pi\sigma}) = \det(A_\pi A_\sigma) = \det(A_\pi)\det(A_\sigma) = \epsilon(\pi)\epsilon(\sigma)$. □

2.35 Theorem. *Every permutation is either odd or even, and not both. The sign of a permutation is 1 if it is even and -1 if it is odd. In particular the sign of a permutation is always in $\{\pm 1\}$.*

Proof. We know that any permutation can be written as a product of transpositions (although this expression is not unique). Also $\epsilon(\pi\sigma) = \epsilon(\pi)\epsilon(\sigma)$. It thus suffices to show that if τ is a transposition then $\epsilon(\tau) = -1$. But if $\tau = (a\ b)$ then A_τ is obtained from the identity matrix by exchanging rows a and b. Now the identity matrix has determinant 1, and exchanging any two rows changes the sign, so $\det A_\tau = -1$. □

Examined 2018 Q3(i)

2.36 Proposition. *For $n > 1$, we have $[S_n : A_n] = 2$, and so $|A_n| = n!/2$.*

Proof. Fix a transposition $\tau \in S_n$, for example $\tau = (1\ 2)$. For any odd permutation $\pi \in S_n$ we have $\pi\tau \in A_n$. Then $\tau^2 = e$, so $\pi = (\pi\tau)\tau \in A_n\tau$. Thus $S_n = A_n \cup A_n\tau$. Thus there are only two cosets of A_n in S_n. □

2.37 Theorem (Leibniz formula)**.** *If $A = (a_{ij})$ is an $n \times n$ matrix, then*

$$\det A = \sum_{\pi \in S_n} \epsilon(\pi) a_{\pi(1),1} a_{\pi(2),2} \dots a_{\pi(n),n}.$$

2.38 Proposition. *If $\theta : G \to H$ is a homomorphism, then (i) $\theta(1) = 1$, or, more precisely, $\theta(1_G) = 1_H$. (ii) $\theta(g^{-1}) = \theta(g)^{-1}$ for $g \in G$.*

Proof. (i) $\theta(1_G)\theta(1_G) = \theta(1_G 1_G) = \theta(1_G) = \theta(1_G)1_H$, so $\theta(1_G) = 1_H$ by cancellation. (ii) $\theta(g^{-1})\theta(g) = \theta(g^{-1}g) = \theta(1_G) = 1_H$, so $\theta(g^{-1}) = \theta(g)^{-1}$. □

2.39 Proposition. *If $\theta : G \to H$ is a homomorphism between two groups, then $\ker\theta$ is a subgroup of G and $\operatorname{im}\theta$ is a subgroup of H.*

Proof. We have $1 \in \ker\theta$. If $g, g' \in \ker\theta$ then $\theta(gg') = \theta(g)\theta(g') = 1 \circ 1 = 1$, so $gg' \in \ker\theta$. If $g \in \ker\theta$ then $\theta(g^{-1}) = \theta(g)^{-1} = 1^{-1} = 1$, so $g^{-1} \in \ker\theta$. Thus $\ker\theta$ is a subgroup of G. We have $\theta(1) = 1$, so $1 \in \operatorname{im}\theta$. If $h, h' \in \operatorname{im}\theta$, then $h = \theta(g)$ and $h' = \theta(g')$ for some $g, g' \in G$. Then $hh' = \theta(g)\theta(g') = \theta(gg') \in \operatorname{im}\theta$. Also $h^{-1} = \theta(g)^{-1} = \theta(g^{-1}) \in \operatorname{im}\theta$. Thus $\operatorname{im}\theta$ is a subgroup of H. □

2.40 Proposition. *If $\theta : G \to H$ is a homomorphism, then θ is injective if and only if $\ker\theta = \{1\}$. In this case θ defines an isomorphism $G \cong \operatorname{im}\theta$.*

Proof. If θ is injective and $x \in \ker\theta$ then $\theta(x) = 1 = \theta(1)$, so since θ is injective, $x = 1$. Thus $\ker\theta = \{1\}$. Conversely suppose that $\ker\theta = \{1\}$. Suppose that $\theta(x) = \theta(y)$. Then $\theta(xy^{-1}) = \theta(x)\theta(y^{-1}) = \theta(x)\theta(y)^{-1} = \theta(x)\theta(x)^{-1} = 1$. Thus $xy^{-1} \in \ker\theta$, so $xy^{-1} = 1$. Thus $x = y$ and θ is injective. Now θ gives an homomorphism $G \to \operatorname{im}\theta$ which is injective and surjective, so it is an isomorphism. □

2.41 Theorem. *(i) The group CUB of rotations preserving a cube is isomorphic to S_4. (ii) The group TET of rotations preserving a regular tetrahedron is isomorphic to A_4. Proof omitted as this is a very specific theorem that has not come up in the last 3 available exams, see 3.2 for proof.*

2.42 Theorem (Cayley's Theorem)**.** *Any group of order n is isomorphic to a subgroup of S_n.*

Proof. Let $G = \{g_1, g_2, \dots, g_n\}$. For each $g_i \in G$, the Latin square property gives a permutation $\pi(g_i) \in S_n$ with $g_i g_j = g_{\pi(g_i)(j)}$ for all j. Now

$$g_i(g_j g_k) = g_i g_{\pi(g_i)(k)} = g_{\pi(g_i)(\pi(g_j)(k))} \quad \text{and} \quad (g_i g_j) g_k = g_{\pi(g_i g_j)(k)}.$$

Thus $\pi(g_i)(\pi(g_j)(k)) = \pi(g_i g_j)(k)$ for all i, j, k, so $\pi(g_i) \circ \pi(g_j) = \pi(g_i g_j)$ for all i, j, so π defines a homomorphism $G \to S_n$. It is injective since if $\pi(g_i) = id$ then $g_i = id$. □

2.43 Theorem. *G is the disjoint union of its conjugacy classes.*

Proof. It suffices to prove that conjugacy is an equivalence relation. So define a relation $\sim$ on G by $x \sim y$ if $y = g^{-1}xg$ for some $g \in G$. Then (reflexive) For any $x \in G$ the condition $x \sim x$ holds since $x = 1^{-1}x1$. (symmetric) If $x \sim y$ then $y = g^{-1}xg$. Then $x = (g^{-1})^{-1}y(g^{-1})$, so $y \sim x$. (transitive) If $x \sim y$ and $y \sim z$ then $y = g^{-1}xg$ and $z = h^{-1}yh$. Then $z = (gh)^{-1}x(gh)$, so $x \sim z$. □

2.44 Proposition. *Conjugate elements have the same order.*

Proof. If $y = g^{-1}xg$ and $n > 0$, then

$$y^n = 1 \Leftrightarrow \underbrace{(g^{-1}xg)(g^{-1}xg)\dots(g^{-1}xg)}_{n \text{ copies}} = 1 \Leftrightarrow g^{-1}x^n g = 1 \Leftrightarrow x^n = 1.$$

□

2.45 Theorem. *The conjugacy class of $x \in G$ has size $|\mathrm{conj}_G(x)| = [G : C_G(x)]$.*

Proof. $g^{-1}xg = (g')^{-1}xg' \Leftrightarrow xg(g')^{-1} = g(g')^{-1}x \Leftrightarrow g(g')^{-1} \in C_G(x) \Leftrightarrow$ the cosets $C_G(x)g$ and $C_G(x)g'$ are equal. Thus the number of different conjugates of x is equal to the number of different cosets of $C_G(x)$ in G. □

2.46 Theorem. *If $\theta : G \to G'$ is a homomorphism then $\ker\theta$ is a normal subgroup of G.*

Proof. If $x \in \ker\theta$ and $g \in G$ then $\theta(g^{-1}xg) = \theta(g)^{-1}\theta(x)\theta(g) = \theta(g)^{-1}\theta(g) = 1$, so $g^{-1}xg \in \ker\theta$. □

2.47 Theorem. *A subgroup H of G is normal if and only if $Hg = gH$ for all $g \in G$, so that the right cosets are the same as the left cosets.*

Proof. If H is normal and $g \in G$ we need to show $Hg \subseteq gH$ and $gH \subseteq Hg$. If $h \in H$ then $g^{-1}hg \in H$ and $hg = g(g^{-1}hg) \in gH$, giving the first inclusion. Also $ghg^{-1} = (g^{-1})^{-1}h(g^{-1}) \in H$, and $gh = (ghg^{-1})g \in Hg$ giving the second inclusion. Conversely, if $Hg = gH$ and $h \in H$, then $hg = gh'$ for some $h' \in H$, so $g^{-1}hg = h' \in H$, so H is normal. □

2.48 Proposition. *Any subgroup of index 2 in a group is normal.*

Proof. Since there are only two right cosets, they must be H and $G \setminus H$. Similarly the left cosets must be H and $G \setminus H$. Thus the right cosets of H are the same as the left cosets. □

2.49 Lemma. *Let H be a subgroup of a group G. The following are equivalent (i) H is a normal subgroup of G. (ii) For all $g, g' \in G$ we have: if $x \in Hg$ and $y \in Hg'$ then $xy \in H(gg')$.*

Proof. Suppose (i) holds. Let $x = hg$ and $y = h'g'$. Then $xy = h(gh'g^{-1})(gg') \in H(gg')$ since $gh'g^{-1} \in H$. Conversely if (ii) holds and $h \in H$, then $g^{-1} \in Hg^{-1}$ and $hg \in Hg$ so $g^{-1}hg \in H(g^{-1}g) = H1 = H$. □

2.50 Theorem (First isomorphism theorem)**.** *If $\theta : G \to G'$ is a homomorphism, then there is an isomorphism $\overline{\theta} : G/\ker\theta \to \mathrm{im}\,\theta$ defined by $\overline{\theta}(Hg) = \theta(g)$, where $H = \ker\theta$.*

Proof. The map $\overline{\theta}$ is well-defined and injective since $Hx = Hy \Leftrightarrow xy^{-1} \in H = \ker\theta \Leftrightarrow \theta(xy^{-1}) = 1 \Leftrightarrow \theta(x)\theta(y)^{-1} = 1 \Leftrightarrow \theta(x) = \theta(y)$. It is clearly surjective, and it is a homomorphism by the definition of the product in G/H. □

Examined 2018 Q3(iii)

2.4 Vector Spaces

2.51 Proposition. *Suppose that V is a vector space over F. (i) If $a \in F$ is any scalar and $\mathbf{0} \in V$ is the zero vector, then $a\mathbf{0} = \mathbf{0}$. (ii) If 0 is the zero element of the field F and $\mathbf{v} \in V$ is any vector, then $0\mathbf{v} = \mathbf{0}$. (iii) If $a \in F$ and $\mathbf{v} \in V$ and $a\mathbf{v} = \mathbf{0}$, then either $a = 0$ or $\mathbf{v} = 0$. (iv) If $\mathbf{v} \in V$ then $(-1)\mathbf{v} = -\mathbf{v}$, and in general $(-a)\mathbf{v} = -(a\mathbf{v})$, for any $a \in F$.*

Proof. These are straightforward consequences from the axioms. For example for (i), observe that $\mathbf{0} + \mathbf{0} = \mathbf{0}$ since $\mathbf{0}$ is the additive identity. Thus $a(\mathbf{0} + \mathbf{0}) = a\mathbf{0}$, so $a\mathbf{0} + a\mathbf{0} = a\mathbf{0}$ by distributivity. Subtracting $a\mathbf{0}$ from both sides gives $a\mathbf{0} = \mathbf{0}$. (ii) observe that $0 + 0 = 0$ in the field F, hence $(0 + 0)\mathbf{v} = 0\mathbf{v}$, so $0\mathbf{v} = 0\mathbf{v} - 0\mathbf{v}$, $0\mathbf{v} = 0(\mathbf{0}) = \mathbf{0}$ (by (i). (iii) Let $a\mathbf{v} = \mathbf{0}$. If $a \neq 0$, then one can divide by a in the field F, so there is an element $a^{-1} = \frac{1}{a} \in F$. Now $a^{-1}(a\mathbf{v}) = \mathbf{v} = \mathbf{0}$ as required. (iv) Consider $(1 + (-1))\mathbf{v}$, $\mathbf{v} + (-1)\mathbf{v} = 0\mathbf{v}$, $(-1)\mathbf{v} = -\mathbf{v}$. ☐

2.52 Theorem (Subspace criterion)**.** *Let V be a vector space over a field F. A subset U of V is a subspace if and only if it satisfies the following properties (i) $\mathbf{0} \in U$. (ii) For all $\mathbf{u}, \mathbf{u}' \in U$ we have $\mathbf{u} + \mathbf{u}' \in U$, and (iii) For all scalars $a \in F$ and elements $\mathbf{u} \in U$ we have $a\mathbf{u} \in U$.*

Proof. Similar to 2.7. ☐

2.53 Proposition. *Given an $m \times n$ matrix A with entries in F, one gets a linear map $\theta_A : F^n \to F^m$ given by $\theta_A(\mathbf{v}) = A\mathbf{v}$. Conversely any linear map $\theta : F^n \to F^m$ is of the form θ_A, where A is the matrix whose columns are $\theta(\mathbf{e}_1), \theta(\mathbf{e}_2), \ldots, \theta(\mathbf{e}_n)$.*

2.54 Proposition. *If $\theta : V \to W$ is a linear transformation, then $\ker\theta$ is a subspace of V and $\operatorname{im}\theta$ is a subspace of W.*

2.55 Lemma. *Given a finite set of vectors $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in a vector space V over F, the mapping $\phi_S : F^n \to V$ given by*

$$\phi_S(a_1, \ldots, a_n)^T = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$

is a linear map.

Proof. Let $\mathbf{a} = (a_1, \ldots, a_n)^T, \mathbf{a}' = (a'_1, \ldots, a'_n)^T \in F^n$. Then $\phi_S(\mathbf{a} + \mathbf{a}') = \phi_S(a_1 + a'_1, \ldots, a_n + a'_n) = (a_1 + a'_1)\mathbf{v}_1 + \cdots + (a_n + a'_n)\mathbf{v}_n = (a_1\mathbf{v}_1 + \ldots a_n\mathbf{v}_n) + (a'_1\mathbf{v}_1 + \cdots + a'_n\mathbf{v}_n) = \phi_S(\mathbf{a}) + \phi_S(\mathbf{a}')$. Also $\phi_S(\lambda\mathbf{a}) = \phi_S(\lambda a_1, \ldots, \lambda a_n)^T = (\lambda a_1)\mathbf{v}_1 + \cdots + (\lambda a_n)\mathbf{v}_n = \lambda(a_1\mathbf{v}_1 + \ldots a_n\mathbf{v}_n) = \lambda\phi_S(\mathbf{a})$. ☐

2.56 Proposition. *$\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent if and only if some $\mathbf{v}_i$ is a linear combination of its predecessors, that is, $\mathbf{v}_i \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}\}$.*

Proof. Suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent, so there is a relation $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$ with some coefficient nonzero. Let i be maximal with $a_i \neq 0$. Then $a_1\mathbf{v}_1 + \cdots + a_i\mathbf{v}_i = \mathbf{0}$, so

$$\mathbf{v}_i = (-\frac{a_1}{a_i})\mathbf{v}_1 + \cdots + (-\frac{a_{i-1}}{a_i})\mathbf{v}_{i-1} \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}\}.$$

Conversely, if $\mathbf{v}_i \in \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{i-1}\}$, then $\mathbf{v}_i = b_1\mathbf{v}_1 + \cdots + b_{i-1}\mathbf{v}_{i-1}$ for some scalars $b_1, \dots, b_{i-1}$, giving a relation $b_1\mathbf{v}_1 + \cdots + b_{i-1}\mathbf{v}_{i-1} - \mathbf{v}_i = \mathbf{0}$. □

2.57 Theorem. *Given vectors $S = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ in F^n, write them as the rows of a matrix A, and row reduce to echelon form giving a matrix B. Then (i)* $\operatorname{span} S$ = *span of the rows of B. This is equal to $F^n \Leftrightarrow B$ has non-zero leading elements in every column. (ii) the rows of A are linearly independent $\Leftrightarrow$ the rows of B are linearly independent $\Leftrightarrow$ B has no rows which are entirely zero. Given without proof.*

2.58 Theorem. *In any vector space, if I is a linearly independent set and S is a spanning set, then $|I| \le |S|$.*

Proof. Write each element of the linearly independent set $I = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ as as a linear combination of the vectors in the spanning set $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, say

$$\begin{array}{ccccccc} \mathbf{w}_1 & = & a_{11}\mathbf{v}_1 & + & \dots & + & a_{1n}\mathbf{v}_n \\ \mathbf{w}_2 & = & a_{21}\mathbf{v}_1 & + & \dots & + & a_{2n}\mathbf{v}_n \\ \cdots & & & & & & \\ \mathbf{w}_m & = & a_{m1}\mathbf{v}_1 & + & \dots & + & a_{mn}\mathbf{v}_n \end{array}$$

The coefficients give an $m \times n$ matrix A. The rows of A are linearly independent because any relation between them would give a relation between the $\mathbf{w}_j$. Thus, after row reducing, the matrix has no zero rows. But this is only possible if the number of rows is $\le$ the number of columns, that is, $m \le n$. □

2.5 Diagonalizability and the orthogonal group.

2.59 Theorem. *Let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a finite subset of V. Then the following are equivalent (i) S is a basis of V (ii) $\phi_S : F^n \to V$ is an isomorphism of vector spaces (iii) every $\mathbf{v} \in V$ can be written in a unique way as a linear combination $\mathbf{v} = a_1\mathbf{v}_1 + \dots + a_n\mathbf{v}_n$.*

Proof. (i)$\Leftrightarrow$(ii). Since $\operatorname{span} S = \operatorname{im} \phi_S$, S spans V if and only if ϕ_S is surjective. Also S is linearly independent if and only if $\ker\phi_S = \{\mathbf{0}\}$, which is if and only if ϕ_S is injective.
(ii)$\Leftrightarrow$(iii). Clear. □

2.60 Theorem. *Any two bases of a vector space have the same number of elements.*

Proof. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ and $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ be bases of V. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is independent and $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ spans, so $k \le n$ by Theorem 2.58. Also $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ is independent and $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ spans, so $n \le k$. Thus $n = k$. □

2.61 Theorem. *(i) In a vector space, any spanning set contains a basis. Thus any spanning set in a vector space of dimension n has $\ge n$ elements, and if it has exactly n, then it is a basis.*
(ii) In a finite-dimensional vector space, any linearly independent set can be extended to a basis. Thus any linearly independent set in a vector space of dimension n has $\le n$ elements, and if it has exactly n, then it is a basis.

Proof. (i) Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a spanning set. If S is linearly independent, then it is already a basis. Thus assume that S is linearly dependent. Then some $\mathbf{v}_i \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}\}$. It follows that $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n\}$ is spanning. Now either S' is linearly independent, so a basis of V, or we can continue in the same way, eliminating further elements. Eventually we obtain a basis of V.

(ii) Let I be the linearly independent set and let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be a basis. If $\mathbf{u}_1 \notin \operatorname{span} I$, replace I by $I \cup \{\mathbf{u}_1\}$. It is still linearly independent. Now if $\mathbf{u}_2 \notin \operatorname{span} I$, replace I by $I \cup \{\mathbf{u}_2\}$, and so on. At the end, we have enlarged I to a linearly independent set whose span contains all elements in the basis, so it is a basis. □

2.62 Theorem. *If W is a subspace of a finite-dimensional vector space V, then W is finite-dimensional and $\dim W \leq \dim V$. Moreover, if $\dim W = \dim V$ then $W = V$.*

Proof. Any linearly independent subset S of W is linearly independent in V so has at most $\dim V$ elements. Thus we can choose one with as many elements as possible. Every element $\mathbf{w} \in W$ is in $\operatorname{span} S$, for otherwise $S \cup \{\mathbf{w}\}$ is linearly independent by Proposition 2.56. Thus S is a basis for W. Straightforward from Theorem 2.61, but omitted. If V has dimension n, then any linearly independent subset of W has $\leq n$ elements, and it is easy to see, using , that a linearly independent subset of W of maximal size must be a basis of W. □

2.63 Theorem. *Two finite-dimensional vector spaces over F are isomorphic if and only if they have the same dimension.*

Proof. If they both have dimension n, then they are both isomorphic to F^n, so they are isomorphic to each other.

Conversely, if $\theta : V \to W$ is an isomorphism, and V is finite-dimensional, with basis $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, then it is easy to see that $\{\theta(\mathbf{v}_1), \ldots, \theta(\mathbf{v}_n)\}$ is a basis for W, so W also has dimension n. □

2.64 Proposition. *If V is a finite-dimensional vector space and U is a subspace of V, then $\dim V/U = \dim V - \dim U$.*

Proof. Take a basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ of U. It can be extended to give a basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ of V. We check that $\{U + \mathbf{v}_1, \ldots, U + \mathbf{v}_\ell\}$ is a basis of V/U.

Span. Any element of V/U is of the form $U + \mathbf{v}$. We can write $\mathbf{v} = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k + b_1\mathbf{v}_1 + \cdots + b_\ell\mathbf{v}_\ell$ for some a_i, b_i. Then

$$U+\mathbf{v} = U+a_1\mathbf{u}_1+\cdots+a_k\mathbf{u}_k+b_1\mathbf{v}_1+\cdots+b_\ell\mathbf{v}_\ell = U+b_1\mathbf{v}_1+\cdots+b_\ell\mathbf{v}_\ell = b_1(U+\mathbf{v}_1)+\cdots+b_\ell(U+\mathbf{v}_\ell).$$

Linear independence. Say $b_1(U + \mathbf{v}_1) + \cdots + b_\ell(U + \mathbf{v}_\ell) = U + \mathbf{0}$. Then $U + b_1\mathbf{v}_1 + \cdots + b_\ell\mathbf{v}_\ell = \mathbf{0}$. Then $b_1\mathbf{v}_1 + \cdots + b_\ell\mathbf{v}_\ell \in U$. Thus there are a_i with $a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k + b_1\mathbf{v}_1 + \cdots + b_\ell\mathbf{v}_\ell = \mathbf{0}$. But then all $a_i = 0$ and $b_i = 0$. □

2.65 Theorem (First isomorphism theorem for vector spaces)**.** *If $\theta : V \to W$ is a linear map, then it induces an isomorphism of vector spaces $\overline{\theta} : V/\ker\theta \to \operatorname{im}\theta$.*

Proof. Same as 2.50 $\square$

2.66 Corollary. *[Rank-nullity formula] If $\theta : V \to W$ is a linear map with V finite-dimensional, then $r(\theta) + n(\theta) = \dim V$.*

Proof. $r(\theta) = \dim \operatorname{im} \theta = \dim V/\ker\theta = \dim V - \dim\ker\theta = \dim V - n(\theta)$. $\square$

2.67 Corollary. *If $\theta : V \to W$ is a linear map with $\dim V = \dim W$, then θ is injective if and only if it is surjective.*

Proof. Surjective $\Leftrightarrow r(\theta) = \dim W \Leftrightarrow r(\theta) = \dim V \Leftrightarrow n(\theta) = 0 \Leftrightarrow$ injective. $\square$

2.68 Proposition. *If $\theta : V \to W$, S is a basis of V and R is a basis of W then $[\theta(\mathbf{v})]_R = A[\mathbf{v}]_S$ for $\mathbf{v} \in V$.*

Proof. If $\mathbf{x} = [\mathbf{v}]_S$, then $\mathbf{v} = \sum_{j=1}^n x_j\mathbf{v}_j$, so

$$\theta(\mathbf{v}) = \theta(\sum_{j=1}^n x_j\mathbf{v}_j) = \sum_{j=1}^n x_j \sum_{i=1}^n a_{ij}\mathbf{w}_i = \sum_{i=1}^n(\sum_{j=1}^n a_{ij}x_j)\mathbf{w}_i = \sum_{i=1}^n (A\mathbf{x})_i\mathbf{w}_i.$$

Thus $[\theta(\mathbf{v})]_R = A\mathbf{x} = A[\mathbf{v}]_S$. $\square$

2.69 Theorem (Change of basis)**.** *Let $\theta : V \to V$ be a linear map from a vector space to itself.*
Let A be the matrix of θ with respect to a basis S of V.
Let A' be the matrix of θ with respect to a basis S' of V.
Then $A' = P^{-1}AP$ where P is the transition matrix from S to S'.

Proof. For $\mathbf{v} \in V$, we have $AP[\mathbf{v}]_{S'} = A[\mathbf{v}]_S = [\theta(\mathbf{v})]_S = P[\theta(\mathbf{v})]_{S'} = PA'[\mathbf{v}]_{S'}$. Since this holds for all $\mathbf{v}$, we must have $AP = PA'$. $\square$

2.70 Theorem. *Let A be an $n \times n$ matrix over F. The following are equivalent:*
(i) A is diagonalizable, meaning that it is similar to a diagonal matrix, so

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

for some invertible matrix $P \in \mathrm{GL}_n(F)$.
(ii) A has n linearly independent eigenvectors (so they form a basis of F^n).
(iii) The characteristic polynomial of A has n roots in F, counted with multiplicity (which always holds if $F = \mathbb{C}$), and for each eigenvalue λ, the geometric multiplicity of λ is equal to the algebraic multiplicity of λ.

Proof. Sketch. (i)⇒(iii) Similar matrices have the same characteristic polynomial, for

$$\chi_{P^{-1}AP}(t) = \det(tI - P^{-1}AP) = \det(P^{-1}(tI - A)P) = \det(tI - A) = \chi_A(t).$$

Thus if (i) holds then $\chi_A(t) = (t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_n)$, so it has n roots in F. Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be the columns of P. Since P is invertible, the set $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis of F^n. Also, letting D be the diagonal matrix of λs, we have $D = P^{-1}AP$, so $AP = PD$. The ith column in this equation gives $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$, so the $\mathbf{v}_i$ are eigenvectors for the λ_i.

(iii)⇒(ii) Combining bases of each of the eigenspaces, one gets n eigenvectors. One can show that this set is linearly independent.

(ii)⇒(i) The matrix of θ_A with respect to this basis of eigenvectors is diagonal. □

2.71 Corollary. *If A is an $n \times n$ matrix, and its characteristic polynomial has n distinct roots in F, then A is diagonalizable.*

Proof. Each eigenvalue has algebraic multiplicity 1, which must therefore also be its geometric multiplicity. □

2.72 Theorem. *Let A be an $n \times n$ matrix. If the characteristic polynomial of A has n roots in F, counted with multiplicity (which always holds if $F = \mathbb{C}$), then A is similar to an upper triangular matrix. Proof not given, but is based up on the following lemma*

2.73 Lemma. *Let A be an $n \times n$ matrix and $\mathbf{v}$ an eigenvector with eigenvalue λ. Extend to a basis $\{\mathbf{v}_1 = \mathbf{v}, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ of F^n, and let P be the matrix with columns $\mathbf{v}_1, \dots, \mathbf{v}_n$. Then $P^{-1}AP$ has block form*

$$\begin{pmatrix} \lambda & * \\ 0 & B \end{pmatrix}$$

where B is an $(n-1) \times (n-1)$ matrix and $*$ is a $1 \times (n-1)$ matrix.

2.74 Proposition. *The determinant of an orthogonal matrix is ± 1.*

Proof. $\det P^T = \det P$, so $P^T P = I$ gives $(\det P)^2 = 1$. □

2.75 Proposition. *A matrix P is orthogonal if and only if its columns are an orthonormal set of vectors.*

Proof. If $\mathbf{v}_i$ is the ith column of P, then $\mathbf{v}_i^T$ is the ith row of P^T, and the (i, j) entry of $P^T P$ is $\mathbf{v}_i^T \mathbf{v}_j$. Thus the set of columns $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is orthonormal if and only if $P^T P = I$. □

2.76 Proposition. *Any orthonormal set of vectors $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly independent.*

Proof. If $a_1\mathbf{v}_1 + \dots + a_k\mathbf{v}_k = \mathbf{0}$ is a linear relation, then for all i we have

$$0 = \mathbf{v}_i \cdot \mathbf{0} = \mathbf{v}_i \cdot (a_1\mathbf{v}_1 + \dots + a_k\mathbf{v}_k) = a_1\mathbf{v}_i \cdot \mathbf{v}_1 + \dots + a_k\mathbf{v}_i \cdot \mathbf{v}_k = a_i.$$

□

2.77 Theorem (Gram-Schmidt process)**.** *Given any linearly independent set $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ in $\mathbb{R}^n$, we can find an orthonormal set $S' = \{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ with the same span. In particular, any subspace of $\mathbb{R}^n$ has an orthonormal basis.*

Proof. We construct

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 \\
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 \\
\mathbf{u}_3 &= \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 \\
&\dots \\
\mathbf{u}_k &= \mathbf{v}_k - \sum_{i=1}^{k-1} \frac{\mathbf{v}_k \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}\mathbf{u}_i
\end{aligned}$$

By construction $\mathbf{u}_j \cdot \mathbf{u}_i = 0$ for all $i < j$. Thus $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is orthogonal. It is easy to see that it is linearly independent and has the same span as S. Now all $\mathbf{u}_i \neq \mathbf{0}$, so we can normalize them, letting $\mathbf{w}_i = \mathbf{u}_i/|\mathbf{u}_i|$, to get an orthonormal set S'. □

2.78 Theorem. *A real symmetric matrix has real eigenvalues, and eigenvectors for distinct eigenvalues are orthogonal. Proof omitted for complexity. See Ommmited Proofs 3.3*

2.79 Lemma. *If a matrix A is a symmetric, then so is $P^{-1}AP$ for P orthogonal.*

Proof. $(P^{-1}AP)^T = P^T A^T (P^{-1})^T = P^{-1}AP$. □

2.80 Theorem. *Any real symmetric matrix A can be diagonalized by an orthogonal matrix, that is, there is an orthogonal matrix P with $D = P^{-1}AP$ diagonal. Equivalently, $\mathbb{R}^n$ has an orthonormal basis consisting of eigenvectors of A. Proof Omitted due to Length, see Ommitted Proofs 3.4*

2.81 Theorem. *If P is an orthogonal matrix, then the map $\theta_P : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{v} \mapsto P\mathbf{v}$ is an isometry. Conversely, any isometry of $\mathbb{R}^n$ that fixes the origin is of this form. Proof Omitted due to Length, see Ommitted Proofs 3.5*

2.82 Lemma. *Every matrix $P \in SO_3(\mathbb{R})$ has 1 as an eigenvalue.*

Proof. It suffices to show that $\det(P - I) = 0$. Recall that $\det P = \det(P^T)$, so $\det(P^T) = 1$. Since P is orthogonal $P^T(P - I) = (I - P)^T$. Then

$$\det(P - I) = \det P^T(P - I) = \det(I - P)^T = \det(I - P).$$

But for any 3×3 matrix, B, $\det(-B) = -\det B$. Thus $\det(P - I) = -\det(P - I)$, so $\det(P - I) = 0$. □

2.83 Theorem. *The rotations of $\mathbb{R}^2$ and $\mathbb{R}^3$ fixing the origin are the linear maps θ_P given by matrices P in $SO_2(\mathbb{R})$ and $SO_3(\mathbb{R})$ respectively. Proof Ommitted due to Length, see Ommitted Proofs 3.6*

2.84 Corollary. *[Euler's Theorem] The composition of two rotations of $\mathbb{R}^3$ about axes through the origin is also a rotation.*

2.85 Theorem. *Every finite subgroup of $SO_3(\mathbb{R})$ is one of the following*
- the group of planar rotations of a regular n-gon ($\cong C_n$)
- the full group of symmetries of a regular n-gon ($\cong D_n$)
- the group of rotations preserving a regular tetrahedron ($\cong A_4$)
- the group of rotations preserving a cube (or regular octahedron) ($\cong S_4$)
- the group of rotations preserving a regular icosahedron (or dodecahedron) ($\cong A_5$)

3 Ommitted Proofs.

Proof. (3.1) Suppose the group is not cyclic, so every element has order 1,2 or 3. Suppose first that there is no element of order 3. Then every element has order 1 or 2, from which it follows that the group is abelian, since $yx = y^{-1}x^{-1} = (xy)^{-1} = xy$. If a and b are distinct elements of order 2 then $\{1, a, b, ab\}$ is a subgroup if G. But this is impossible by Lagrange's Theorem. Thus there is an element of order 3, say r. Then $H = \{1, r, r^2\}$ is a subgroup of G. Let x be an element not in this subgroup. Then $G = H \cup Hx = \{1, r, r^2, x, rx, r^2x\}$. If x has order 3 then we can't have $x^2 \in \{1, x, rx, r^2x\}$, so $x^2 \in \{r, r^2\}$, but then $x = (x^2)^2 \in \{r^2, r^4\} = \{r^2, r\}$, contrary to $x \notin \{1, r, r^2\}$. Thus x has order 2. Similarly rx and r^2x have order 2. Then $rxrx = 1$, so, multiplying on the left by r^2 and on the right by x, one gets $xr = r^2x$. This is enough to fill in the multiplication table for $G = \{1, r, r^2, x, rx, r^2x\}$, giving the same table as D_3. □

Proof. (3.2) (i) There are four long diagonals (through opposite vertices). We number them 1,2,3,4. Consider the map $\theta : G \to S_4$ sending a rotation to the permutation it induces of the long diagonals. This is a homomorphism. The identity rotation is sent to id, and it is easy to see that none of the other rotations are sent to id. (Rotations about a long diagonal are sent to 3-cycles, rotations about an axis through face centres are sent to 4-cycles or products of two transpositions, and rotations about an axis through two edge midpoints are sent to transpositions.) Thus the homomorphism is injective. Therefore $\operatorname{im}\theta \cong G$, so it has 24 elements. Thus we must have $\operatorname{im}\theta = S_4$. (ii) Number the vertices of the tetrahedron 1,2,3,4. Consider the map $\theta : TET \to S_4$ sending any rotation to the induced permutation of the vertices. This is clearly injective, and one can check it has image equal to A_4. □

Proof. (3.3) Suppose that $\lambda \in \mathbb{C}$ is an eigenvalue with associated eigenvector $\mathbf{v} \in \mathbb{C}^n$. We compute $\overline{\mathbf{v}}^T A\mathbf{v}$ in two ways. We have $A\mathbf{v} = \lambda\mathbf{v}$, so $\overline{\mathbf{v}}^T A\mathbf{v} = \overline{\mathbf{v}}^T \lambda\mathbf{v} = \lambda \sum_{i=1}^n |v_i|^2$. On the other hand, starting with $A\mathbf{v} = \lambda\mathbf{v}$, taking the conjugate, and using A real, gives $A\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$. Now taking the transpose and using the fact that A is symmetric, we get $\overline{\mathbf{v}}^T A = \overline{\lambda}\overline{\mathbf{v}}^T$. Thus $\overline{\mathbf{v}}^T A\mathbf{v} = \overline{\lambda}\overline{\mathbf{v}}^T\mathbf{v} = \overline{\lambda} \sum_{i=1}^n |v_i|^2$. Thus $\lambda \sum_{i=1}^n |v_i|^2 = \overline{\lambda} \sum_{i=1}^n |v_i|^2$, so λ is real.

If $A\mathbf{v} = \lambda\mathbf{v}$ and $A\mathbf{w} = \mu\mathbf{w}$, then $\mathbf{v}^T A\mathbf{w} = \mathbf{v}^T\mu\mathbf{w} = \mu\mathbf{v}\cdot\mathbf{w}$. But also $\mathbf{v}^T A$ is the transpose of $A\mathbf{v}$, so it is $\lambda\mathbf{v}^T$, so $\mathbf{v}^T A\mathbf{w} = \lambda\mathbf{v}^T\mathbf{w} = \lambda\mathbf{v}\cdot\mathbf{w}$. Thus $\mu\mathbf{v}\cdot\mathbf{w} = \lambda\mathbf{v}\cdot\mathbf{w}$, so $\mathbf{v}\cdot\mathbf{w} = 0$. □

Proof. (3.4) Take an eigenvalue λ of A. It is real, so has an eigenvector $\mathbf{v}\in\mathbb{R}^n$. We may assume that $|\mathbf{v}| = 1$. We can extend this to a basis $S = \{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ of $\mathbb{R}^n$ where $\mathbf{v}_1 = \mathbf{v}$, and, by the Gram-Schmidt process, we may assume that S is an orthonormal basis of $\mathbb{R}^n$. Let P_0 be the matrix whose columns are the vectors $\mathbf{v}_i$. It is an orthogonal matrix. If $\theta_A : \mathbb{R}^n\to\mathbb{R}^n$ is the linear map $\mathbf{v}\mapsto A\mathbf{v}$, then $P_0^{-1}AP_0$ is the matrix of θ_A with respect the basis S. Since $\theta_A(\mathbf{v}) = \lambda\mathbf{v}$, it takes upper block shape, so

$$P_0^{-1}AP_0 = \begin{pmatrix}\lambda & *\\ 0 & B\end{pmatrix}$$

Since P_0 is an orthogonal matrix, $P_0^{-1}AP_0$ is symmetric. Thus the $*$ term is zero, so the matrix is block diagonal. Now B is symmetric and smaller so by induction there is an orthogonal matrix Q with $C = Q^{-1}BQ$ diagonal. Then

$$Q' = \begin{pmatrix}1 & 0\\ 0 & Q\end{pmatrix}$$

is orthogonal, hence so is $P = P_0Q'$, and

$$P^{-1}AP = (Q')^{-1}(P_0^{-1}AP_0)Q' = \begin{pmatrix}1 & 0\\ 0 & Q^{-1}\end{pmatrix}\begin{pmatrix}\lambda & 0\\ 0 & B\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & Q\end{pmatrix} = \begin{pmatrix}\lambda & 0\\ 0 & C\end{pmatrix},$$

which is diagonal. For the last part, an orthogonal matrix diagonalizes A if and only if its columns are an orthonormal basis of eigenvectors of A. □

Proof. For any $\mathbf{v}\in\mathbb{R}^n$ we have $|P\mathbf{v}| = |\mathbf{v}|$ since

$$|P\mathbf{v}|^2 = (P\mathbf{v})\cdot(P\mathbf{v}) = (P\mathbf{v})^T(P\mathbf{v}) = \mathbf{v}^TP^TP\mathbf{v} = \mathbf{v}^T\mathbf{v} = \mathbf{v}\cdot\mathbf{v} = |\mathbf{v}|^2.$$

Now the distance between $\theta_P(\mathbf{v})$ and $\theta_P(\mathbf{w})$ is

$$|\theta_P(\mathbf{v}) - \theta_P(\mathbf{w})| = |P\mathbf{v} - P\mathbf{w}| = |P(\mathbf{v}-\mathbf{w})| = |\mathbf{v}-\mathbf{w}|.$$

Conversely suppose that θ is an isometry fixing the origin. Then for any $\mathbf{v},\mathbf{w}$ we have $|\theta(\mathbf{v}) - \theta(\mathbf{w})| = |\mathbf{v}-\mathbf{w}|$, so

$$(\theta(\mathbf{v}) - \theta(\mathbf{w}))\cdot(\theta(\mathbf{v}) - \theta(\mathbf{w})) = (\mathbf{v}-\mathbf{w})\cdot(\mathbf{v}-\mathbf{w}).$$

In particular, taking $\mathbf{w} = \mathbf{0}$ and using $\theta(\mathbf{0}) = \mathbf{0}$, we get $\theta(\mathbf{v})\cdot\theta(\mathbf{v}) = \mathbf{v}\cdot\mathbf{v}$. Expanding the displayed formula above, and substituting in this formula and the corresponding one for $\mathbf{w}$, gives

$$\theta(\mathbf{v})\cdot\theta(\mathbf{w}) = \mathbf{v}\cdot\mathbf{w}.$$

Thus $\{\theta(\mathbf{e}_1),\ldots,\theta(\mathbf{e}_n)\}$ is an orthonormal basis of $\mathbb{R}^n$, so the matrix P whose columns are the $\theta(\mathbf{e}_j)$ is orthogonal. Now for any vector $\mathbf{v} = (v_1,\ldots,v_n)$ we can write $\theta(\mathbf{v}) = \lambda_1\theta(\mathbf{e}_1) + \cdots + \lambda_n\theta(\mathbf{e}_n)$ for some scalars $\lambda_1,\ldots,\lambda_n$. Then $\lambda_i = \theta(\mathbf{v})\cdot\theta(\mathbf{e}_i) = \mathbf{v}\cdot\mathbf{e}_i = v_i$, so $\theta(\mathbf{v}) = v_1\theta(\mathbf{e}_1) + \cdots + v_n\theta(\mathbf{e}_n) = P\mathbf{v}$. □

Proof. (3.6) For $\mathbb{R}^2$ this is clear. Sketch for $\mathbb{R}^3$. Say P is in $SO_3(\mathbb{R})$. It has 1 as an eigenvalue. Take an eigenvector of length 1 and extend to an orthonormal basis of $\mathbb{R}^3$. This gives an an orthogonal matrix Q with $Q^{-1}PQ$ having upper triangular block form

$$Q^{-1}PQ = \begin{pmatrix} 1 & * \\ 0 & B \end{pmatrix}.$$

But this matrix is orthogonal (since P and Q are), which implies that the $*$ block must be zero. Now the block B must be in $SO_2(\mathbb{R})$, so a 2×2 rotation matrix. Then the matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & B \end{smallmatrix}\right)$ is a rotation about the axis $(1, 0, 0)^T$, and P is the corresponding rotation in the coordinate system given by the columns of Q^{-1}. $\square$