SUPPLEMENTARY NOTES: CHAPTER 1

1. Groups

A group G is a set with single binary operation which takes two elements $a, b \in G$ and produces a third, denoted ab and generally called their product. (Mathspeak: We have a mapping $G \times G \to G$ in which the image of the pair a, b is denoted by ab.) In general $ab \neq ba$. This multiplication must satisfy the following axioms:

(1) *Associativity*: If $a, b, c \in G$, then $(ab)c = a(bc)$.
(2) *Existence of a unit element*: There exists a "neutral" element, e, usually called the unit element, such that for any $a \in G$ we have $ea = ae = a$.
(3) *Existence of inverses*: For every group element a there exists an element denoted a^{-1} such that $aa^{-1} = e = a^{-1}a$.

The third axiom tacitly implies that the inverse of an element is unique. It is, and the unit element is also unique. These are simple exercises.

If $ab = ba$ for all $a, b \in G$ then the group is called commutative or Abelian (in honor of Niels Henrik Abel, 1802-1829)

The group law allows us to multiply only two elements at a time, so there are two ways to form the product of three elements. The associative law says that these are equal, so we can simply write abc without parentheses for the product.

Exercise: Write down the five ways of forming a product of four elements and prove that they are equal. Now prove (by induction), that all ways of forming a product of n elements are equal for any ln. A product of the form $a_1 a_2 \cdots a_n$ is therefore unambiguous; we do not have to use parentheses. We will see a conceptual proof of this.

An Abelian group is sometimes written in additive form: We write $a + b$ instead of ab, denote the unit element by 0, so $a + 0 = 0 + a = a$ and the (additive) inverse of a by $-a$. When an abelian group is written this way it is sometimes called an additive group. The most important example of an additive group is the ordinary **integers** $\{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ with addition as the operation, usually denoted by $\mathbb{Z}$ (from the German, *Zahl*, number).

Group that we will need to understand when dealing with determinants is the **permutation group** of a set X, denoted Sym_X, but to describe this group adequately we first we need some basic ideas about sets and mappings.

If we have a mapping $f : X \to Y$ from a set X to a set Y and another mapping $g : Y \to Z$, then their composite gf is the combined map

$$X \xrightarrow{f} Y \xrightarrow{g} Z$$

which sends an element $x \in X$ to $g(f(x))$. In the present context we will call this the 'product' of f and g, but don't confuse it with he concept of product when we are dealing with functions. (There the functions have realtor complex values, and by the product function we usually mean $f(x)g(x)$.) The sets X, Y and Z don't have to be different sets; we could be talking about maps from one set X to itself.

Suppose now that we have sets X, Y, Z, W (which don't all have to be different – they could all be the same set X) and mappings $f : X \to Y$, $g : Y \to Z$, $h : Z \to W$. We can picture this by the following diagram

$$X \xrightarrow{f} Y \xrightarrow{g} Z \xrightarrow{h} W.$$

Then the two triple composites $h(gf)$ and $(hg)f$ both mean exactly the same thing, send $x \in X$ to $h(g(f(x)))$. The associative law is immediate and we don't need an parentheses. The same is true for the composite of any number of maps.

A map f from a set X to a set Y (written as $f : X \to Y$) is called one-to-one or 1-1 if distinct elements of X are sent to distinct elements of Y. Such a map is often also commonly called an "injection". A map $f : X \to Y$ is called onto if every $y \in Y$ is the image of some $x \in X$, that is, if there is an $x \in X$ such that $f(x) = y$. Such a map is also commonly called a surjection. If f is both 1-1 and onto, something which is commonly called a bijection, then for every $y \in Y$ there is exactly one $x \in X$ with $f(x) = y$, so we can define a mapping from Y back to X by sending y back to its 'preimage' x. This map is denoted f^{-1}. (Don't confuse this with $1/f$ when one is dealing with functions.) It is the inverse of f in the sense that $f^{-1}f(x) = x$ for all $x \in X$, that is $f^{-1}f$ is the 'identity' map id_X, and $ff^{-1}(y) = y$ for all $y \in Y$, so $ff^{-1} = \mathrm{id}_Y$. Notice that if f has an inverse in this sense then it must be a bijection. Why? A bijection from a set X to itself is called a permutation of X.

The group Sym_X is just the set of all permutations of X. The product of permutations (in the sense in which we are now using product) has just been shown to be associative. There is an identity element in Sym_X, namely id_X, and every $f \in \mathrm{Sym}_X$ has an inverse. Notice that if X is finite then a mapping $f : X \to X$ is 1-1 if and only if (abbreviated 'iff') it is also onto, but in the infinite case this doesn't necessarily hold. Give examples.

Notice that $(fg)^{-1} = g^{-1}f^{-1}$, for if $f(x) = y$ and $g(y) = z$ then to get the inverse operation you must first take z back to y and then y back to x. ("The inverse operation of putting on your jacket and then your coat is first to take off the coat and then the jacket.") This is true in any group, $(ab)^{-1} = b^{-1}a^{-1}$. Notice that if you write out $abb^{-1}a^{-1}$ then first b and b^{-1} cancel and then a and a^{-1} can cancel.

If we have two groups, G_1 and G_2 then a homomorphism (or these days, simply a morphism) from G_1 to G_2 is a map $f : G_1 \to G_2$ which "preserves" or "respects" products. This means that $f(ab) = f(a)f(b)$ for all $A, b \in G_1$. In modern terminology, a homomorphism that is 1-1 is called a monomorphism and one that is onto is an epimorphism.

Exercise: Prove that (1) any morphism $f : G_1 \to G_2$ preserves the unit element, i.e., if e_1 is the unit element of G_1 and e_2 that of G_2, then $f(e_1) = e_2$, and that it also preserves inverses, i.e., that $f(a^{-1}) = f(a)^{-1}$ for all $a \in G_1$.

If $f : G_1 \to G_2$ which is both 1-1 is a group morphism and is also a bijection, then its inverse map is also a group morphism. Prove this. A morphism which has in inverse morphism is called an isomorphism, so this f is an isomorphism. "Morphism" is a very broad concept in mathematics. It means a mapping which preserves whatever kind of structure we are dealing with. In this case it happens to be group structure. A group morphism which is 1-1 and onto as it happens turns out to be an isomorphism. This is also true for morphisms of vector spaces, which we will encounter in Chap. 3. For some complicated kinds of mathematical

structures a morphism which is a bijection has an inverse map whites not necessarily a morphism, but we will not have to deal with such structures in this course.

Isomorphic groups G_1 and G_2 cannot be told apart by any internal structure. They are both sets, and the objects of one may be painted red and the other blue, but corresponding objects multiply in the same way. A mapping between sets which is 1-1 is commonly called an injection, a mapping which is onto is a surjection, one which is both is a bijection. The terminology is originally French and mathematicians have had to adapt to it from the older 1-1 and onto. Notice that if X and Y are finite sets with the same number of elements then there is a bijection between them. In fact, to say that X has n elements means that there is a bijection between X and the set of integers $\{1, 2, \ldots, n\}$. In general, however, the only meaning we can give to the statement that two sets have the same number of elements is that there exists a bijection between them. An infinite set X is called countable if there is a bijection between it and the set $\{1, 2, 3, \cdots\}$ of all positive integers, but there are infinite sets which are so big that this is not possible, for example the set of all real numbers $\mathbb{R}$.

Suppose that we have a bijection $\phi : Y \to X$ between two sets. (Humor me as to the direction.) Then we can construct an isomorphism $Sym_X \to \mathrm{Sym}_Y$ as follows. If $f : X \to X$ is an element of Sym_X then send f to the composite map $\phi f \phi^{-1}$. (That is, if you wan to know where an element $y \in Y$ goes, pull it back to X by the inverse of ϕ, opera on it by f inside X, and then push it forward by ϕ to Y again. It is easy to check that this really is a morphism. Suppose that $f_1, f_2 \in SymX$. We must show that $\phi f_1 f_2 \phi^{-1} = \phi f_1 \phi^{-1} \cdot \phi f_2 \phi^{-1}$, where I have put a $\cdot$ in the middle to make reading the right side easier, but it is clear that the $\phi^{-1} \cdot \phi$ in the middle simply cancel out.

It follows that if X and Y both have n elements (where n is some positive finite number), then they are isomorphic, and both are isomorphic to the symmetric group on the set $\{1, \ldots, n\}$. The latter group is simply called "the" symmetric group or "the" permutation group on n elements and denoted S_n. It will play a major role in the study of determinants. Its size grows very rapidly with n, since it contains $n!$ elements. The number of elements in a group G is commonly called its order and in older texts was usually denoted by $|G|$. Recently, however, $\#G$ has come into favor since the absolute value sign has acquired so many different uses.

The group axioms given above are due to Arthur Cayly (1821-1895) who also made the following fundamental observation. Suppose that G is any group and that $a \in G$. Define a map $f_a : G \to G$by setting $f_a(x) = ax$ for all $x \in G$. It is a permutation of the underlying set of elements of G because it has an inverse, namely $f_{a^{-1}}$/ Check this. The map f_a doesn't respect the multiplication in G, but the map which sends $a \in G$ to $f_a \in \mathrm{Sym}\, G$ is a morphism of G into Sym_G. In fact this is precisely what the associative law says. For if x is an arbitrary element of G, then by definition $fabx = (ab)x$, while $f_a f_b x = a(bx)$, and these are equal, so we have $fab = f_a f_b$. Moreover, the map $a \mapsto f_a$ (read this as "a maps to f_a) is a monomorphism, for if $f_a = f_b$, then, in particular $f_a(1) = f_b(1)$, i.e., $a1 = b1$ or $a = b$. Now the image of G inside Sym_G, denote it by f_G, satisfies the full associative law since all of Sym_G does, one doesn't need any parentheses to make multiplication unambiguous. Therefore, so does G.

This is the sort of abstract reasoning which mathematicians love and which you might not like in this particular instance since the argument by induction is certainly

much simpler to understand. On the other hand, it sometimes works magic as we will see when we define determinants.

2. Rings

Hoffman and Kunze give the formal definition of a ring at the beginning of Chap. 5 on determinants. Read their section 5,1. In brief, a ring is a set K with two operations, addition and multiplication; the sum of x and y is denoted by $x + y$ and their product by xy. Under addition, K is a commutative group. The multiplication is associative but not necessarily commutative. The two operations are linked by the distributive laws,

$$x(y + z) = xy + xz, \quad (y + z)x = yx + zx.$$

If $xy = yx$ for all $X, y \in K$ then K is called commutative. If there is an element 1 in K such that $1x = x1 = x$ for al x then this element is unique Why? and Hoffman and Kunze say that K is a ring with identity. The more recent terminology is to call K a unital ring. HK use the notation K for any ring (but most of the time is commutative). I will use a slightly different convention. An arbitrary ring will be denoted by K, but for a commutative unital ring I will always write $\mathbf{k}$.

The most important example of a commutative unital ring is the ordinary integers, $\mathbb{Z}$. Notice that division is not possible in this ring; the quotient of two integers is not an integer. However, there are two elements in this ring which have inverses in the ring, namely 1 and -1. Elements of a commutative unital ring which do have inverses in the ring are generally called units (but there is only one 'unit element'). In anon-commutative ring they are just called invertible elements. The $n \times n$ matrices with coefficients in a ring K form a non-commutative ring which will be discussed below.

3. Fields

A field $\mathbb{F}$ is a commutative unital ring in which every element other than zero is a unit. Briefly, one can add, subtract, multiply, and divide (except by 0) just as we can with real numbers. The fields with which you may be most familiar are the real numbers, generally denoted by $\mathbb{R}$, and the complex numbers $\mathbb{C}$. But notice that the set of rational numbers, i.e.(not necessarily proper) fractions m/n, where m and n are integers with $n \neq 0$ is already a field, denoted $\mathbb{Q}$ (for quotient), It is a subfield (i.e., a field contained in) $\mathbb{R}$, which in turn is a subfield of C. The field $\mathbb{Q}$ has no subfields other than $\mathbb{Q}$ itself Prove this. but there are many interesting intermediate fields between $\mathbb{Q}$ and $\mathbb{R}$. HK give an example, p. 4 Example 4. of a field which is generally denoted by $\mathbb{Q}(\sqrt{2}$. There are no intermediate fields between $\mathbb{R}$ and $\mathbb{C}$. This will be very easy to prove once we have introduced the concept of the dimensions a vector space, but if you already know what that is, you may turn in the proof for extra credit.

As HK point out, p.3, there are fields which do not behave at all like subfields of $\mathbb{C}$, and in fact they may have only a finite number elements. These are fundamental in an area of mathematics called Number Theory. They were discovered by one of brightest mathematical minds of all time, Évariste Galois (1811-1831) who died as a result of a duel over a prostitute at the age of 20. The smallest of these "Galois" fields contains only two elements, 0 and 1 with the rule $1 + 1 = 0$. (This field is generally denoted $\mathbb{F}_2$.) We do not need to know about such fields in this course,

where you may assume that all our fields are subfields of $\mathbb{C}$ (including $\mathbb{C}$ itself, which will be very important), but it may at some point be useful for you to know that they exist.

4. Matrix rings

An important class of non-commutative rings comes from the multiplication of matrices with coefficients in a ring K. For many purposes K can be an arbitrary ring, not necessarily commutative, so for the moment (unlike HK) we make no assumptions about K. An $m \times n$matrix A with coefficients in K is a rectangular array of elements of K arranged in m rows and n columns. The element in the (i, j) position is denoted A_{ij}. (HK are a little more careful, see p. 141.) A $1 \times n$ matrix is called a row vector of length n; an $n \times 1$ matrix is a column vector of length n. (Later, when K is a field, see below, we may say 'dimension' n instead of length.) Suppose now that $\boldsymbol{\alpha}$ is a row vector of length n and that $\boldsymbol{\beta}$ is a column vector of the same length n, both with coefficients in K, so

$$\boldsymbol{\alpha} = (a_1, \dots, a_n), \quad \boldsymbol{\beta} = \begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix}, \quad \text{where} \quad a_1, \dots, a_n, b_1, \dots, b_n \in K.$$

Their product is then defined to be

$$\boldsymbol{\alpha}\boldsymbol{\beta} = a_1 b_+ a_2 b_2 + \cdots + a_n b_n = \sum_{k=1}^{n} a_k b_k,$$

using the standard summation notation. Denote the ith row of the $m \times n$ matrix A by $A_{i,\bullet}$, a row vector length n. Suppose now that B is an $n \times p$ matrix and denote its j column by $B_{\bullet,j}$, a column vector of also of length n. Then the product AB is defined by setting

$$(AB)_{ij} = A_{i,\bullet} B_{\bullet,j} = \sum_{k=1}^{n} A_{ik} b_{kj}.$$

That is, the (i, j) entry in AB is the product of the ith row of A with the jth column of B. Notice that AB is an $m \times p$ matrix. This multiplication is associative in the following sense. If A is an $m \times n$ matrix, B an $n \times p$ matrix and C a $p \times q$ matrix then $(AB)C$ and $A(BC)$ are both well-defined $m \times q$ matrices and we have $(AB)C = A(BC)$. HK prove this on p. 19 under the assumption that K is a field, because we will need that when row=reducing matrices, but the proof actually requires no more than that K be a ring.

We can add and subtract matrices if they have the same dimensions: If A and B are both $m \times n$ then $(A \pm B)_{ij} = A_{ij} \pm B_{ij}$ and there is a neutral element, the $m \times n$ matrix all of whose entries are 0. This matrix is also frequently denoted simply by 0 without specifying its dimensions. The set of all $m \times n$ matrices with coefficients in K, denoted $K^{m,n}$ is therefore an additive group.

Consider now the set $K^{n,n}$ of square $n \times n$ with coefficients in K. Since we can form products of these, it has both an addition and a multiplication and the distributive laws hold Prove this. so it is a ring. If K is unital, then this ring also has a unit element denoted I or I_n when we want to be explicit about its size, which is the matrix whose entries on the main diagonal are all equal to one and all

whose other entries are 0. In particular, $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Another way to describe it is with the very useful Kronecker delta function, defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

Here i and j are integers between 1 and n. The (i, j) entry of I is just δ_{ij}.

All the non-zero entries of I are on the diagonal. Any matrix having non-zero entries only on the diagonal is called a diagonal matrix. An $n \times n$ diagonal matrix with diagonal entries $c_1, \ldots, c_n$ is denoted $\mathrm{diag}(c_1, \ldots, c_n)$. The add and multiply in a simple way:

$$\mathrm{diag}(c_1, \ldots, c_n) + \mathrm{diag}(c'_1, \ldots, c'_n) = \mathrm{diag}(c_1 + c'_1, \ldots, c_n + c'_n) \tag{1}$$

$$\mathrm{diag}(c_1, \ldots, c_n)\,\mathrm{diag}(c'_1, \ldots, c'_n) = \mathrm{diag}(c_1 c'_1, \ldots, c_n c'_n). \tag{2}$$

Notice (from the second line) that diagonal matrices commute, and that if none of the c_i is zero then the inverse of $\mathrm{diag}(c_1, \ldots, c_n$ is $\mathrm{diag}(c_1^{-1}, \ldots, c_n^{-1})$. If A is an $m \times n$ matrix, then $diag(c_1, \ldots, c_m)A$ is obtained by multiplying the ith row of A by c_i, and $A\,\mathrm{diag}(c'_1, \ldots, c'_n)$ is obtained by multiplying the jth column of A by c'_j. (I have tried to be careful about the indices.)

Elementary matrices. Here, in more explicit form are the elementary matrices, cf p. 20 of HK, corresponding to the elementary row operations on p.6 1. If $c \neq 0$ then $\mathrm{diag}(1, \ldots, 1, c\,(i\text{th place}), 1, \ldots, 1)$ is the elementary matrix which, when multiplied from the left, multiplies the ith row of A by c, and leaves the others unchanged. Its inverse is obtained by replacing c by c^{-1}. 2. The matrix whose entries are all 0 except for a 1 in the (i, j) place is denoted by e_{ij}. (Its dimensions are usually unmentioned and determined by the context.) The elementary matrix $(I + ce_{ij})$ by adds c times the jth row of A to its ith row. Its inverse is $I - ce_{ij}$. 3. The elementary matrix which interchanges the ith and jth rows of A (by multiplication from the left) is obtained from the identity matrix I by interchanging the ith and jth rows of I. The elementary row operations on A can therefore all be described as multiplication from the left by corresponding elementary matrices.

A permutation which interchanges two elements of a set X, leaving all the others fixed, is called a transposition. If we have a transposition t and do it twice then nothing has moved, t^2 is just the identity element of Sym_X; the inverse of a transposition is itself. It follows follows that if we have a product of transpositions, $t_1 t_2 \cdots t_k$ then its inverse is the same product in reverse order, $(t_1 t_2 \cdots t_k)^{-1} = t_k \cdots t_2 t_1$. When X is the set of integers $\{1, \ldots, n\}$, then the transposition which interchanges i and j is denoted simply by (i, j), and the identity elements denoted by e.

We will need the following convention: a product of zero elements in a group is the identity element of the group. (The reason is that if you multiply a product of k elements by a product of zero elements you should still just have the original product of k elements.)

Theorem 1. *If X is finite then every element of Sym_X is a product of transpositions.*

Proof. The proof will be by induction on the number of elements in X. If this is n then we can assume, wlog (without loss of generality) that X is just

the set of integers $\{1, \ldots, n\}$, because Sym_X is isomorphic to S_n. There is noting to prove when $n = 1$ since S_1 contains only the identity. The theorem is trivial when$n = 2$ since S_2 consists only of e and (1,2). So we may assume that $n > 3$ and suppose that the theorem is true for $n - 1$. Let σ be an arbitrary element of S_n. We wish to show that σ is a product of transpositions. Suppose that $\sigma(n) = n$ then σ is really only permuting $1, \ldots, n-1$, so we know by the induction hypothesis that σ is a product of transpositions. So suppose that $\sigma(n) =$ where $m \neq n$. Then $(m, n)\sigma$ leaves n fixed (since σ carried n to m and (m, n) carried it back to n. Therefore $(m, n)\sigma$ is a product of transpositions, $t_1 t_2 \cdots t_k$. But then $\sigma = (m, n)(m, n)\sigma = 9m, n)t_1 t_2 \cdots t_k$, a product of transpositions. □

So every permutation of $\{1, \ldots, n\}$ can be written as a product of transpositions. This can generally be done in more than one way, but we will prove later that if some permutation σ can be written as a product of an even number of transpositions, then any expression for σ as a product of transpositions will always have an even number of factors, and similarly for odd. Permutations are therefore accordingly called even or odd. A single transposition is odd.

A permutation matrix P is one which is obtained from the identity by a permutation of the rows of the identity matrix. If this permutation is σ, then PA is the matrix obtained from A by permuting its rows according to σ. Prove this. Prove that any permutation matrix canals be obtained by permuting the columns of I, and that you get the same matrix if you permute the rows of I by σ or its columns by $(\sigma)^{-1}$.