

# NOTES ON LIE GROUPS AND ALGEBRAS

ALEX NELSON

## CONTENTS

1. Introduction	1
2. Tangent Vectors	1
3. An Introduction to Infinitesimals	2
3.1. A Brief Aside on Complex Numbers	2
3.2. And Now For Something Slightly Different	2
4. Tangent Vectors and Infinitesimals Together	3
5. Matrix Groups	3
6. Lie Algebras	5
6.1. Lie Bracket	6
Appendix A. A Few Remarks On Groups	7
A.1. Homomorphisms and Isomorphisms	8
A.2. Centralizers, Normalizers, Stabilizers, etc.	9
Appendix B. A Review of Matrices	10
B.1. Matrix Operations	10
B.2. Euclidean and Einstein Summation Conventions	11
B.3. Transposes of Matrices	12
References	13

## 1. INTRODUCTION

We follow an approach similar to Artin [Art91].

## 2. TANGENT VECTORS

**Definition 1.** Let  $S \subseteq \mathbb{R}^n$  be a set. If our set  $S$  is the locus of zeros of one or more polynomial functions  $f(x_1, \dots, x_n)$ , it is called a **real algebraic set**:

$$(1) \quad S = \{x \in \mathbb{R}^n : f(x) = 0\}.$$

**Example 2.** The set  $S \subset \mathbb{R}^3$  defined by the zeros of the polynomial

$$(2) \quad f(x, y, z) = x^2 + y^2 + z^2 - 1$$

defines a sphere of radius 1. QEF.

**Example 3.** The set  $S \subset \mathbb{R}^3$  defined by the zeros of the polynomial

$$(3) \quad f(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1$$

(where  $a, b, c \in \mathbb{R}^+$ ) defines an ellipsoid. QEF.

**Lemma 4.** Let  $S$  be a real algebraic set in  $\mathbb{R}^n$ , defined as the locus of zeros of one or more polynomial functions  $f(x)$ . The tangent vectors to  $S$  at  $x$  are orthogonal to the gradients  $f(x)$ .

*Algebraic Sets often taught as “surfaces” in calculus courses in universities*

## 3. AN INTRODUCTION TO INFINITESIMALS

**3.1. A Brief Aside on Complex Numbers.** Recall with complex numbers (the set  $\mathbb{C}$ ) we are dealing with a new “number”

$$(4) \quad i^2 + 1 = 0 \quad \Rightarrow \quad i = \sqrt{-1}$$

and so we have a new number system with elements of the form

$$(5) \quad (x + iy) \in \mathbb{C} \quad \text{where } x, y \in \mathbb{R}.$$

We can formally define multiplication using the polynomial multiplication:

$$(6) \quad (a + ib)(x + iy) = (ax - by) + i(ay + bx)$$

which results in a complex number.

**3.2. And Now For Something Slightly Different.** Suppose instead of Eq (4) we worked with

$$(7) \quad \varepsilon^2 = 0 \quad \text{and} \quad \varepsilon \neq 0$$

which means we have elements called **infinitesimals** of the form

$$x + \varepsilon y.$$

We can formally define multiplication in the same way we did for complex numbers

$$(8) \quad (a + b\varepsilon)(x + y\varepsilon) = ax + (xb + ya)\varepsilon + by\varepsilon^2 \overset{0}{\leftarrow}$$

Note that  $\varepsilon^2 = 0$  is the defining characteristic of our infinitesimal element!

Observe what mayhem we may do with this. We have some favorite Taylor series

$$(9) \quad f(x + h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!} h^n$$

which becomes with infinitesimals

$$(10) \quad f(x + \varepsilon) = f(x) + f'(x)\varepsilon + \mathcal{O}(\varepsilon^2) \overset{0}{\leftarrow}$$

Similarly for

$$(11) \quad \frac{1}{1+x} = \sum_{n=0}^{\infty} (-x)^n = 1 - x + x^2 - \dots$$

we get

$$(12) \quad \frac{1}{1+\varepsilon} = 1 - \varepsilon.$$

For the famous euler’s constant

$$(13) \quad e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

we get

$$(14) \quad e^\varepsilon = 1 + \varepsilon + \mathcal{O}(\varepsilon^2/2) \overset{0}{\leftarrow}$$

For trigonometric functions, which can be obtained by Euler’s identity

$$(15) \quad e^{ix} = \cos(x) + i \sin(x)$$

we deduce both

$$(16) \quad \cos(\varepsilon) = 1 + \mathcal{O}(\varepsilon^2) \overset{0}{\leftarrow}$$

and

$$(17) \quad \sin(\varepsilon) = \varepsilon + \mathcal{O}(\varepsilon^3)^{\nabla 0}.$$

Thus we find that in this instance

$$(18) \quad \tan(\varepsilon) = \frac{\sin(\varepsilon)}{\cos(\varepsilon)} = \varepsilon/1$$

which is remarkable!

#### 4. TANGENT VECTORS AND INFINITESIMALS TOGETHER

Recall Lemma (4) that tangent vectors are orthogonal to  $\nabla f(x)$ . Given a point  $x$  of  $\mathbb{R}^n$  and a vector  $\vec{v} \in \mathbb{R}^n$ , the sum  $x + \vec{v}\varepsilon$  is a vector with infinitesimal entries which we think of intuitively as *an infinitesimal change in  $x$  in the direction of  $\vec{v}$* . Observe the Taylor expansion of  $f(x + \vec{v}\varepsilon)$  yields by Eq (10)

$$(19) \quad f(x + \vec{v}\varepsilon) = f(x) + \varepsilon \vec{v} \cdot \nabla f(x)$$

since higher order terms vanish by our use of infinitesimal generator  $\varepsilon$ .

We want to find tangent vectors to some point  $x \in \mathbb{R}^n$ , that is vectors such that

$$(20) \quad \vec{v} \cdot \nabla f(x) = 0.$$

If we demand that  $x \in S$  is part of the algebraic set, then we necessarily have

$$(21) \quad f(x) = 0$$

since we specifically defined algebraic sets to be the sets of zeros of polynomials.

If we notice these two things, we have

$$(22) \quad f(x + \vec{v}\varepsilon) = f(x) + \varepsilon \vec{v} \cdot \nabla f(x) = \varepsilon \vec{v} \cdot \nabla f(x)$$

since by Eq (21) we have  $f(x) = 0$  we are left with the right hand side. Since we want to find all such vectors  $\vec{v}$  that are orthogonal to the gradient of  $f$ , we then can find it through the relationship

$$(23) \quad f(x + \vec{v}\varepsilon) = 0$$

instead of solving a differential system of equations. We will be using this property a lot to find tangent vectors.

Here we must reiterate the beauty of this system. To find tangent vectors, we simply are looking for vectors  $\vec{v}$  such that

$$(24) \quad f(x + \varepsilon \vec{v}) = 0$$

where  $f(x) = 0$  generates an algebraic set  $S$ . By Taylor expanding, we see that

$$(25) \quad f(x + \varepsilon \vec{v}) = f(x)^{\nabla 0} + \varepsilon \vec{v} \cdot \nabla f(x)$$

and moreover, if this is zero then we have found the tangent vectors to the point  $x \in S$ !

#### 5. MATRIX GROUPS

If the reader is unfamiliar with either group theory or matrices, it is advised to refer to the appropriate appendix first. Whenever we discuss matrices, it is assumed to be a square matrix.

Allow us to first introduce the *general linear group*. If we have two matrices which are invertible  $A$  and  $B$ , then their product is invertible. We know this because if  $\det(A) \neq 0$  then  $A$  is invertible, and

$$(26) \quad \det(AB) = \det(A) \det(B).$$

Thus the only way that  $AB$  is not invertible would be if

$$(27) \quad \det(AB) = 0$$

but

$$(28) \quad \det(A) \neq 0 \quad \text{and} \quad \det(B) \neq 0$$

which means that Eq (27) is no longer true. Thus we have a contradiction.

We therefore conclude the product of two invertible matrices is invertible. We have a group! The identity element is the identity matrix, and the binary operation is matrix multiplication. This group is called the **general linear group** of  $n \times n$  matrices over the field  $\mathbb{F}$ , or more succinctly just  $GL_n(\mathbb{F})$ . We usually have  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{C}$ .

*Meaning of “special”*

We can then normalize the matrices by a simple trick. If we have an  $n \times n$  matrix  $A$  with a determinant not equal to one, we can simply map it to

$$(29) \quad A \mapsto \frac{1}{\sqrt[n]{\det(A)}} A$$

which preserves the property of unit determinant. The subgroup of  $GL_n(\mathbb{F})$  with this extra condition is called the **special linear group** and denoted  $SL_n(\mathbb{F})$ .

The next property of matrices that we can investigate would be orthogonality. That is, when we diagonalize a matrix  $X$ , we end up multiplying it by  $AXA^{-1}$ . The property of orthogonal matrices that we are interested in simply is

$$(30) \quad A^T = A^{-1}.$$

The subgroup of  $GL_n(\mathbb{F})$  that satisfies such a property is the **orthogonal group** denoted as  $O_n(\mathbb{F})$ . And similarly, the subgroup of  $SL_n(\mathbb{F})$  that satisfies the orthogonality property is the **special orthogonal group** denoted  $SO_n(\mathbb{F})$ .

The favorite property of quantum physicists, unitarity (“self-adjointness”), also forms a group. The basic property is this, for an  $n \times n$  matrix  $A$  with complex entries, we have

$$(31) \quad \bar{A}^T = A^{-1}$$

the complex conjugate of the transpose is the inverse of  $A$ . This forms the **unitary group** denoted as  $U_n(\mathbb{C})$ . We can form the **special unitary group** by the normalization routine outlined above, and we denote this group as  $SU_n(\mathbb{C})$ .

The most bizarre condition on a  $2n \times 2n$  matrix  $M$  (note even dimensions!) is that of being symplectic. What does it mean anyways? It means that  $M$  satisfies

$$(32) \quad M^T \Omega M = \Omega$$

where (typically)

$$(33) \quad \Omega = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

What is the consequence of this? Well, it’s a bit stronger than orthogonality because

$$(34) \quad M^{-1} = \Omega^{-1} M^T \Omega$$

holds. We can form  $M$  into block matrix form where

$$(35) \quad M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where  $A, B, C, D$  are all  $n \times n$  matrices satisfying the following properties:

$$(36a) \quad A^T D - C^T B = I$$

$$(36b) \quad A^T C = C^T A$$

$$(36c) \quad D^T B = B^T D.$$

The symplectic property of matrices is not as well known as the orthogonality condition or self-adjointness.

We have this neat little table of matrix groups:

$GL_n(\mathbb{R})$	all invertible matrices with real entries
$SL_n(\mathbb{R})$	all matrices in $GL_n(\mathbb{R})$ with determinant 1
$O_n(\mathbb{R})$	all matrices with their inverse equal to their transpose
$SO_n(\mathbb{R})$	all matrices in $O_n(\mathbb{R})$ with determinant 1
$U_n$	all unitary $n \times n$ matrices
$SU_n$	all matrices in $SL_n(\mathbb{C})$ and in $U_n$ .
$SP_{2n}(\mathbb{R})$	$P \in GL_{2n}(\mathbb{R})$ such that $P^T J P = J$ for a given antisymmetric matrix $J$ (usually the one in Eq (33))

## 6. LIE ALGEBRAS

The notion of a Lie algebra is fairly straightforward, but to reiterate we just generalized the notion of a tangent vector to a point  $x$  in an algebraic set  $S$  generated by the zeros of the function  $f$  to be the vector  $\vec{v}$  such that

$$(37) \quad f(x + \varepsilon \vec{v}) = 0.$$

We will use this notion to find the “tangent vectors” of the matrix groups tangent to the “point”  $I$  (the identity matrix).

**Example 5.** We have certain conditions on the groups which we have outlined in function form, e.g. the orthogonal group requires that

$$(38) \quad A^T = A^{-1}.$$

We have the tangent vectors then satisfy

$$(39) \quad (I + \varepsilon V)^T = (I + \varepsilon V)^{-1}$$

which reduces to

$$(40) \quad \varepsilon V^T = -\varepsilon V.$$

How did we do this step? Well, with matrices if the spectral radius is less than  $\pm 1$ , then we can do the following

$$(41) \quad (I - X)^{-1} = I + X + \frac{X^2}{2!} + \dots$$

and since  $\varepsilon^2 = 0$  we have it only to first order. Thus

$$(42) \quad (I + \varepsilon V)^{-1} = I - \varepsilon V$$

and we are content.

This implies that

$$(43) \quad \varepsilon(V^T + V) = 0$$

and since

$$(44) \quad \varepsilon \neq 0$$

we have *necessarily*

$$(45) \quad (V^T + V) = 0 \Rightarrow V^T = -V.$$

Thus we are dealing with traceless matrices. QEF.

It often occurs that the same infinitesimal tangents can correspond to different groups. For example, for the special linear group we have the condition

$$(46) \quad \det(X) = 1$$

but we have the infinitesimal tangents determined by

$$(47) \quad \det(I + \varepsilon A) = 1.$$

We know for “small”  $\varepsilon$  that

$$(48) \quad \det(I + \varepsilon A) = 1 + \varepsilon \operatorname{tr}(A)$$

and to have this be equal to one we need

$$(49) \quad \operatorname{tr}(A) = 0$$

which is satisfied by Eq (45).

**Theorem 6.** Let  $A$  be a real,  $n \times n$  matrix. Then the following are equivalent

- (1)  $\operatorname{tr}(A) = 0$ ;
- (2)  $\exp(tA)$  is a one-parameter subgroup of  $SL_n(\mathbb{R})$ ;
- (3)  $A$  is in the Lie Algebra of  $SL_n(\mathbb{R})$ ;
- (4)  $A$  is an infinitesimal tangent to  $SL_n(\mathbb{R})$  at  $I$ .

Note note note that the infinitesimal tangents includes as a subset the Lie Algebras. We denote the Lie Algebra of a Matrix Group  $G$  as  $\operatorname{Lie} G$ . We also denote the infinitesimal tangents of  $G$  by  $\operatorname{Inf}(G)$ . We denote the one-parameter subgroup of  $G$  as  $\operatorname{Exp}(G)$ . We have

$$(50) \quad \operatorname{Exp}(G) \subset \operatorname{Lie}(G) \subset \operatorname{Inf}(G).$$

So we have the most general set being the tangents of  $I$  in  $G$ .

**6.1. Lie Bracket.** The Lie Algebra of a linear group has additional structure, an operation called a **Lie Bracket**. We formally induce a Lie Bracket on these collections of matrices by the expression

$$(51) \quad [A, B] = AB - BA$$

where  $A$  and  $B$  are matrices. This satisfies the **Jacobi Identity**

$$(52) \quad [A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

We see also that if  $A, B$  are skew-symmetric (antisymmetric), then

$$(53) \quad [A, B]^T = (AB - BA)^T = B^T A^T - A^T B^T = BA - AB = -[A, B].$$

That is, it results in a skew-symmetric matrix.

The bracket is important because it is the infinitesimal version of the commutator in Lie groups  $PQP^{-1}Q^{-1}$ . To see this, we introduce two infinitesimal parameters  $\epsilon$  and  $\delta$  such that

$$(54) \quad \epsilon^2 = \delta^2 = 0$$

and

$$(55) \quad \epsilon\delta = \delta\epsilon.$$

Recall that the inverse of  $I + A\epsilon$  is  $I - A\epsilon$ . So if  $P = I + A\epsilon$  and  $Q = I + B\delta$ , the commutator expands to

$$(56) \quad (I + A\epsilon)(I + B\delta)(I - A\epsilon)(I - B\delta) = I + (AB - BA)\epsilon\delta.$$

Intuitively we can see this because the commutator of infinitesimal elements in  $G$  is still in  $G$ !

**Definition 7.** A **Lie Algebra**  $V$  over a field  $\mathbb{F}$  is a vector space together with a law of composition

$$\begin{aligned} V \times V &\rightarrow V \\ v, w &\mapsto [v, w] \end{aligned}$$

called the **bracket** having the following properties:

- (1) (Bilinearity)  $[v_1 + v_2, w] = [v_1, w] + [v_2, w]$ ,  $[cv, w] = c[v, w]$ ,  $[v, w_1 + w_2] = [v, w_1] + [v, w_2]$ ,  $[v, cw] = c[v, w]$ ;
- (2) (Antisymmetry)  $[v, v] = 0$ ;
- (3) (Jacobi Identity)  $[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$ ;

for all  $u, v, w \in V$  and all  $c \in \mathbb{F}$ .

## APPENDIX A. A FEW REMARKS ON GROUPS

We will follow Dummit and Foote [DF04] here when introducing groups to be as general as possible.

**Definition 8.** A **binary operation**  $\star$  on a set  $G$  is a function

$$(57) \quad \star : G \times G \rightarrow G.$$

For any  $a, b \in G$  we shall write  $a \star b$  instead of  $\star(a, b)$ .

**Example 9.** The operator of addition defined on integers is a binary operation.

**Definition 10.** A binary operation  $\star$  on  $G$  is **associative** if for all  $a, b, c \in G$  we have

$$(58) \quad a \star (b \star c) = (a \star b) \star c.$$

**Remark 11.** Associativity means we can screw around with the order of evaluation, or more to the point we can screw around with parentheses as we please. Most things (the vast majority of things) are associative in math.

**Definition 12.** If  $\star$  is a binary operation on a set  $G$  we say elements  $a, b$  of  $G$  **commute** if

$$(59) \quad a \star b = b \star a.$$

We say  $\star$  (or  $G$ ) is **commutative** if this holds for all  $a, b \in G$ .

**Remark 13.** Let  $\star$  be a binary operation defined on  $G$  and  $H \subset G$  be a subset. If for all  $a, b \in H$  we have  $a \star b \in H$ , then  $H$  is said to be **closed** under  $\star$ . Observe further that if  $\star$  is associative on  $G$ , then necessarily it is associative on  $H$ .

**Definition 14.** A **group** is an ordered pair  $(G, \star)$  where  $G$  is a set and  $\star$  is a binary operation on  $G$  satisfying:

- (1)  $(a \star b) \star c = a \star (b \star c)$  for all  $a, b, c \in G$ , i.e.  $\star$  is associative;
- (2) There exists an element  $e \in G$  called the **identity** of  $G$  such that for all  $a \in G$  we have  $a \star e = e \star a = a$ ;
- (3) For each  $a \in G$  there is an element  $a^{-1} \in G$  called the **inverse** of  $a$  such that  $a \star a^{-1} = a^{-1} \star a = e$ .

**Definition 15.** Let  $(G, \star)$  be a group. If

$$(60) \quad a \star b = b \star a$$

holds for all  $a, b \in G$ , then the group is called **abelian**.

**Remark 16.** Abelian merely means the group has a commutative binary operation, or equivalently the group is commutative.

**Theorem 17.** Let  $G$  be a group with the operation  $\star$ , then

- (1) The identity of  $G$  is unique;
- (2) For each  $a \in G$ ,  $a^{-1}$  is unique;
- (3)  $(a^{-1})^{-1} = a$  for all  $a \in G$ ;
- (4)  $(a \star b)^{-1} = (b^{-1}) \star (a^{-1})$
- (5) For any  $a_1, \dots, a_n \in G$ , the expression  $a_1 \star \dots \star a_n$  is independent of how the expression is bracketed (“generalized associativity law”).

*Change of Notation*

Note that the notation  $a \star b$  gets tedious to write, so we instead just write  $a \cdot b$  or  $ab$ .

**Theorem 18.** Let  $G$  be a group and  $a, b \in G$ . The equations  $ax = b$  and  $ya = b$  have unique solutions for  $x, y \in G$ . In particular, the left and right cancellation laws hold in  $G$ , i.e.

- (1) if  $au = av$ , then  $u = v$ , and
- (2) if  $ub = vb$ , then  $u = v$ .

**Definition 19.** For  $G$  a group and  $x \in G$ , the **order** of  $x$  is defined to be the smallest positive integer  $n$  such that

$$(61) \quad x^n = 1$$

and denote this integer by  $|x|$ . In this case,  $x$  is said to be of order  $n$ . If no positive power of  $x$  is the identity, the order of  $x$  is defined to be infinity and  $x$  is said to be of infinite order.

This last definition is useful only when dealing with finite groups, which we may possibly deal with.

**Definition 20.** Let  $G = \{g_1, \dots, g_n\}$  be a finite group with  $g_1 = e$ . The **multiplication table** or **group table** of  $G$  is the  $n \times n$  matrix whose  $i, j$  entry is the group element  $g_i g_j$ .

**Example 21.** Let our group be  $\mathbb{Z}_2 = \{0, 1\}$  with the usual multiplication. Then the multiplication table is

$$(62) \quad \begin{array}{c|cc} \times & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

**Definition 22.** Let  $G$  be a group. If  $H \subset G$  is closed under the binary operator  $\star$ , and  $H$  satisfies all the axioms of a group, then  $H$  is called a **subgroup**.

**Definition 23.** The binary operation on  $G$  defined by the map  $G \times G \rightarrow G$  defined by

$$(63) \quad (g, x) \mapsto gxg^{-1}$$

is called **conjugation**.

**Definition 24.** A subgroup  $N$  of  $G$  is called a **normal subgroup** if it has the following property: For every  $a \in N$  and every  $b \in G$ , the conjugate  $bab^{-1}$  is in  $N$ .

#### A.1. Homomorphisms and Isomorphisms.

**Definition 25.** Let  $(G, \star)$  and  $(H, \diamond)$  be groups. A map  $\varphi : G \rightarrow H$  such that

$$(64) \quad \varphi(x \star y) = \varphi(x) \diamond \varphi(y)$$

(for all  $x, y \in G$ ) is called a **homomorphism**.



**Example 26.** The map

$$(65) \quad \{x^i, p_j\} = \delta_j^i \mapsto [\hat{x}^i, \hat{p}_j] = i\hbar\delta_j^i$$

used in quantization of classical systems is an example of a homomorphism.

**Definition 27.** A Homomorphism  $\varphi : G \rightarrow H$  is a **isomorphism** if and only if  $\varphi$  is bijective. We say  $G$  and  $H$  are **isomorphic** and denote it as  $G \cong H$ .

**Remark 28.** An isomorphism is a homomorphism that can be undone. We can think of it as a sort of equivalence relation. That is, if  $G \cong H$  then we can intuitively think of  $G$  and  $H$  as being “equivalent”.

**Definition 29.** The **kernel of a linear operator**  $T$  is the set of all arguments which are mapped to zero, i.e. the additive identity.

**Definition 30.** The **kernel of a homomorphism**  $\varphi : G \rightarrow G'$  is the set of elements of  $G$  mapped to the identity in  $G'$ :

$$(66) \quad \text{Ker}(\varphi) = \{a \in G : \varphi(a) = 1\}$$

which can be described as  $\varphi^{-1}(1)$  – the inverse image of the identity. The kernel is a subgroup because (for  $a, b \in G$ )

$$(67) \quad \varphi(ab) = \varphi(a)\varphi(b) = 1 \cdot 1 = 1$$

thus  $ab \in \text{Ker}(\varphi)$ .

## A.2. Centralizers, Normalizers, Stabilizers, etc.

**Definition 31.** Let  $H$  be a subgroup of  $G$ . A **left coset** is a subset of the form

$$(68) \quad aH = \{ah : h \in H\}.$$

Observe that the subgroup  $H$  itself is a coset, i.e.  $H = 1H$ .

**Example 32.** If we have a collection of vectors over the field  $\mathbb{F}_2$  such that they all satisfy some system of equations (they are eigenvectors of a given operator), then we may construct a coset by adding some constant vector  $\vec{v}$  to all of these vectors.

**Corollary 33.** The left cosets of a subgroup partition the group.

It is customary to denote the set of cosets of a normal subgroup  $N$  of  $G$  by the notation

$$(69) \quad G/N = \text{set of cosets of } N \text{ in } G.$$

We will also use the symbols  $\bar{G} = G/N$ .

We can now study a set  $S$  on which the group  $G$  operates, then decompose  $S$  into “orbits” which we will now define.

**Definition 34.** The **orbit** of  $s \in S$  is the set

$$(70) \quad O_s = \{gs : g \in G\}.$$

It is a subset of  $S$ .

**Definition 35.** The **stabilizer** of an element  $s \in S$  is the subgroup  $G_s$  of  $G$  of elements leaving  $s$  fixed

$$(71) \quad G_s = \{g \in G : gs = s\}.$$

**Definition 36.** The stabilizer of an element  $x \in G$  for the operation of conjugation is called the **centralizer** of  $x$  and is denoted  $Z(x)$ :

$$(72) \quad Z(x) = \{g \in G : gxg^{-1} = x\} = \{g \in G : gx = xg\}.$$

The centralizer is the set of group elements which commute with  $x$ .

## APPENDIX B. A REVIEW OF MATRICES

**B.1. Matrix Operations.** There are two main matrix operations: matrix addition and matrix multiplication. Both are interesting.

First a digression. When working with matrices, we usually have the components represented by a value. We also want to be as general as possible, so these values are represented by variables. But the English language only has 26 letters...so anything beyond a 5 by 5 matrix is beyond hope of representation if we just use  $a, b, \dots, x, y, z$ . What to do? We will use **indices** to indicate what entry we are talking about! So each matrix has only one letter, and two indices. For example, a 2 by 2 matrix could be represented by

$$(73) \quad A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

*Indices are dummy variables*

where the  $i$  index tracks the **rows** and the  $j$  index tracks the **columns**. **NOTE NOTE NOTE** that the indices are *dummy variables*! This means that  $b_{ij}$  can be rewritten as  $b_{mn}$  (or with any other pair of letters one would like) and it would refer to the same thing.

So a 3 by 3 matrix can be represented by

$$(74) \quad B = [b_{kl}] = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

and so on and so forth. It is increasingly common to see in Differential Geometry simply  $b_{kl}$  to be used instead of  $B$ , so we can explicitly calculate things out.

Matrix addition is merely componentwise addition. So if  $A = [a_{ij}]$  and  $B = [b_{kl}]$  then

$$(75) \quad A + B = [a_{ij} + b_{ij}] = \begin{bmatrix} a_{11} + b_{11} & \dots \\ \vdots & \ddots \end{bmatrix}.$$

Matrix addition is straightforward, it's done component-wise.

The other operation, matrix multiplication, is a bit trickier! If  $A$  is an  $m$ -by- $n$  matrix and  $B$  is an  $n$ -by- $p$  matrix, then their product  $C = AB$  is an  $m$ -by- $p$  matrix defined component wise by

$$(76) \quad C = [c_{ij}] = \left[ \sum_k^n a_{ik} b_{kj} \right]$$

where  $A = [a_{ij}]$  and  $B = [b_{kj}]$ .

Also note that we may think of a matrix as being composed of column vectors or row vectors. Why should we think of it like this? Well, matrix multiplication drastically simplifies 100 fold. Observe that it suddenly becomes little more than finding the dot product of the  $i^{\text{th}}$  row of  $A$  with the  $j^{\text{th}}$  column of  $B$  to find the component  $c_{ij}$ .<sup>1</sup>

For a generalized example, consider the following illustration:

$$(77) \quad \mathbf{AB} = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \end{bmatrix} * \begin{bmatrix} B_1 & B_2 & B_3 & \dots \end{bmatrix} = \begin{bmatrix} (A_1 \cdot B_1) & (A_1 \cdot B_2) & (A_1 \cdot B_3) & \dots \\ (A_2 \cdot B_1) & (A_2 \cdot B_2) & (A_2 \cdot B_3) & \dots \\ (A_3 \cdot B_1) & (A_3 \cdot B_2) & (A_3 \cdot B_3) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

where  $A_i$  are row vectors and  $B_k$  are column vectors.

<sup>1</sup>The rule I always remember is "row-dot-column".

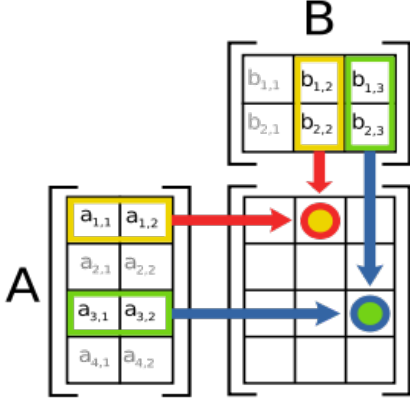


FIGURE 1. The “row dot column” technique of matrix multiplication illustrated.

We will explicitly compute an example of matrix multiplication

$$(78) \quad \begin{bmatrix} 1 & 0 & 2 \\ -1 & 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 \times 3 + 0 \times 2 + 2 \times 1 & 1 \times 1 + 0 \times 1 + 2 \times 0 \\ -1 \times 3 + 3 \times 2 + 1 \times 1 & -1 \times 1 + 3 \times 1 + 1 \times 0 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix}$$

**Example 37.** Given two vectors  $\vec{a} = a_i, \vec{b} = b_j$  their dot product is explained through the use of matrix multiplication

$$(79) \quad \vec{a} \cdot \vec{b} = \sum_{i=1}^n a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

or more generally, if we are using complex-valued vectors

$$(80) \quad \vec{a} \cdot \vec{b} = \sum_{i=1}^n a_i \bar{b}_i$$

where  $\bar{b}_i$  are the complex conjugate components of  $\vec{b}$ . QEF.

Let us note some properties of Matrix multiplication: it is associative

$$(81) \quad A(BC) = (AB)C$$

it is distributive

$$(82) \quad A(B + C) = AB + AC$$

and

$$(83) \quad (A + B)C = AC + BC.$$

It is compatible with scalar multiplication too (let  $c$  be a scalar)

$$(84) \quad \begin{aligned} c(AB) &= (cA)B \\ (Ac)B &= A(cB) \\ (AB)c &= A(Bc). \end{aligned}$$

**B.2. Euclidean and Einstein Summation Conventions.** There is a convention which Einstein invented (or so I am told, I may be completely wrong!) where repeated indices are summed over. This occurs if and only if at least one of the repeating indices is a subscript and at least one is a superscript. For example, the dot product is represented as

$$(85) \quad \vec{a} \cdot \vec{b} = a_i b^i = \sum_i a_i b^i.$$

Do not confuse the superscripted  $b$  vector for “ $b$  to the  $i^{\text{th}}$  power”! It is used to indicate which vectors are **covariant** and which vectors are **contravariant**.<sup>2</sup> Here covariant vectors are linear functionals (see §4). Covariant vectors are denoted by the subscript indices, and contravariant vectors are denoted by the superscript indices.

We will resist the urge to go into details about the notion of covariance and contravariance until §4, feel free to skip ahead to find out more about it.

On the other hand, if we just sum whenever we see two indices of any kind with the same variable, this is called **Euclidean Summation Convention**<sup>3</sup>. For instance

$$(86) \quad a_{ijk}b_{klm} = \sum_k a_{ijk}b_{klm}$$

is done by Euclidean convention, but this would never happen in Einstein convention!

**Remark 38.** Most of the time, Euclidean summation convention and Einstein summation convention are both used in physics and math texts. Care should be used in the future!

**B.3. Transposes of Matrices.** When we have a matrix  $A$ , we can take its transpose, which is (for a 2 by 2 matrix)

$$(87) \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix}^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

Note how the diagonal entries stay the same as we swap the off diagonal components.

We can take the transpose of rectangular matrices too. Observe

$$(88) \quad \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.$$

Recall that we represent vectors as a column vector, which is a column matrix with 1 entry per row. That is, for a 2 dimensional vector we have

$$(89) \quad \vec{v} = \begin{bmatrix} a \\ b \end{bmatrix}$$

and a 3 dimensional vector:

$$(90) \quad \vec{w} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

We too can take their transposes, and we end up with a *row vector*!

There are some important things to remember about transposes. First of all, it's an involution. What does this mean? Well, if you do it twice, you get back the original matrix. That is

$$(91) \quad (A^T)^T = A.$$

Go ahead and try to prove this one for yourself.

Secondly, it's **linear**! That is, given two matrices  $A$  and  $B$ , we have

$$(92) \quad (A + B)^T = A^T + B^T$$

<sup>2</sup>There is some confusion and debate about the use of these terms in higher mathematics, since in category theory they mean the exact opposite of what they mean in the physics application of differential geometry. We will use the physics conventions here.

<sup>3</sup>The origin of this phrase is, as far as the author is aware, from Misner, Thorne, and Wheeler's *Gravitation*, Chapter 12.3, page 294: “(‘Euclidean’ index conventions; repeated space indices to be summed even if both are down; dot denotes time derivative)”.

which is trivial but an example will be given:

$$(93) \quad \left( \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} w & x \\ y & z \end{bmatrix} \right)^T = \left( \begin{bmatrix} a+w & b+x \\ c+y & d+z \end{bmatrix} \right)^T = \begin{bmatrix} a+w & c+y \\ b+x & d+z \end{bmatrix}$$

and

$$(94) \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix}^T + \begin{bmatrix} w & x \\ y & z \end{bmatrix}^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} + \begin{bmatrix} w & y \\ x & z \end{bmatrix} = \begin{bmatrix} a+w & c+y \\ x+b & d+z \end{bmatrix}$$

which is precisely what we just got for  $(A+B)^T$ ! So we find it holds true for  $2 \times 2$  matrices.

Now, there is a rather counter-intuitive property that

$$(95) \quad (AB)^T = B^T A^T.$$

This is precisely the property that we want to see!

#### REFERENCES

- [Art91] M. Artin. *Algebra*. Prentice-Hall (1991).  
 [DF04] D. Dummit, R. M. Foote. *Abstract Algebra*. John Wiley and Sons, third edition (2004).  
*E-mail address:* `pqnelson@gmail.com`  
*URL:* `https://pqnelson.github.io/tebkb/`