Brief Notes concerning Lie Groups and Algebras:

for Physics 495, *An Introduction to Special Relativity*

Daniel Finley, Fall, 2001

25 September, 2001

1. A *group* G is a set together with a (binary) mapping $\varphi:\ G\times G \to G$ such that for all $a, b, c \in G$

a.) $\varphi(a, b) \in G\,,\qquad \{\text{closure}\},$

b.) $\varphi[\varphi(a, b), c] = \varphi[a, \varphi(b, c)]\,,\qquad \{\text{associativity}\},$

c.) $\exists\ e \in G$ such that $\varphi(e, a) = a = \varphi(a, e)\,,\qquad \{\text{identity}\},$

d.) $\forall a \in G, \exists!\ a^{-1} \in G$ such that $\varphi(a, a^{-1}) = e = \varphi(a^{-1}, a)\,,\qquad \{\text{unique inverse}\}.$

Notes:

i.) The mapping is often referred to as a "product."

ii.) In practice, the group mapping is usually indicated simply by the use of juxtaposition of the two elements in question, occasionally by the use of $\circ$, or with $+$ when the mapping is commutative, i.e., when $\varphi(a, b) = \varphi(b, a)$.

iii.) The Cartesian product of two sets, A and B, is denoted by $A \times B$; the Cartesian product of two or more sets, say p of them, is defined as simply the set of all p-tuples of objects chosen from the sets in question, in the order given:

$$A \times B \times C \text{ is defined as } \{(a, b, c) \mid a \in A, b \in B, c \in C\}\,,$$

where the right-hand side, above, is read as "the set of all possible triples of quantities, where we choose an element from A, then one from B, and then one from C." Therefore a statement such as $f: G\times H \to J$ is to be read as f is a mapping that takes one object from the set G, and a second object from the set H and then gives one an object from the set J. It is usual to require some sort of requirements about how the mapping does this, and then to specify that by saying, for instance, that the mapping f is continuous, or differentiable, or smooth, i.e., arbitrarily-often differentiable, or analytic, etc. I take all our mappings to be smooth, at least, unless other specified.

iv.) Some "almost-standard" mathematical shorthand notation is used in the definition above:

$$\begin{cases} a \in A & \text{means that "a is an element of the set A";} \\ \exists\ e & \text{means that "there exists a quantity called } e\text{";} \\ \exists! & \text{means that "there exists } \textit{uniquely} \ \ldots \text{ ";} \\ \forall a & \text{means "for all possible choices of a quantity } a\text{";} \\ \ni & \text{is sometimes used as an abbreviation for "such that".} \end{cases}$$

2. A *homomorphism* of two groups is a mapping $f : G_1 \to G_2$, which preserves the product, i.e., $f(g) \diamondsuit f(h) = f(g \circ h)$, where we have indicated the product in G_1 with $\circ$ and the product in G_2 with $\diamondsuit$.

2a.) An *isomorphism* of two groups is a homomorphism that is one-to-one.

2b.) A *representation* of a group, G is a homomorphism of G into another group, with that second group being a group of *linear operators* on some space. [Obviously that space needs to be one that allows linear operators.] Typical examples of linear operators of this type are (i) matrices over some vector space, (ii) differential operators over some function space.

3. A Lie group is a group containing continuously many elements for which the product mapping is separately and jointly continuous.

4. A *field* is a set, K, with two binary mappings, $\mu : K \times K \to K$ and $\alpha : K \times K \to K$, with the following requirements on them:

a.) K is a group with respect to α, and is commutative; we refer to the identity for this group structure as f_0;

b.) the set K is also a group with respect to μ, **except that the element f_0 does not have an inverse**; this structure may or may not be commutative; we denote its identity by f_1;

c.) the two mappings satisfy a *distributive law* for all $a, b, c \in K$:

$$\mu[a, \alpha(b, c)] = \alpha[\mu(a, b), \mu(a, c)] \quad \text{and} \quad \mu[\alpha(a, b), c] = \alpha[\mu(a, c), \mu(b, c)] \ .$$

Notes:

i.) Usually μ is represented by $\circ$, or perhaps simply juxtaposition, while α, being necessarily commutative, is usually represented by $+$. In such a case the distributive requirement looks rather simpler:

$a \circ (b + c) = ab + ac$ and $(a + b)c = ac + bc$.

ii.) The most commonly used fields are the real numbers, $\mathbb{R}$, the complex numbers, $\mathbb{C}$, and the quaternions, $\mathcal{Q}$, with the first two, of course, having their multiplication as commutative.

5. A *vector space* is a set of objects, V, along with an associated field, K, together with two mappings, $A : V \times V \longrightarrow V$, called "addition of vectors," and $M : K \times V \longrightarrow V$, called "scalar multiplication," which have the following constraints, where we use $+$ and $\circ$ for the operations within the field K, and take arbitrary elements $\alpha, \beta \in K$ and $v, w \in V$:

a.) V with A is a commutative group; we refer to its identity as $\vec{0}$;

b.) the following "distributivity laws" hold with respect to the interaction of the two mappings:

b1. $M[\alpha, M(\beta, v)] = M(\alpha \circ \beta, v)$;

b2. $M(f_1, v) = v$, and $M(f_0, v) = \vec{0}$;

b3. $M[\alpha, A(v, w)] = A[M(\alpha, v), M(\alpha, w)]$;

b4. $M(\alpha + \beta, v) = A[M(\alpha, v), M(\beta, v)]$.

Notes:

i.) It is common to denote the "addition" map in a vector space by the same symbol as the addition map in its associated field, i.e., by the symbol "+". If we do that, and also denote multiplication in the field and by scalars by simple juxtaposition, then the four distributivity requirements above take on the form

$$(b1)\ \alpha(\beta v) = \alpha\beta v\ , \qquad (b2)\ 1v = v\ , \quad 0v = \vec{0}\ ,$$

$$(b3)\ \alpha(v + w) = \alpha v + \alpha w\ , \quad (b4)\ (\alpha + \beta)v = \alpha v + \beta v\ .$$

ii.) The space of tangents to all curves passing through any particular point in spacetime forms a vector space. The set of all such tangent vectors at all points in (at least) some neighborhood of points is called a tangent bundle, and I refer to it as $\mathfrak{T}$. To refer to just the single vector space at a point P on our spacetime, we append a subscript, namely $\mathfrak{T}_P$.

6. The *dual vector space* for a given vector space, V, is often denoted by V^*, and is defined as the set of all continuous, linear mappings from the vector space into the real numbers, i.e.,

$$\phi \in V^* \implies \quad \phi : V \to \mathbb{R} \, , \quad \text{such that}$$

$$\phi(\alpha a + b) = \alpha\phi(a) + \phi(b) \, , \quad \forall \alpha \in \mathbb{R} \;\text{ and }\; a, b \in V \, .$$

Notes:

i.) The dual space V^* is a vector space itself; one can check this by, for instance, assigning $(\phi + \psi)(a) \equiv \phi(a) + \psi(a)$. Here, of course the $+$ on the left-hand side is the newly-defined addition in V^*, while that on the right-hand side is addition in $\mathbb{R}$.

ii.) The dual space to the dual space, i.e., $(V^*)^*$, is the same as V in all finite-dimensional cases, and some infinite-dimensional ones.

iii.) Elements of the dual space are often called *co-vectors*.

iv.) When the vector space V is the space of tangent vectors to curves over spacetime, we refer to its elements as *tangent vectors* and the bundle of such vector spaces by the symbol $\mathfrak{T}$, or also by $\mathfrak{T}^1$; in that case it is most usual to refer to the elements of the dual space as *1-forms,* and the bundle of all those dual spaces over the same neighborhood by the symbol Λ or Λ^1.

7. The *tensor product* of two vector spaces is a coupling of elements from the two vector spaces, with requirements on equivalency on the couplings that preserve the notions of linearity so that the tensor product is itself a vector space. To be more precise, let A and B be the two vector spaces, and define their tensor product, $A \otimes B$, by the existence of a map, ϕ, as follows, where $a, b \in A$ and $v, w \in B$, while $\alpha, \beta \in K$:

$$\phi : A \times B \to A \otimes B \ , \ \text{ such that}$$

$$\phi(a+b, v) = \phi(a, v) + \phi(b, v) \ \text{ and } \ \phi(a, v+w) = \phi(a, v) + \phi(a, w) \ ,$$

$$\phi(\alpha a, \beta v) = \phi(\alpha\beta a, v) \ = \phi(a, \alpha\beta v) = \alpha\beta\phi(a, v) \ ,$$

Notes:

i.) I have only defined the tensor product, above, for two vector spaces that share a common field, K, of scalars; obviously more generality could have been maintained, but it is totally unnecessary for our needs.

ii.) A somewhat different approach to defining the tensor product of A and B is first to suppose that

$$\{e_i\}_1^a \text{ is a basis for } A, \text{ and } \ \{v_r\}_1^b \text{ is a basis for } B;$$

$$A \otimes B \text{ is a vector space with basis } \{e_i v_r \mid i = 1, \ldots, a \ , r = 1, \ldots, b\} \ ,$$

where no definition at all is given for the juxtaposition of any of the two pairs of basis vectors in the definition of the basis. [Such a product is then given some name such as "tensor product" or "dyad product," etc.] It is easy to check that such a definition, where the elements of $A \otimes B$ are simply all possible linear combinations of the new basis elements—there being ab of them—does satisfy the definition above, and is, perhaps (?), more intuitive, and certainly more heuristic.

iii.) Once we can create the tensor product of two vector spaces, it is straightforward to take that product and take a tensor product with a third, etc., etc., so that one may take the tensor product of an arbitrary, finite, number of vector spaces with no problems. It can be shown that the requirements of linearity insure that the resulting multiple tensor products are associative!

iv.) A tensor over spacetime, then, may be thought of as an element of the tensor product of some number of copies of the tangent vector space, $\boldsymbol{\mathfrak{T}}^1$, and its dual, Λ^1, the space of 1-forms. If a given tensor $T \in \boldsymbol{\mathfrak{T}}^1 \otimes \ldots \boldsymbol{\mathfrak{T}}^1 \otimes \Lambda^1 \otimes \ldots \Lambda^1$, where there are r copies of $\mathfrak{T}^1$ in the product and s copies of Λ^1 in the product, then we say that the tensor is *"of type* $[r, s]$, or that it has r contravariant indices and s covariant indices.

8. An *algebra* is a vector space, A, over its associated field, K, together with an additional product mapping into itself, $C : A \times A \to A$, which "respects" the underlying vector space structures, i.e., the addition and scalar multiplication there. We formalize this restriction with the following *bilinearity* requirements. Here I use "+" for the addition map in the vector space and the symbol $\bullet$ to denote this additional product, and let a, b, c be elements of the algebra, A, and $\alpha, \beta \in K$:

$$a \bullet b \in A\,, \qquad \text{"closure"};$$

$$(a + b) \bullet c = \; a \bullet c + b \bullet c \;\text{ and }\; a \bullet (b + c) = a \bullet b + a \bullet c\,,$$

"bilinearity with respect to addition";

$$(\alpha a) \bullet b = \alpha(a \bullet b) = \; a \bullet (\alpha b)\,, \text{ "bilinearity w.r.t. scalar multiplication"}.$$

Notes:

i.) There are many different types of (useful) algebras.

ii.) For instance, the set of ordinary square matrices form an *associative algebra* under the usual matrix multiplication.

iii.) A newer one for us is the Lie algebra, defined next.

7. A *Lie algebra*, $\mathcal{G}$, is an algebra where the product operation is skew-symmetric, and therefore not associative but rather is required to satisfy the Jacobi identity. It is usual to refer to this product as the *Lie product*, but also, often, just the commutator product. As usual we use $\bullet$ to denote the algebraic product operation—remembering that it is now **not associative**—and let a, b, and c, be elements of the algebra $\mathcal{G}$:

a.) $a \bullet b + b \bullet a = 0 \; \in \mathcal{G}\,,$ "skew symmetry";

b.) $a \bullet (b \bullet c) + b \bullet (c \bullet a) + c \bullet (a \bullet b) = 0\,,$ "Jacobi identity."

Notes:

i.) skew-symmetry implies non-associativity, i.e., $a \bullet (b \bullet c) \neq (a \bullet b) \bullet c$.

ii.) The usual commutator of 2 matrices, i.e., $[A, B] \equiv AB - BA$, where A and B are two square matrices, satisfies the requirements above for a Lie product.

Notice that this means that $a \bullet (a \bullet b)$ is the same as $[A, [A, B]]$, while of course $(a \bullet a) \bullet b$ is written as $[[A, A], B] = 0$.

This allows one to write expressions such as $A^3 \bullet B \equiv [A, [A, [A, B]]]$; however, a much more common notation is to use the *adjoint operator* over a Lie algebra:

$$(\mathrm{ad\,A})B \equiv [A, B] \equiv A \bullet B \ ,$$

$$(\mathrm{ad\,A})^n B = [A, [A, [A, \ldots, [A, B] \ldots]]] \quad \text{with } n \ A\text{'s in the sequence.}$$

iii.) Since the Lie algebra is a vector space, we may choose a basis for it and then write any other element as a linear combination of those basis elements. Let $\{\mathbf{X}_i\}_1^n$ be a basis set for an n-dimensional Lie algebra. Then any two arbitrary elements of the Lie algebra may be written in the form

$$A = A^i \mathbf{X}_i \ , B = B^j \mathbf{X}_j \implies \quad A \bullet B = A^i B^j \{X_i \bullet X_j\} \ .$$

Therefore we see that it is sufficient to know the commutators of the basis elements. It is usual to define the quantities ${C_{ij}}^k$, which are called the commutation coefficients of the algebra, and, to within choice of basis, characterize that Lie algebra:

$$X_i \bullet X_j \equiv {C_{ij}}^k X_k \ \ , \text{or} \ \ [X_i \, , \, X_j] \equiv {C_{ij}}^k X_k \ .$$

iv.) The Baker-Campbell-Hausdorff theorem describes the behavior of the product of the exponentials of any two elements of an algebra, expressing that result in terms of a single exponential of an infinite sequence of multiple commutators of those elements:

$$e^{sA}\, e^{tB} \ = e^{H(sA,tB)} \ ,$$

$$\begin{aligned} H(sA, tB) = sA + tB + \tfrac{1}{2} st[A, B] &+ \tfrac{1}{12} st(s[A, [A, B]] + t[B, [B, A]]) \\ &+ \tfrac{1}{48} s^2 t^2 ([A, [B, [B, A]]] - [B, [A, [A, B]]]) + \ \ldots \quad , \end{aligned}$$

where all the other terms involve **only** various different repeated commutators. Because of this theorem, the exponentials of elements of a Lie algebra form a group, defining that product (of the exponentials) by the addition in the algebra. The proofs as to exactly how this "heuristic" statement is implemented are Lie's famous Three Theorems.

v.) A very simple, but not completely trivial, example of a Lie group and its Lie algebra is the (1-dimensional) group of translations of functions of one real variable. The group of translations of that one variable, say x, may be described as the set

$$T \equiv \{T_a \mid a \in \mathbb{R}\} \quad \text{such that} \quad (T_a[f])(x) = f(x+a) \ .$$

Using Taylor's theorem and the standard expansion of the exponential function, we may rewrite this as

$$f_a(x) \equiv (T_a[f])(x) \equiv f(x+a) = e^{a\frac{d}{dx}} f(x) = e^{a\check{t}} f(x) \ ,$$

so that a basis of this (1-dimensional) Lie algebra is just the differential operator $t \equiv d/dx$.