APPENDIX A

TENSORS AND DIFFERENTIAL FORMS ON VECTOR SPACES

Since only so much of the vast and growing field of differential forms and differentiable manifolds will be actually used in this survey, we shall attempt to briefly review how the calculus of exterior differential forms on vector spaces can serve as a replacement for the more conventional vector calculus and then introduce only the most elementary notions regarding more topologically general differentiable manifolds, which will mostly be used as the basis for the discussion of Lie groups, in the following appendix.

Since exterior differential forms are special kinds of tensor fields – namely, completely-antisymmetric covariant ones – and tensors are important to physics, in their own right, we shall first review the basic notions concerning tensors and multilinear algebra. Presumably, the reader is familiar with linear algebra as it is usually taught to physicists, but for the "basis-free" approach to linear and multilinear algebra (which we shall not always adhere to fanatically), it would also help to have some familiarity with the more "abstract-algebraic" approach to linear algebra, such as one might learn from Hoffman and Kunze [**1**], for instance.

1. Tensor algebra. – A tensor algebra is a type of algebra in which multiplication takes the form of the tensor product.

a. Tensor product. – Although the tensor product of vector spaces can be given a rigorous definition in a more abstract-algebraic context (See Greub [**2**], for instance), for the purposes of actual calculations with tensors and tensor fields, it is usually sufficient to say that if V and W are vector spaces of dimensions n and m, respectively, then the *tensor product* $V \otimes W$ will be a vector space of dimension nm whose elements are finite linear combinations of elements of the form $\mathbf{v} \otimes \mathbf{w}$, where $\mathbf{v}$ is a vector in V and $\mathbf{w}$ is a vector in W. The tensor product $\otimes$ then takes the form of a bilinear map $V \times W \to V \otimes W$, $(\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} \otimes \mathbf{w}$. In particular, that kind of product is not closed, since the tensor product of two vectors will belong to a different vector space.

Bilinearity means that the map is linear in each factor individually, but not collectively linear. Hence:

$$(\alpha \mathbf{v} + \beta \mathbf{v}') \otimes \mathbf{w} = \alpha \mathbf{v} \otimes \mathbf{w} + \beta \mathbf{v}' \otimes \mathbf{w}, \tag{1.1}$$

$$\mathbf{v} \otimes (\alpha \mathbf{w} + \beta \mathbf{w}') = \alpha \mathbf{v} \otimes \mathbf{w} + \beta \mathbf{v} \otimes \mathbf{w}'. \tag{1.2}$$

One can also see that this means that the tensor product is right and left distributive over vector addition.

Because of this bilinearity, if $\{\mathbf{e}_i , i = 1, \ldots, n\}$ is a basis for V and $\{\mathbf{f}_a , a = 1, \ldots, m\}$ is a basis for W then $\{\mathbf{e}_i \otimes \mathbf{f}_a , i = 1, \ldots, n, a = 1, \ldots, m\}$ will constitute a basis for $V \otimes W$.

Hence, if $\mathbf{t}$ is an element of $V \otimes W$ then it can be expressed as a linear combination of the basis elements in the form:

$$\mathbf{t} = t^{ia}\,\mathbf{e}_i \otimes \mathbf{f}_a \equiv \sum_{i=1}^{n}\sum_{a=1}^{m} t^{ia}\,\mathbf{e}_i \otimes \mathbf{f}_a\,. \tag{1.3}$$

The numbers t^{ia} are referred to as the *components* of $\mathbf{t}$ with respect to the chosen basis on $V \otimes W$. Most of the literature of theoretical physics, even up to the present era, deals exclusively with the components of tensors, although if brevity be the soul of wit then one can easily see that dealing with the intrinsic objects, such as $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{t}$, can add a certain clarity and conciseness to one's mathematical expressions, even if it does involve a small investment of abstraction in the process.

One can also use the bilinearity of the tensor product to express the tensor product $\mathbf{v} \otimes \mathbf{w}$ in terms of components. Suppose that $\mathbf{v} = v^i\,\mathbf{e}_i$ and $\mathbf{w} = w^a\,\mathbf{f}_a$ ([1]). One will then have:

$$\mathbf{v} \otimes \mathbf{w} = v^i\,w^a\,\mathbf{e}_i \otimes \mathbf{f}_a\,; \tag{1.4}$$

i.e., the components of $\mathbf{v} \otimes \mathbf{w}$ with respect to the chosen basis will be $v^i\,w^a$.

We now see that there are two distinct types of elements in $V \otimes W$, namely, *decomposable* elements which have the form $\mathbf{v} \otimes \mathbf{w}$, and *indecomposable* elements, which have the more general form of finite linear combinations of decomposable elements, such as in (1.3). The fact that not all elements are decomposable is due to the fact that linear combinations of decomposable elements do not have to be decomposable. In the case of exterior algebra, which we shall discuss shortly, the decomposable elements will sometimes define quadric hypersurfaces in the tensor product space, rather than vector subspaces.

b. Contravariant tensors. – A common situation in tensor algebra (as well as in physics) is when the vector space W is the vector space V. One can then refer to the elements of $V \otimes V$ as *second-rank contravariant* tensors (over V). The term "second-rank" refers to the fact that there are two copies of V in the tensor product. Hence, a basis can be defined by $\{\mathbf{e}_i \otimes \mathbf{e}_j,\ i, j = 1, \ldots, n\}$ and components will look like t^{ij}:

$$\mathbf{t} = t^{ij}\,\mathbf{e}_i \otimes \mathbf{e}_j\,. \tag{1.5}$$

The term "contravariant" refers to the way that the components transform under a change of basis. In particular, if:

$$\overline{\mathbf{e}}_i = \mathbf{e}_j A_i^j \tag{1.6}$$

is a change of linear basis in V (so A_i^j is an invertible matrix) then the components of $\mathbf{v}$ and $\mathbf{w}$ (which were v^i and w^i with respect to $\mathbf{e}_i$) will now be:

([1]) From now on, we shall invoke the "summation convention," which is often attributed to Einstein, namely, whenever a superscript agrees with a subscript, one sums over all defined values of the index in question. In the occasional situations where it is necessary to refer to components with doubled indices, such as the diagonal elements of matrices, the convention will usually be rescinded explicitly.

$$\bar{v}^i = \tilde{A}^i_j v^j , \qquad \bar{w}^i = \tilde{A}^i_j w^j \tag{1.7}$$

with respect to $\bar{\mathbf{e}}_i$. The notation $\tilde{A}^i_j$ refers to the inverse of the matrix A^i_j, so this type of transformation is referred to as *contravariant.*

From the bilinearity of the tensor product, the resulting transformation of the components of $\mathbf{v} \otimes \mathbf{w}$ will be:

$$\bar{v}^i \, \bar{w}^j = \tilde{A}^i_k \, \tilde{A}^j_l v^k w^l , \tag{1.8}$$

and more generally (also due to the bilinearity of the tensor product), the components t^{ij} of $\mathbf{t}$, as in (1.5), will transform to:

$$\bar{t}^{ij} = \tilde{A}^i_k \, \tilde{A}^j_l t^{kl} . \tag{1.9}$$

Hence, one can say that they are *doubly-contravariant.*

Since the tensor product is also associative:

$$(\mathbf{a} \otimes \mathbf{b}) \otimes \mathbf{c} = \mathbf{a} \otimes (\mathbf{b} \otimes \mathbf{c}), \tag{1.10}$$

one can define higher tensor products $V \otimes \ldots \otimes V$ of a finite number of copies of V, and the elements of $\otimes_k V = V \otimes \ldots \otimes V$ when there are – say – k copies of V are then referred to as *rank-k contravariant tensors* over V. Hence, a basis for $\otimes_k V$ can be given by all tensor products of the form $\mathbf{e}_{i_1} \otimes \ldots \otimes \mathbf{e}_{i_k}$, the components of a general element $\mathbf{t} \in \otimes_k V$ will take the form $t^{i_1 \cdots i_k}$, and they will transform to:

$$\bar{t}^{\,i_1 \cdots i_k} = \tilde{A}^{i_1}_{j_1} \cdots \tilde{A}^{i_k}_{j_k} \, t^{\,j_1 \cdots j_k} \tag{1.11}$$

under a change of basis on V.

Clearly, the dimension of $\otimes_k V$ will be n^k.

One can also form tensors of mixed rank over V by forming finite linear combinations of elements in various $\otimes_k V$'s for different values of k. For instance, one might form the linear combination $\mathbf{a} + \mathbf{b} \otimes \mathbf{c}$. Such expressions cannot generally be simplified further, unless some further structure is imposed upon $\otimes_k V$, which will usually be another algebra, for us. A tensor that does not have mixed type is referred to as *homogeneous*, although in most cases, that will be tacitly implied.

The direct sum $\otimes_0 V \oplus \otimes_1 V \oplus \ldots \oplus \otimes_k V \oplus \ldots$ of all the vector spaces $\otimes_k V$ as k varies from 0 ($\otimes_0 V \equiv \mathbb{R}$) to infinity will denoted by simply $\otimes_* V$. It will be referred to as the *contravariant tensor algebra* over V, and will clearly be infinite-dimensional. One sees that the tensor product then makes the vector space $\otimes_* V$ into an *algebra*, since it defines a bilinear binary product on the vector space. It will also be associative and possess a unity element (namely, 1) as an algebra. In addition, since the tensor product of a homogeneous tensor of rank k with one of rank l will be a homogeneous tensor of rank $k + l$, one refers to the algebra $\otimes_* V$ as a *graded* algebra.

c. Covariant tensors. – The dual space V^* to the vector space V (viz., the vector space of all linear functionals on V) is itself a vector space, so one can still define tensors of any rank over it. Hence, $\otimes^k V \equiv \otimes_k V^* = V^* \otimes \ldots \otimes V^*$ (k copies) is a vector space of dimension n^k, and if $\{\theta^i, i = 1, \ldots, n\}$ is a basis for V^* then a basis for $\otimes^k V$ can be defined by $\theta^{i_1} \otimes \ldots \otimes \theta^{i_k}$, and a general element $t \in \otimes^k V$ will then take the form:

$$t = t_{i_1 \cdots i_k}\, \theta^{i_1} \otimes \ldots \otimes \theta^{i_k} . \tag{1.12}$$

The scalar components $t_{i_1 \cdots i_k}$ are then the components of a rank-k covariant tensor over V. Under a change of basis on V^*:

$$\bar{\theta}^i = \tilde{A}^i_j\, \theta^j , \tag{1.13}$$

the components $t_{i_1 \cdots i_k}$ will transform to:

$$\bar{t}_{i_1 \cdots i_k} = A^{j_1}_{i_1} \cdots A^{j_k}_{i_k} t_{j_1 \cdots j_k} . \tag{1.14}$$

Of particular interest is the case in which the basis θ^i for V^* is *reciprocal* to the basis $\mathbf{e}_i$ for V. In that case, one will have, by definition:

$$\theta^i (\mathbf{e}_j) = \delta^i_j = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases} \tag{1.15}$$

δ^i_j is then the *Kronecker delta symbol.*

In order for a reciprocal basis to go to a reciprocal basis under a change of basis $\mathbf{e}_i$, such as in (1.6), one must have:

$$\delta^i_j = \bar{\theta}^i (\bar{\mathbf{e}}_j) = B^i_l\, \theta^l (\mathbf{e}_k A^k_j) = B^i_l\, A^k_j\, \theta^l (\mathbf{e}_k) = B^i_l\, A^k_j\, \delta^l_k = B^i_k\, A^k_j .$$

Hence, B^i_j can only be $\tilde{A}^i_j$. One then says that the basis θ^i transforms *contragrediently* to the basis $\mathbf{e}_i$, and in fact, so do the components of tensors over V^*.

One can also form the direct sum $\otimes^* V \equiv \otimes_* V^* = \otimes^0 V \oplus \otimes^1 V \oplus \ldots \oplus \otimes^k V \oplus \ldots$ over all k and call it the *algebra of covariant tensors over V.* It will once more be an infinite-dimensional graded associative algebra with unity.

d. Tensors mixed variance. – One can take the tensor product $\otimes^k_l V \equiv (\otimes^k V) \otimes (\otimes_l V)$ and obtain a vector space of dimension n^{k+l} whose elements are finite linear combinations of (homogeneous) elements of the form $\alpha^{i_1} \otimes \ldots \otimes \alpha^{i_k} \otimes \mathbf{v}_{j_1} \otimes \ldots \otimes \mathbf{v}_{j_l}$, in which the α's are linear functionals on V, and the $\mathbf{v}$'s are vectors in V. Such an element will then be a tensor that is k-times covariant and l-times contravariant. Hence, a basis for $\otimes^k_l V$ can be defined by the tensor products $\theta^{i_1} \otimes \ldots \otimes \theta^{i_k} \otimes \mathbf{e}_{i_1} \otimes \ldots \otimes \mathbf{e}_{i_l}$, and a general element $t \in \otimes^k_l V$ will take the component form:

$$t = t^{j_1\cdots j_l}_{i_i\cdots i_k}\;\theta^{i_1}\otimes \ldots \otimes \theta^{i_k}\otimes \mathbf{e}_{i_1}\otimes \ldots \otimes \mathbf{e}_{i_l}\,. \tag{1.16}$$

Under a change of basis (1.6) for V and the contragredient change of the reciprocal basis for V^*, the components $t^{j_1\cdots j_l}_{i_i\cdots i_k}$ will become:

$$\overline{t}^{\,j_1\cdots j_l}_{i_i\cdots i_k} = A^{r_1}_{i_1}\cdots A^{r_k}_{i_k}\tilde{A}^{j_1}_{s_1}\cdots \tilde{A}^{j_l}_{s_l}\,t^{s_1\cdots s_l}_{r_i\cdots r_k}\,. \tag{1.17}$$

The direct sum of all $\otimes^k_l V$ for all k and l then becomes an associative algebra with unity that is doubly-graded and is referred to as the *full tensor algebra* over V.

Of particular interest in the mixed tensor algebra is the vector space $\otimes^1_1 V = V^* \otimes V$. That is because if $\mathbf{e}_i$ is a basis for V and θ^j is a basis for V^* (such as the reciprocal basis) then the components of a tensor $t \in \otimes^1_1 V$ will be $n \times n$ matrices t^i_j. Indeed, one can define an isomorphism of $V^* \otimes V$ with the vector space End V of all linear maps from V to itself (i.e., endomorphisms) by taking each decomposable element $\alpha \otimes \mathbf{v}$ to the linear map that takes any vector $\mathbf{w} \in V$ to the vector $\alpha(\mathbf{w})\,\mathbf{v}$. The general indecomposable element in $V^* \otimes V$ is a finite linear combination of elements of the form $\alpha \otimes \mathbf{v}$, so one extends the basic association to a complete map $V^* \otimes V \to \text{End } V$ "by linearity," as they say; viz., one takes finite linear combinations of decomposable elements in $V^* \otimes V$ to the corresponding linear combinations of the corresponding decomposable elements in End V. The fact that the resulting linear map is also an isomorphism is probably easiest to see by choosing a basis for both vector spaces and noting that the map simply associates the *tensor components* t^i_j with the *matrix* t^i_j, which otherwise looks like the identity map.

e. Contraction of covariant tensors with contravariant ones. – There is a basic bilinear pairing of an element $\alpha \in V^*$ with an element $\mathbf{v} \in V$ to form a scalar that amounts to the evaluation of the linear functional α on the vector $\mathbf{v}$ to produce the scalar $\alpha(\mathbf{v})$. When one chooses bases for V and V^*, the components α_i will combine with the components v^i to give:

$$\alpha(\mathbf{v}) = \alpha_i\, v^i. \tag{1.18}$$

One can also regard this as a linear map $\otimes^1_1 V \to \mathbb{R}$ that takes the second-rank mixed tensor $\alpha \otimes \mathbf{v}$ to the scalar $\alpha(\mathbf{v})$, so it will take the components $\alpha_j\, v^i$ to the scalar $\alpha_i\, v^i$. This process can be generalized to a linear map from $\otimes^k_l V$ to $\otimes^{k-1}_{l-1} V$ that takes the decomposable element $\alpha^{i_1}\otimes \ldots \otimes \alpha^{i_k}\otimes \mathbf{v}_{j_1}\otimes \ldots \otimes \mathbf{v}_{j_l}$ to the decomposable element:

$$\alpha^{i_r}(\mathbf{v}_{i_s})\;\alpha^{i_1}\otimes \ldots\otimes \hat{\alpha}^{i_r}\otimes \ldots \otimes\alpha^{i_k}\otimes\mathbf{v}_{j_1}\otimes \ldots \otimes \hat{\mathbf{v}}_{i_s}\otimes\ldots\otimes\mathbf{v}_{j_l}\,,$$

in which the caret denotes the removal of that term from the tensor product. One then extends the map of decomposable elements to indecomposable ones by linearity. In particular, one can perform this process on the basis elements of $\otimes_l^k V$ to produce basis elements for $\otimes_{l-1}^{k-1} V$.

The usual way that one encounters this process in physics is in terms of components. If the components of a tensor $t \in \otimes_l^k V$ are $t^{i_1 \cdots i_r \cdots i_l}_{j_1 \cdots j_s \cdots j_k}$ then the components of the contraction of the covariant index i_r with the covariant index j_s will be:

$$t^{i_1 \cdots \hat{i}_r \cdots i_l}_{j_1 \cdots \hat{j}_s \cdots j_k} = \sum_{i=1}^{n} t^{i_1 \cdots i \cdots i_l}_{j_1 \cdots i \cdots j_k} . \tag{1.19}$$

One refers to the process of setting a superscript and a subscript equal to each other and summing over them as the *contraction* of a mixed tensor of rank $k+l$ to produce a mixed tensor of rank $(k-1)+(l-1)$.

A particularly useful example of the contraction of a mixed-rank tensor is when one applies the process to the components A^i_j of a matrix A, which one can also regard as the components of a tensor in $V^* \otimes V$. The resulting contraction will give the trace of the matrix A;

$$\mathrm{Tr}\, A = A^i_i . \tag{1.20}$$

f. Multilinear functionals. – Since elements of V^* can be evaluated on vectors in V to produce scalars in a linear way, one finds that elements of $V^* \otimes V^*$ can be evaluated on pairs of vectors $(\mathbf{v}, \mathbf{w})$ in $V \times V$ to produce scalars in a *bilinear* way. Specifically, if $\alpha \otimes \beta \in V^* \otimes V^*$ then:

$$(\alpha \otimes \beta)\,(\mathbf{v}, \mathbf{w}) = \alpha(\mathbf{v})\, \beta(\mathbf{w}). \tag{1.21}$$

Hence, the second-rank covariant tensor $\alpha \otimes \beta$ can also be regarded as a bilinear functional on V. If we denote the vector space of all bilinear functionals on V by $L^2\, V$ then the map $V^* \otimes V^* \to L^2\, V$ that one gets by taking each decomposable element $\alpha \otimes \beta$ in $V^* \otimes V^*$ to the bilinear function $\alpha \otimes \beta$ in $L^2\, V$ and extending by linearity is actually a linear isomorphism.

One can easily extend this argument to higher-order tensors by saying that if $\alpha_1 \otimes \ldots \otimes \alpha_k \in \otimes^k V$ then it will also define the k-linear functional on V:

$$(\alpha_1 \otimes \ldots \otimes \alpha_k)\,(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \alpha_1(\mathbf{v}_1) \ldots \alpha_k(\mathbf{v}_k). \tag{1.22}$$

Hence, one gets a linear isomorphism of $\otimes^k V$ with $L^k V$ by extending the previous map in the obvious way.

In order to get a linear isomorphism of $\otimes_k V$ with $L_k V \equiv L^k V^*$, one simply inverts the order of things in the left-hand side of (1.22) to get:

$$(\mathbf{v}_1 \otimes \ldots \otimes \mathbf{v}_k)\,(\alpha_1, \ldots, \alpha_k) = \alpha_1(\mathbf{v}_1) \ldots \alpha_k(\mathbf{v}_k). \tag{1.23}$$

Tensoring $\otimes^k V$ with $\otimes_l V$ will then give a linear isomorphism of the mixed tensor space $\otimes_l^k V$ with the vector space $L_l^k V$ of multilinear functionals on $V \times \ldots \times V \times V^* \times \ldots \times V^*$, in which there are k copies of V and l copies of V^*.

g. Converting contravariant tensors into covariant ones. – Although the vector space V^* is linearly isomorphic to the vector space V, there is not a canonically-defined isomorphism, in general, unless one imposes some other structure on V. In general, one could simply choose an isomorphism $C : V \to V^*$, $\mathbf{v} \mapsto C(\mathbf{v})$, which is referred to as a *correlation* in projective geometry (or really the corresponding map of the projective spaces that are associated with V and V^*). If $\mathbf{e}_i$ is a basis for V and θ^j is the reciprocal basis for V^* then the component matrix C_{ij} will be invertible. One can also regard the linear isomorphism C as an element of $\otimes^2 V$; i.e., a doubly-covariant tensor on V.

However, there is nothing to say that the tensor C is symmetric or antisymmetric. In the event that C is *symmetric*, so:

$$C(\mathbf{v}, \mathbf{w}) = C(\mathbf{w}, \mathbf{v}) \qquad \text{for all } \mathbf{v}, \mathbf{w} \in V, \tag{1.24}$$

the component matrix C_{ij} will also be symmetric in its indices:

$$C_{ij} = C_{ji}\,. \tag{1.25}$$

Since we have also assumed that the matrix is invertible, C can be used as the basis for a scalar product on V (or *metric*, as it is usually called in the theory of relativity).

If C is *antisymmetric*, so:

$$C(\mathbf{v}, \mathbf{w}) = -C(\mathbf{w}, \mathbf{v}) \qquad \text{for all } \mathbf{v}, \mathbf{w} \in V, \tag{1.26}$$

then the component matrix C_{ij} will also be antisymmetric in its indices:

$$C_{ij} = -C_{ji}\,. \tag{1.27}$$

Due to the invertibility of C that will define a *symplectic structure* on V. Such a thing can exist in finite dimensions only when the dimension of V is even.

If the components of a vector $\mathbf{v} \in V$ are v^i with respect to some basis then the components of $C(\mathbf{v})$ with respect to the reciprocal basis will be:

$$v_i = C_{ij}\, v^j\,. \tag{1.28}$$

When C refers to a metric, one calls this process "lowering an index using the metric."

One can also "raise an index" by means of the inverse matrix C^{ij}:

$$C^{ik}\, C_{kj} = C_{jk}\, C^{ki} = \delta^i_j\,. \tag{1.29}$$

One will then get:

$$v^i = C^{ij}\, v_j\,. \tag{1.30}$$

This process can be repeated as many times as one pleases for the components of higher rank tensors. For example:

$$T_{ij} = C_{ik}\, T^{k}{}_{j}\,, \quad T^{ij} = C^{jk}\, T^{i}{}_{k}\,. \tag{1.31}$$

Notice that when the resulting matrices T_{ij} or T^{ij} are not symmetric, it becomes necessary to specify whether the index being raised or lowered will go into the first or second position in the result.

h. Complex tensor algebras. – So far, we have been tacitly assuming that the field of scalars that act upon V by scalar multiplication, and thus the field of scalars that the components of vectors belong to is the field $\mathbb{R}$ of real numbers. In fact, most of what was said can be extended to the field $\mathbb{C}$ of complex numbers without modification. Typically, all that one must do is specify that linearity or multilinearity means $\mathbb{C}$-linearity or $\mathbb{C}$-multilinearity; i.e., with linear (multilinear, resp.) respect to the scalar field $\mathbb{C}$.

Of course, the extension from real to complex vector spaces will introduce some new structures (they are typically related to the existence of the complex conjugation map on complex numbers), and we shall simply discuss those issues as they become relevant in the main body of the text.

2. Exterior algebra. – The full tensor algebra $\otimes V$ over a vector space contains some important subalgebras that are defined by the symmetries of completely-covariant or completely-contravariant tensors. In the simplest case of $\otimes^2 V$, one can decompose that vector space into a direct sum $S^2V \oplus A^2V$, where S^2V is composed of all symmetric, doubly-covariant tensors, and A^2V is composed of all antisymmetric ones. In fact, one can decompose a given element $t \in \otimes^2 V$ into its symmetric and antisymmetric parts by polarization:

$$t = t^+ + t^-, \tag{2.1}$$

in which:

$$t^+(\mathbf{v}, \mathbf{w}) = \tfrac{1}{2}[t(\mathbf{v}, \mathbf{w}) + t(\mathbf{w}, \mathbf{v})], \qquad t^-(\mathbf{v}, \mathbf{w}) = \tfrac{1}{2}[t(\mathbf{v}, \mathbf{w}) - t(\mathbf{w}, \mathbf{v})]. \tag{2.2}$$

This is usually introduced in the context of physics and engineering in its component form:

$$t_{(ij)} = \tfrac{1}{2}(t_{ij} + t_{ji})\,, \qquad t_{[ij]} = \tfrac{1}{2}(t_{ij} - t_{ji})\,. \tag{2.3}$$

Hence, one can think of these definitions as defining linear projections $\otimes^2 V \to S^2 V$ and $\otimes^2 V \to A^2 V$ onto complementary subspaces.

A corresponding discussion will also apply to the contravariant case of $\otimes_2 V$, $S_2 V$, and $A_2 V$.

a. Complete symmetrization and antisymmetrization of tensors. – One defines the complete symmetrization of a rank-k completely-covariant tensor T by way of:

$$T^{+}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \frac{1}{k!}\sum_{\pi} T(\mathbf{v}_{\pi(1)}, \ldots.\mathbf{v}_{\pi(k)}), \tag{2.4}$$

in which $\pi(i)$ is a permutation of the numerals $12\ldots k$, so the sum is over all such permutations. Hence, one is essentially "averaging" the values of $T(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ over all permutations of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

The corresponding expression for the components of T is:

$$T_{(i_1\cdots i_k)} = \frac{1}{k!}\sum_{\pi} T_{\pi(i_1)\cdots\pi(i_k)}. \tag{2.5}$$

If we denote the vector space of completely-symmetric, completely-covariant tensors of rank k on V by S^kV then this process of complete symmetrization will define a linear projection $\otimes^k V \to S^kV$.

The complete antisymmetrization of T then takes the form:

$$T^{-}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \frac{1}{k!}\sum_{\pi} \mathrm{sign}(\pi) T(\mathbf{v}_{\pi(1)}, \ldots.\mathbf{v}_{\pi(k)}), \tag{2.6}$$

in which:

$$\mathrm{sign}(\pi) = \begin{cases} + & \text{when } \pi \text{ is an even permutation} \\ - & \text{when } \pi \text{ is an odd permutation.} \end{cases} \tag{2.7}$$

The corresponding component expression is:

$$T_{[i_1\cdots i_k]} = \frac{1}{k!}\sum_{\pi} \mathrm{sign}(\pi) T_{\pi(i_1)\cdots\pi(i_k)}. \tag{2.8}$$

Hence, if we denote the vector space of all completely-antisymmetric, completely covariant tensors of rank k in V by A^kV then complete antisymmetrization will define a linear projection $\otimes^k V \to A^kV$. The elements of A^kV are referred to as *algebraic k-forms* on V.

Analogous constructions and statements apply to the completely-contravariant tensors on V, and the elements of the vector space A_kV that is obtained by complete antisymmetrization will be referred to as *k-vectors*, such as *bivectors*, *trivectors*, or *multivectors*, in general.

Although one can speak of completely symmetrizing and completely antisymmetrizing completely-covariant tensors of rank k and completely-contravariant tensors of rank k, the decomposition of $\otimes_k V$ and $\otimes^k V$ into direct sums of subspaces will involve more than just the two summands ([1]). Hence, we shall concentrate on only the completely-symmetric and completely-antisymmetric subspaces, for now.

b. Symmetric product. – The complete symmetrization of the tensor product of k vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$, namely:

$$\mathbf{v}_1 \odot \ldots \odot \mathbf{v}_k = \frac{1}{k!}\sum_{\pi} \mathbf{v}_{\pi(1)} \otimes \ldots \otimes \mathbf{v}_{\pi(k)}, \tag{2.9}$$

is called the *symmetric product* of the vectors. For example, when one has two vectors $\mathbf{v}$, $\mathbf{w}$:

$$\mathbf{v} \odot \mathbf{w} = \tfrac{1}{2}(\mathbf{v} \otimes \mathbf{w} + \mathbf{w} \otimes \mathbf{v}). \tag{2.10}$$

The components of any completely-symmetric tensor t will then be completely-symmetric in their indices:

$$t^{i_1 \cdots i_k} = t^{(i_1 \cdots i_k)} \equiv \frac{1}{k!}\sum_{\pi} t^{\pi(i_1)\cdots\pi(i_k)}. \tag{2.11}$$

Although the completely-symmetric tensors on a vector space do not get quite as much attention from differential geometry as the completely-antisymmetric ones, it is important to note two cases in which one does encounter completely-symmetric tensors.

The first one is in the context of mixed partial derivatives. One recalls that as long as one differentiates a k-times continuously-differentiable function $f(x^1, \ldots, x^n)$ of n variables, one will have:

$$f_{,i_1,\ldots,i_k} \equiv \frac{\partial^k f}{\partial x^{i_1} \cdots \partial x^{i_k}} = \frac{\partial^k f}{\partial x^{\pi(i_1)} \cdots \partial x^{\pi(i_k)}}. \tag{2.12}$$

Part of the reason that this fact gets passed over is that modern differential geometry starts with differentiable manifolds that are more topologically general that ones that look like open subsets of $\mathbb{R}^n$, so partial derivatives are no longer regarded as fundamental constructions, as opposed to "covariant derivatives."

Another ways that completely-symmetric tensors can appear in differential geometry is due to the fact that a completely-symmetric k-linear functional on a vector space V of dimension n can be associated with a homogeneous polynomial in the components x^i of any vector $\mathbf{x} \in V$ when one chooses a basis $\mathbf{e}_i$ for V. Specifically, if $T(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ is the completely-symmetric k-linear functional then the corresponding polynomial is defined by:

$$\mathrm{P}[x^1, \ldots, x^n] = T(\mathbf{x}, \ldots, \mathbf{x}) = \sum_{i_1,\ldots,i_k} T_{i_1 \cdots i_k} x^{i_1} \cdots x^{i_k}. \tag{2.13}$$

([1]) In fact, these direct-sum decompositions are closely related to the Clebsch-Gordan decomposition in the theory of representations of Lie groups and Lie algebras and their associated Young tableaux.

An important special case of this is the way that a symmetric, doubly-covariant tensor, such as a metric, can be associated with a quadratic form:

$$Q[\mathbf{x}] = T(\mathbf{x}, \mathbf{x}). \tag{2.14}$$

In order to get inhomogeneous polynomials in **x**, one would have to go to mixed completely-symmetric tensors in $S^*V \equiv \mathbb{R} \oplus S^1 \oplus \ldots$ One calls that algebra, when it is given the completely-symmetrized tensor product $\odot$, the *symmetric algebra* over V^*, and similarly, S_*V will be the *symmetric algebra* over V. Like $\oplus^*$ and $\oplus_*$, S^* and S_* will also be infinite-dimensional.

Since we shall mostly be concerned with completely-antisymmetric tensors in this study, we shall reserve any further comments regarding the symmetric algebras to their specific applications and focus upon the antisymmetric algebras.

c. Exterior product. – The complete-antisymmetrization of the tensor product of k vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$, namely:

$$\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k = \frac{1}{k!}\sum_{\pi} \operatorname{sign}\pi\, \mathbf{v}_{\pi(1)} \otimes \ldots \otimes \mathbf{v}_{\pi(k)}, \tag{2.15}$$

is called the *exterior product* of the vectors; it is also called their *wedge product*. A similar definition applies to the exterior product of covectors (i.e., elements $\alpha^1, \ldots, \alpha^k$ in V^*):

$$\alpha^1 \wedge \ldots \wedge \alpha^k = \frac{1}{k!}\sum_{\pi} \operatorname{sign}\pi\ \alpha^{\pi(1)} \otimes \ldots \otimes \alpha^{\pi(k)}. \tag{2.16}$$

One then sees that $\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k$ will be an element of A_kV, while $\alpha^1 \wedge \ldots \wedge \alpha^k$ will be an element of A^kV.

In particular, the exterior product of two vectors **v** and **w** in V will be:

$$\mathbf{v} \wedge \mathbf{w} = \tfrac{1}{2}(\mathbf{v} \otimes \mathbf{w} - \mathbf{w} \otimes \mathbf{v}) = -\,\mathbf{w} \wedge \mathbf{v}. \tag{2.17}$$

Just as the tensor product is associative, so is the exterior product:

$$(\mathbf{u} \wedge \mathbf{v}) \wedge \mathbf{w} = \mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w}), \tag{2.18}$$

so the expressions on the left-hand sides of (2.15) and (2.16) do not require parentheses in order to be well-defined.

A useful property of $\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k$ is that it will vanish iff $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly-dependent. Otherwise, they will define a k-frame that spans a k-plane. In particular:

$$\mathbf{v} \wedge \mathbf{w} = 0 \qquad \text{iff} \qquad \mathbf{w} = \lambda\,\mathbf{v}. \tag{2.19}$$

Although the exterior product $\mathbf{v} \wedge \mathbf{w}$ of two non-collinear vectors is anticommutative, when one groups $\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_{k+l}$ into the product $(\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k) \wedge (\mathbf{v}_{k+1} \wedge \ldots \wedge \mathbf{v}_{k+l})$ of a k-vector $\mathbf{A} = \mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k$ with an l-vector $\mathbf{B} = \mathbf{v}_{k+1} \wedge \ldots \wedge \mathbf{v}_{k+l}$, the product $\mathbf{A} \wedge \mathbf{B}$ can either be commute or anticommute, and in fact:

$$\mathbf{A} \wedge \mathbf{B} = (-1)^{kl}\, \mathbf{B} \wedge \mathbf{A}. \tag{2.20}$$

One establishes this by counting the number of transpositions of terms in $\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_{k+l}$ that it takes to permute all of the elements in $\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k$ past all of the elements in $\mathbf{v}_{k+1} \wedge \ldots \wedge \mathbf{v}_{k+l}$. This generalizes to the case of indecomposable multivectors.

A useful special case for the sake of calculation is when $\mathbf{A}$ and $\mathbf{B}$ are both bivectors:

$$\mathbf{A} \wedge \mathbf{B} = \mathbf{B} \wedge \mathbf{A}. \tag{2.21}$$

If $\mathbf{e}_i$ defines a basis on V then $\mathbf{v} = v^i\, \mathbf{e}_i$ and $\mathbf{w} = w^j\, \mathbf{e}_j$, so:

$$\mathbf{v} \wedge \mathbf{w} = \tfrac{1}{2}(v^i\, \mathbf{e}_i \otimes w^j\, \mathbf{e}_j - w^j\, \mathbf{e}_j \otimes v^i\, \mathbf{e}_i) = v^i\, w^j \tfrac{1}{2}(\mathbf{e}_i \otimes \mathbf{e}_j - \mathbf{e}_j \otimes \mathbf{e}_i) = v^i\, w^j\, \mathbf{e}_i \wedge \mathbf{e}_j\,,$$

but since $\mathbf{e}_i \wedge \mathbf{e}_j$ is antisymmetric in i and j, one must antisymmetrize the components, as well:

$$\mathbf{v} \wedge \mathbf{w} = \tfrac{1}{2}(v^i\, w^j - v^j\, w^i)\, \mathbf{e}_i \wedge \mathbf{e}_j\,. \tag{2.22}$$

One's initial reaction to this expression might be to say that all $\mathbf{e}_i \wedge \mathbf{e}_j$ define a basis for A^2V, and $\tfrac{1}{2}(v^i\, w^j - v^j\, w^i)$ are the components of $\mathbf{v} \wedge \mathbf{w}$ with respect to that basis, but the fact that the exterior product is antisymmetric would make:

$$\mathbf{e}_i \wedge \mathbf{e}_j = -\,\mathbf{e}_j \wedge \mathbf{e}_i\,, \qquad \mathbf{e}_i \wedge \mathbf{e}_i = 0, \tag{2.23}$$

so not all of the bivectors $\mathbf{e}_i \wedge \mathbf{e}_j$ would be non-zero and linearly-independent, but only half of the ones with $i \neq j$. Hence, the leading factor of $\tfrac{1}{2}$ in the summation on the right-hand side of (2.22) only serves to correct for the redundancy in the basis, and the components of $\mathbf{v} \wedge \mathbf{w}$ with respect to the redundant basis $\mathbf{e}_i \wedge \mathbf{e}_j$ will be:

$$(\mathbf{v} \wedge \mathbf{w})^{ij} = v^i\, w^j - v^j\, w^i. \tag{2.24}$$

That means that the number of linearly-independent elements in the set $\{\mathbf{e}_i \wedge \mathbf{e}_j\,,\ i \neq j\}$, and thus the dimension of A^2V, will be equal to one-half the number of non-diagonal elements in an $n \times n$ matrix, namely:

$$\dim A^2V = \tfrac{1}{2} n\,(n-1). \tag{2.25}$$

Of course, one notes that the $(\mathbf{v} \wedge \mathbf{w})^{ij}$ in (2.24) are also the components $(\mathbf{v} \times \mathbf{w})^k$ of the vector product $\mathbf{v} \times \mathbf{w}$ when one cyclically permutes ijk. Hence, the exterior product amounts to a generalization of that cross product that can be defined for vector spaces of any dimension, and not just 3. However, there is another face to the cross product on $\mathbb{R}^3$ that leads in a different direction, namely, the fact that it defines a Lie algebra on $\mathbb{R}^3$ that is isomorphic to the Lie algebra of infinitesimal three-dimensional Euclidian rotations. Therefore, one should not abandon the cross product as a "quaint anachronism" just yet.

In the more general case of $k > 2$, one finds that $\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k$ is non-zero iff the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly-independent; i.e., they define a k-frame on V. Thus, since one cannot find more than n linearly-independent vectors in an n-dimensional vector space, one must have that $\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k = 0$ whenever $k > n$; i.e., all k-vectors vanish for $k > n$.

If $\{\mathbf{e}_i, i = 1, \ldots, n\}$ is a basis for V then the number of linearly-independent exterior products $\mathbf{e}_{i_1} \wedge \ldots \wedge \mathbf{e}_{i_k}$ of k of them will be equal to the number of ways of choosing k members of that set of n elements when the order is irrelevant; i.e., the number of combinations of n things taken k at a time. Hence, the dimension of A^kV will be:

$$\dim A^kV = \binom{n}{k} = \frac{n!}{k!(n-k)!}. \tag{2.26}$$

That already indicates a duality in terms of dimension, since k and $n - k$ appear symmetrically in the denominator. Hence:

$$\dim A^kV = \dim \mathrm{A}^{n-k}V. \tag{2.27}$$

The general k-vector $\mathbf{A}$ can then be written in the component form:

$$\mathbf{A} = \frac{1}{k!} A^{i_1 \cdots i_k} \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_k}, \tag{2.28}$$

in which the components $A^{i_1 \cdots i_k}$ are completely-antisymmetric.

In the extreme case of $k = n$, if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are n linearly-independent vectors on an n-dimensional vector space V then they will define an n-frame on V; i.e., a basis. Any other set $\{\bar{\mathbf{v}}_1, \ldots, \bar{\mathbf{v}}_n\}$ of n linearly-independent vectors on V will define another basis and must then be related to the first one by an invertible linear matrix A_i^j:

$$\bar{\mathbf{v}}_i = \mathbf{v}_j A_i^j, \tag{2.29}$$

so:

$$\bar{\mathbf{v}}_1 \wedge \ldots \wedge \bar{\mathbf{v}}_n = (\mathbf{v}_{i_1} \wedge \ldots \wedge \mathbf{v}_{i_n}) A_1^{i_1} \cdots A_n^{i_n}.$$

However, one finds that the indices $i_1 \ldots i_n$ all amount to permutations of $1 \ldots n$, so if we introduce the totally-antisymmetric Levi-Civita symbol:

$$\varepsilon_{i_1\cdots i_n} = \begin{cases} +1 & \text{if } i_1\cdots i_n \text{ is an even permutation of } 12\ldots n \\ -1 & \text{if } i_1\cdots i_n \text{ is an odd permutation of } 12\ldots n \\ 0 & \text{otherwise} \end{cases} \tag{2.30}$$

then the expression for $\bar{\mathbf{v}}_1 \wedge \ldots \wedge \bar{\mathbf{v}}_n$ can be expressed in the form:

$$\bar{\mathbf{v}}_1 \wedge \ldots \wedge \bar{\mathbf{v}}_n = \frac{1}{n!}\varepsilon_{i_1\cdots i_n} A_1^{i_1} \cdots A_n^{i_n} \mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_n = (\det A)\, \mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_n . \tag{2.31}$$

This not only says that there is only one linearly-independent vector in A^nV (i.e., A^nV is one-dimensional) but that the concept of the exterior product has much in common with the theory of the determinant ([1]).

The algebra that is defined on A^*V comes from the extension of the exterior product to products of k-vectors and l-vectors to give $k+l$-vectors. Namely, if $\mathbf{A} = \frac{1}{k!}A^{i_1\cdots i_k}\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_k}$ is a k-vector and $\mathbf{B} = \frac{1}{l!}B^{j_1\cdots j_l}\mathbf{e}_{j_1} \wedge \cdots \wedge \mathbf{e}_{j_l}$ is an l-vector then the $k+l$-vector $\mathbf{A} \wedge \mathbf{B}$ will be:

$$\mathbf{A} \wedge \mathbf{B} = \frac{1}{(k+l)!} A^{[i_1\cdots i_k} B^{j_1\cdots j_l]}\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_k} \wedge \mathbf{e}_{j_1} \wedge \cdots \wedge \mathbf{e}_{j_l} . \tag{2.32}$$

Hence, A_*V with $\wedge$, like $\otimes_*V$ with $\otimes$, is a graded, associative algebra with unity. However, unlike $\otimes_*V$, the direct sum $A_*V = \mathbb{R} \oplus V \oplus A_2V \oplus \ldots \oplus A_nV$ will terminate after a finite number of summands.

Although we have been discussing the exterior algebra over V (i.e., multivectors), the same considerations will apply analogously to the exterior algebra A^*V over V^* (i.e., algebraic k-forms).

d. Interior product. – The exterior product allows one to define a linear operator $e_{\mathbf{A}}$: $A_lV \to A_{k+l}V$, $\mathbf{B} \to \mathbf{A} \wedge \mathbf{B}$ that takes l-vectors to $k+l$-vectors when one is given a k-vector $\mathbf{A}$; this basically amounts to left-multiplication by $\mathbf{A}$. Similarly, if one is given a k-form α then one can define a linear operator $e_\alpha : A^lV \to A^{k+l}V$, $\beta \to \alpha \wedge \beta$.

If one is given a k-vector $\mathbf{A}$ then one can also define a linear map $i_{\mathbf{A}} : A^lV \to A^{l-k}V$, that reduces the rank of an l-form α ($l > k$) to $l - k$ by defining the *interior product* of α by $\mathbf{A}$ to be the $l - k$-vector $i_{\mathbf{A}}\alpha$ such that if $\mathbf{B}$ is any $k–l$-vector then:

$$(i_{\mathbf{A}}\alpha)(\mathbf{B}) = \alpha(\mathbf{A} \wedge \mathbf{B}). \tag{2.33}$$

Note that here we are making use of the fact that the l-form α can be regarded as a linear functional on l-vectors, as well as an l-linear functional on vectors.

([1]) For more on this relationship, one might confer the aforementioned book on linear algebra by Hoffman and Kunze [**1**].

An analogous construction will give one the interior product of an l-vector $\mathbf{A}$ by a k-form α ($k < l$) to product an l–k-vector. Namely, $i_\alpha : A_l V \to A_{l-k} V$ will take the l-vector $\mathbf{A}$ to the l-k-vector $i_\alpha \mathbf{A}$ with the property that if β is any k-form then:

$$\beta(i_\alpha \mathbf{A}) = (\alpha \wedge \beta)(\mathbf{A}). \tag{2.34}$$

In the simplest case where $k = 1$, so $\mathbf{v}$ is a vector, if one can represent α in the form $\alpha_1 \wedge \ldots \wedge \alpha_k$, in which the α_i's are all 1-forms, then:

$$i_{\mathbf{v}}\alpha = \sum_{i=1}^{k} (-1)^{i+1} \alpha_1 \wedge \cdots \wedge i_{\mathbf{v}}\alpha_i \wedge \cdots \wedge \alpha_k = \sum_{i=1}^{k} (-1)^{i+1} \alpha_i(\mathbf{v})\, \alpha_1 \wedge \cdots \wedge \widehat{i_{\mathbf{v}}\alpha_i} \wedge \cdots \wedge \alpha_k ; \tag{2.35}$$

once more, the caret implies omission. In particular:

$$i_{\mathbf{v}}(\alpha \wedge \beta) = \alpha(\mathbf{v})\, \beta - \beta(\mathbf{v})\, \alpha . \tag{2.36}$$

If a k-form α is expressed in the component form $\alpha = \frac{1}{k!} \alpha_{i_1 \cdots i_k} \theta^{i_1} \wedge \cdots \wedge \theta^{i_1}$ for some choice of coframe θ^i then the components of $i_{\mathbf{v}}\alpha$ will be simply:

$$(i_{\mathbf{v}}\alpha)_{i_1 \cdots i_{k-1}} = \alpha_{i_1 \cdots i_k} v^{i_k} . \tag{2.37}$$

As usual, analogous constructions will apply to the interior product of an l-vector $\mathbf{A}$ by a k-form α.

e. Poincaré isomorphism. – Of particular interest is the case in which one takes the interior product of an n-form on an n-dimensional vector space with a k-vector $\mathbf{A}$ or dually, an n-vector with a k-form α. One will then have linear maps $i_{\mathbf{A}} : A^n \to A^{n-k}$ and $i_\alpha : A_n \to A_{n-k}$. Hence, if one chooses a particular non-zero n-form V then one can define a linear map $\# : A_k \to A^{n-k}$, $\mathbf{A} \mapsto i_{\mathbf{A}} V$. One finds that since $i_{\mathbf{A}} V = 0$ iff $\mathbf{A} = 0$, the map will be one-to-one, and since the two vector spaces A_k and A^{n-k} have the same dimension, the map # will be a linear isomorphism, which one calls the *Poincaré isomorphism*. The inverse isomorphism $\#^{-1} : A^{n-k} \to A_k$ can also be obtained by choosing a non-zero n-vector $\mathbf{V}$ such that $V(\mathbf{V}) = 1$, and defining $\#^{-1}$ to take the n–k-form α to the k-vector $i_\alpha \mathbf{V}$.

Since the vector spaces A^n and A_n are one-dimensional, any two choices of non-zero $\mathbf{V}$ and V will differ by a non-zero scalar. Such a choice of non-zero n-vector or n-form on an n-dimensional space is referred to as a *volume element* for the space. That is because if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly-independent then they will span an n-dimensional parallelepiped such that the non-zero number $V(\mathbf{v}_1, \ldots, \mathbf{v}_n) = V(\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_n)$ can be regarded as its volume (relative to V).

A common way of obtaining a volume element is to first choose an n-frame $\mathbf{e}_i$ for V and its reciprocal coframe θ^i for V^*. One can then define:

$$\mathbf{V} = \mathbf{e}_1 \wedge \ldots \wedge \mathbf{e}_n = \frac{1}{n!}\varepsilon^{i_1\cdots i_n}\mathbf{e}_{i_1}\wedge\cdots\wedge\mathbf{e}_{i_n}\,, \tag{2.38}$$

$$V = \theta^1 \wedge \ldots \wedge \theta^n = \frac{1}{n!}\varepsilon_{i_1\cdots i_n}\theta^{i_1}\wedge\cdots\wedge\theta^{i_n}\,, \tag{2.39}$$

in which $\varepsilon^{i_1\cdots i_n}$ and $\varepsilon_{i_1\cdots i_n}$ are the contravariant and covariant completely-antisymmetric Levi-Civita symbols for n dimensions.

For instance, if $\mathbf{v} = v^i\,\mathbf{e}_i$ is a vector in V then the components of its Poincaré dual will be:

$$(\#\,\mathbf{v})_{i_1\cdots i_{k-1}} = \varepsilon_{i_1\cdots i_n}v^{i_n}\,, \tag{2.40}$$

and if $\mathbf{A} = \mathbf{v} \wedge \mathbf{w} = \frac{1}{2}(v^i\,w^j - v^j\,w^i)\,\mathbf{e}_i \wedge \mathbf{e}_j$ is a bivector then the components of its dual will be:

$$(\#\,\mathbf{A})_{i_1\cdots i_{k-2}} = \tfrac{1}{2}\varepsilon_{i_1\cdots i_{n-1}i_n}v^{i_{n-1}}w^{i_n}\,. \tag{2.41}$$

Thus, the Levi-Civita symbols can be regarded as either the components of the volume elements with respect to a chosen basis or the components of the matrix of the Poincaré isomorphism.

An important point to notice is that when one performs successive interior products of V (or $\mathbf{V}$) the order of exterior multiplication will get inverted; i.e.:

$$i_{\mathbf{a}}\,i_{\mathbf{b}}\,V = i_{\mathbf{b}\wedge\mathbf{a}}\,V. \tag{2.42}$$

It is essential to recognize that although # is a linear isomorphism of vector spaces, it is not, however, an *algebra* isomorphism. That is, it does not take exterior products to exterior products.

$$\#\,(\mathbf{A} \wedge \mathbf{B}) \neq \#(\mathbf{A}) \wedge \#(\mathbf{B}). \tag{2.43}$$

In fact, the ranks of the resulting forms do not even match up, since if $\mathbf{A}$ is a k-form and $\mathbf{B}$ is an l-form, so $\mathbf{A} \wedge \mathbf{B}$ is a $k+l$-form, then $\#(\mathbf{A} \wedge \mathbf{B})$ will be an $n-k-l$-form, but $\#(\mathbf{A})$ will be an $n-k$-*form* and $\#(\mathbf{B})$ will be an $n-l$-*form*, so that would make $\#(\mathbf{A}) \wedge \#(\mathbf{B})$ a $2n-k-l$-form. Those two ranks would be consistent iff $n = 0$.

f. Hodge duality. – When one has chosen a scalar product for $\mathbb{R}^n$ (i.e, a metric tensor g), along with a volume element V, one can use the linear isomorphism of $\mathbb{R}^n$ with its dual space $\mathbb{R}^{n*}$ that is defined by the metric to define isomorphisms $* : \Lambda^k \to \Lambda^{n-k}$ that one refers to as *Hodge duality*.

In particular, the metric g defines a linear isomorphism $\iota_g : \mathbb{R}^n \to \mathbb{R}^{n*}$ that takes every vector $\mathbf{v}$ to the linear functional (i.e., covector) $\iota_g\,(\mathbf{v})$, which has the property that for every vector $\mathbf{w}$:

$$\iota_g(\mathbf{v}) = g(\mathbf{v}, \mathbf{w}). \tag{2.44}$$

Its component form is simply the process of lowering the index on v^i using the metric g:

$$v_i = g_{ij} v^j. \tag{2.45}$$

The inverse linear isomorphism $\iota_g^{-1} : \mathbb{R}^{n*} \to \mathbb{R}^n$ is easiest to describe by saying that it raises the index on v_i:

$$v^i = g^{ij} v_j \qquad (g^{ik} g_{kj} = g_j g^{ki} = \delta^i_j). \tag{2.46}$$

The linear isomorphism ι_g can be extended to linear isomorphisms $\iota_g : \otimes_k \to \otimes^k$ by tensoring the vector isomorphisms on the decomposable elements:

$$\iota_g(\mathbf{v}_1 \otimes \ldots \otimes \mathbf{v}_k) = \iota_g(\mathbf{v}_1) \otimes \ldots \otimes \iota_g(\mathbf{v}_k), \tag{2.47}$$

and extending by linearity to the indecomposable ones.

Similarly, these isomorphisms restrict to corresponding ones $\iota_g : A_k \to A^k$ when one replaces the tensor products with exterior products.

In order to define the Hodge star isomorphism $* : A^k \to A^{n-k}$, one needs only to compose the inverse isomorphism $\iota_g^{-1} : A^k \to A_k$ with the Poincaré isomorphism $\# : A_k \to A^{n-k}$ that is defined by a choice of volume element V:

$$* = \# \cdot \iota_g^{-1}. \tag{2.48}$$

One of the common applications of the * operator is in electromagnetism, where one applies it to the electromagnetic field strength 2-form $F = \frac{1}{2} F_{\mu\nu}\, dx^\mu \wedge dx^\nu$ to get the 2-form $*F$, whose components with respect to the same natural coframe will be:

$$*F_{\mu\nu} = \tfrac{1}{2} \varepsilon_{\mu\nu\kappa\lambda} F^{\kappa\lambda}, \tag{2.49}$$

in which we have implicitly raised the indices on $F_{\mu\nu}$ using the metric tensor g.

Since * is defined for all k, one can iterate it and get:

$$*^2 = \# \cdot \iota_g^{-1} . \# \cdot \iota_g^{-1} = \operatorname{sign}(g)\, (-1)^{n(n-p)} I, \tag{2.50}$$

in which we have defined sign (g) to be the product of its diagonal elements in an orthonormal frame. For instance, when g is Euclidian, sign (g) = + 1, and when g is four-dimensional Minkowski space (with either sign convention), sign (g) = – 1.

Hence, for three-dimensional Euclidian space:

$$*^2 = I, \tag{2.51}$$

while the corresponding statement for four-dimensional Minkowski space would be:

$$*^2 = \begin{cases} +I & k \text{ odd} \\ -I & k \text{ even} \end{cases}. \tag{2.52}$$

Hence, one sees from (2.50) that the inverse of * will be:

$$*^{-1} = \text{sign}\,(g)\,(-1)^{n\,(n-p)}\,*. \tag{2.53}$$

One can also define the Hodge star isomorphism $* : A_k \to A_{n-k}$ on k-vectors, instead of k-forms, by the opposite sequence of maps:

$$* = \iota_g^{-1} \cdot \#, \tag{2.54}$$

although, admittedly, that isomorphism does not get as much attention in mainstream differential geometry.

3. Exterior derivative. – The exterior derivative operator on exterior differential forms generalizes not only the gradient, curl, and divergence of vector fields on $\mathbb{R}^3$, but also something that they called the "bilinear covariant" in studies of the Pfaff equation that were made in the Nineteenth Century.

In order to define it, we first generalize from differentiable ([1]) vector fields on $\mathbb{R}^3$ to *exterior differential k-forms on* $\mathbb{R}^n$, which we define to be differentiable maps $\alpha : \mathbb{R}^n \to A^k\,\mathbb{R}^n$. Similarly, we shall denote the (infinite-dimensional) vector space of all such maps by $\Lambda^k(\mathbb{R}^n)$, or simply Λ^k, when there is no risk of confusion. In particular, 0-forms in this case will be simply differentiable maps on $\mathbb{R}^n$, while 1-forms will be covector fields on $\mathbb{R}^n$. Furthermore, the term "exterior differential k-form" will usually be abbreviated to simply "k-form" whenever the algebraic k-forms are no longer at issue.

The *exterior derivative* operator is a linear map $d_\wedge : \Lambda^k \to \Lambda^{k+1}$ that amounts to an antisymmetrized differential. As it turns out (cf., Warner [**3**], Cartan [**4**]), it can be defined uniquely by the following properties:

a. $d_\wedge f = df$ when f is a 0-form.

b. It is an *anti-derivation*. Hence if α is a differential k-form and β is a differential l-form then:

([1]) We shall typically say "differentiable" to mean "sufficiently differentiable"; i.e., continuously differentiable to as many orders as is required by the analysis. Often, it is simplest to say "smooth" to mean "continuously-differentiable to all orders."

$$d_\wedge(\alpha \wedge \beta) = d_\wedge\alpha \wedge \beta + (-1)^k \alpha \wedge d_\wedge\beta. \tag{3.1}$$

c. Its square always vanishes:

$$d_\wedge d_\wedge \alpha = 0 \qquad \text{for all } \alpha. \tag{3.2}$$

Of particular interest is the component form of $d_\wedge\alpha$ when $\alpha = \alpha_i\, \theta^i$ is a 1-form:

$$d_\wedge\alpha = d_\wedge(\alpha_i\, \theta^i) = d\alpha_i \wedge \theta^i + \alpha_i\, d_\wedge\theta^i\,. \tag{3.3}$$

Here, we must address the difference between the 1-forms θ^i (i.e., coframe fields on $\mathbb{R}^n$) for which $d_\wedge\theta^i$ vanishes and the ones that do not.

When the exterior derivative $d_\wedge\alpha$ of a k-form α vanishes, it will be called *closed*. When α is itself the exterior derivative $d_\wedge\beta$ of some k–1-form β, it will be called *exact*. From (3.2), every exact form will be closed. The converse happens to be true for $\mathbb{R}^n$ only because it has a very elementary topology (viz., it is *contractible*). The latter statement is called the *Poincaré lemma*, since it first showed up in his work on "analysis situs," which was the precursor to homology theory. When the vector space $\mathbb{R}^n$ is replaced with a more topologically-general differentiable manifold (see the next Appendix), one can say only that every closed form is locally exact; i.e., there is some neighborhood of each point upon which that statement will be true.

When one defines a coordinate system $\{x^i,\ i = 1, \ldots, n\}$ for $\mathbb{R}^n$ [so $x^i : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto x^i(x)$ is differentiable], the 1-forms dx^i will be exact and linearly-independent; hence, one will also have $d_\wedge dx^i = 0$. One calls dx^i the *natural coframe field* that is associated with the coordinate system. However, not all coframe fields are natural, since there exist coframe fields θ^i for which $d_\wedge\theta^i$ is non-vanishing for some values of i. Such coframe fields are called *anholonomic*, while the case in which all $d_\wedge\theta^i$ vanish is *holonomic*; from the fact that $d_\wedge d_\wedge = 0$, all natural coframe fields are holonomic; the converse is true, but harder to prove.

If we return to (3.3) in the case of a natural frame field (so $\theta^i = dx^i$) then we can continue our calculations and get:

$$d_\wedge\alpha = d_\wedge(\alpha_i\, dx^i) = d\alpha_i \wedge dx^i = \partial_j\, \alpha_i\, dx^j \wedge dx^i = \tfrac{1}{2}(\partial_i\, \alpha_j - \partial_j\, \alpha_i)\, dx^i \wedge dx^j. \tag{3.4}$$

Hence, the components of $d_\wedge\alpha$ with respect to this natural coframe field will be $\partial_i\, \alpha_j - \partial_j\, \alpha_i$, which are also the components (up to a permutation of the indices) of the curl of the vector field whose components are α^i, if one can ignore the difference between covariant and contravariant components, which is the case for Euclidian vector spaces.

Going on to $d_\wedge\alpha$ when $\alpha = \frac{1}{2}\,\alpha_{ij}\, dx^i \wedge dx^j$ is a 2-form, one gets:

$$d_\wedge\alpha = d_\wedge(\tfrac{1}{2}\,\alpha_{ij}\, dx^i \wedge dx^j) = \tfrac{1}{2}\, d\alpha_{ij} \wedge dx^i \wedge dx^j = \tfrac{1}{2}\,\partial_k\, \alpha_{ij}\, dx^k \wedge dx^i \wedge dx^j,$$

and after dropping the factor of 1/2 and completely antisymmetrizing the components $\partial_k \alpha_{ij}$, one will get:

$$d_\wedge \alpha = \frac{1}{3!}(\partial_i \alpha_{jk} + \partial_j \alpha_{ki} + \partial_k \alpha_{ij})\, dx^k \wedge dx^i \wedge dx^j, \tag{3.5}$$

but since the α_{ij} are already antisymmetric in ij, each of them will be counted twice, and the components of $d_\wedge \alpha$ will become $\frac{1}{3}(\partial_i \alpha_{jk} + \partial_j \alpha_{ki} + \partial_k \alpha_{ij})$.

In three dimensions, one can represent α_{jk} as $\varepsilon_{ijk}\, \alpha_i$, and:

$$d_\wedge \alpha = (\partial_i \alpha^i)\, V, \tag{3.6}$$

which clearly agrees with the divergence of the vector field on $\mathbb{R}^3$ whose components are α^i.

4. Divergence operator. – The divergence operator can also be generalized to n dimensions, just like the curl, and in fact, one only needs to have a volume element V in order to define the generalized divergence operator in terms of the exterior derivative operator.

First, we define the (infinite-dimensional) vector spaces $\Lambda_k\, \mathbb{R}^n$ of differentiable functions $\mathbf{A} : \mathbb{R}^n \to A_k$, which we call *k-vector fields on* $\mathbb{R}^n$. Next, we assume that we have chosen a volume element V on $\mathbb{R}^n$, so we can define the Poincaré isomorphisms # : $\Lambda_k \to \Lambda^{n-k}$ as the obvious extensions of the isomorphisms from A_k to A^{n-k}.

Basically, one defines the divergence operator div : $\Lambda_k \to \Lambda_{k-1}$ to be the "adjoint" of $d_\wedge$ under the Poincaré isomorphism:

$$\mathrm{div} = \#^{-1} \cdot d_\wedge \cdot \#. \tag{4.1}$$

div is then a linear map between the vector spaces Λ_k and Λ_{k-1} that is the composition of maps:

$$\Lambda_k \xrightarrow{\#} \Lambda^{n-k} \xrightarrow{d_\wedge} \Lambda^{n-k+1} \xrightarrow{\#^{-1}} \Lambda_{k-1}$$

When one applies this definition of div to a vector field $\mathbf{v} = v^i\, \partial_i$, in which $\partial_i = \partial / \partial x^i$ is the natural frame field that is associated with the coordinate system x^i, one will get the following function on $\mathbb{R}^n$:

$$\#^{-1} \cdot d_\wedge \cdot \#(\mathbf{v}) = \#^{-1} \cdot d_\wedge i_\mathbf{v} V = (d_\wedge i_\mathbf{v} V)(\mathbf{V}). \tag{4.2}$$

In component form, that will be:

$$\#^{-1} \cdot d_\wedge \cdot \#\, (v^i\, \partial_i) = \#^{-1} \cdot d_\wedge (v^i\, \#\partial_i) = \#^{-1}(dv^i \wedge \#\partial_i) = \#^{-1}(\partial_j v^i\, dx^j \wedge \#\partial_i)$$

$$= \partial_j v^i \#^{-1}(dx^j \wedge \#\partial_i) = \partial_j v^i \#^{-1}[\delta^i_j V] = (\partial_i v^i) \#^{-1}(V) = \partial_i v^i,$$

which condenses down to:

$$\text{div } \mathbf{v} = \partial_i v^i. \tag{4.3}$$

Therefore, this generalized divergence operator certainly agrees with the conventional divergence of vector fields.

One finds that div has a dual property to one of $d_\wedge$ in that:

$$\text{div} \cdot \text{div} = \#^{-1} \cdot d_\wedge \cdot \#\#^{-1} \cdot d_\wedge \cdot \# = \#^{-1} \cdot d_\wedge\, d_\wedge \cdot \# = 0, \tag{4.4}$$

but since the # isomorphism does not preserve the exterior product, one finds that div is not an anti-derivation, as is $d_\wedge$.

One can then define dual notions to closed and exact forms by saying that a k-vector field **A** is *co-closed* iff div **A** = 0 and is *co-exact* iff there is a k+1-vector field **B** such that **A** = div **B**. Hence, every co-exact multivector field is co-closed, and since we are considering only fields on $\mathbb{R}^n$, the converse will also be true, by the Poincaré lemma.

It is more conventional in mainstream differential geometry to define the "codifferential" operator $\delta: \Lambda^k \to \Lambda^{k-1}$ by using the Hodge star *, which involves the use of a metric g, in place of the Poincaré isomorphism #; i.e.:

$$\delta = *^{-1} \cdot d_\wedge \cdot * = \text{sign}(g)\,(-1)^{k(n-k)} * d_\wedge *, \tag{4.5}$$

in which the sign depends upon the signature type of the metric, the dimension of V, and the parity of k.

δ will then have the property that:

$$\delta^2 = 0, \tag{4.6}$$

which it inherits from $d_\wedge$, like the divergence operator.

As we will see in a later section, since the divergence operator is most intrinsically associated with the volume element, the introduction of a metric in order to define a codifferential operator makes it somewhat less than a natural construction.

5. Integration of differential forms. – To some, the true origin of the theory of exterior differential forms is in the theory of integration, since the integrand in the integral of a function $f(x^1, \ldots, x^n)$ on $\mathbb{R}^n$ over a region $D \subset \mathbb{R}^n$, namely:

$$\int_D f(x^1, \ldots, x^n)\, dx^1 \cdots dx^n ,$$

can be regarded as an exterior differential n-form on $\mathbb{R}^n$ (or at least on D). The reason that one typically does not represent that integrand in the form $f(x^1, \ldots, x^n)\, dx^1 \wedge \ldots \wedge dx^n$

is simply that the order of integration is typically fixed, so the sign of the integral will come from the specific order of integrations. However, if one changes the order of integrations then the sign of the integral would change in a manner that would be consistent with the introduction of the exterior product between the 1-forms in $dx^1 \wedge \ldots \wedge dx^n$. Since the latter n-form can be regarded as a volume element V_n on $\mathbb{R}^n$, one can also think of the integral above as the integral of an integral of an n-form over an n-dimensional region D in $\mathbb{R}^n$:

$$\int_D f(x^1, \ldots, x^n)\, dx^1 \cdots dx^n = \int_D f\, V_n\,. \tag{5.1}$$

The simplest n-dimensional region D in $\mathbb{R}^n$ is probably an n-cube $I^n = [0, 1] \times \ldots \times [0, 1]$. One can then use Fubini's theorem to convert the integral $\int_{I^n} f\, V_n$ into a succession of integrals over each individual coordinate:

$$\int_{I^n} f\, V_n = \int_{I^n} f(x^1, \ldots, x^n)\, dx^1 \cdots dx^n = \int_0^1 dx^n \int_0^1 dx^{n-1} \cdots \int_0^1 f\, dx^1\,. \tag{5.2}$$

Hence, if we can convert the more general integral (5.1) into a sum of integrals of this form then the integral of an n-form over an n-dimensional region D in $\mathbb{R}^n$ can be reduced to elementary integrations of the kind that one learns about in multivariable calculus.

The first step is to define a differentiable, singular cubic n-simplex in $\mathbb{R}^n$ to be a differentiable map $\sigma_n : I^n \to \mathbb{R}^n$, $x^i \mapsto \bar{x}^i$. (Although I^n is not a differentiable manifold, since differentiation is a local operation, one can extend σ_n to a differentiable map on an open neighborhood of I^n in any manner that one chooses and restrict the extension back to I^n.)

Since the use of the adjective "singular" in the context of singular simplexes implies the possibility that the map σ_n might not be an embedding, and in fact, it might map even to a single (zero-dimensional) point, one might think of the case in which the dimension is preserved as the "non-singular" case, although it is more traditional to refer to a non-singular n-simplex as a "n-cell." Hence, it will be tacitly assumed that σ_n is an embedding, and we shall drop the use of the adjective "singular" accordingly.

If the region $D \subset \mathbb{R}^n$ is the image of such a simplex σ_n then one define the integrand of an n-form $f\, V_n$ on D to equal the pull-back of $f\, V_n$ to I^n by σ_n:

$$\begin{aligned}\sigma_n^*(f\, V_n) &= f(\bar{x}^1(x), \ldots \bar{x}^n(x))\, \sigma_n^* V_n = f(\bar{x}^1(x), \ldots \bar{x}^n(x))\, \sigma_n^* V_n \\ &= f(\bar{x}^1(x), \ldots \bar{x}^n(x)) \det[\partial_j \bar{x}^i]\, dx^1 \wedge \ldots \wedge dx^n\,.\end{aligned}$$

Thus, we define the integral of $f\, V_n$ "on σ_n" to be:

$$\int_{\sigma_n} f\, V_n \equiv \int_{I^n} \sigma_n^*(f\, V_n) = \int_{I^n} f(\bar{x}^1(x), \ldots \bar{x}^n(x)) \det[\partial_j \bar{x}^i]\, dx^1 \wedge \ldots \wedge dx^n\,. \tag{5.3}$$

We then extend the type of region D to something that can be covered by the images of a *differentiable cubic n-chain* in $\mathbb{R}^n$, which is a finite formal linear combination of n-simplexes in $\mathbb{R}^n$:

$$c_n = \sum_a \lambda_a\, \sigma_n(a)\,. \tag{5.4}$$

One then extends the integral over an n-simplex to an integral over an n-chain by using linearity of integration:

$$\int_{c_n} f\, V_n \equiv \sum_a \lambda_a \int_{\sigma_n(a)} f\, V_n\,. \tag{5.5}$$

As it turns out, the kinds of regions D in $\mathbb{R}^n$ that can be broken up into a collection of deformed n-cubes that are identified at their faces are more topologically general than one might expect. Rather than go into the details at this point, we simply say that we shall typically confine our integrals to only such regions.

The n-cube I^n has $2n$ faces (namely, two for each dimension), which can be collectively regarded as its boundary ∂I^n. More precisely, when one gives all of the faces of I^n – say – the outward-pointing normal as an orientation, one can also regard opposite faces as having opposite signs. ∂I^n can then be represented as a formal sum of signed faces:

$$\partial I^n = \sum_{i=1}^{n} F_i(1) - F_i(0)\,. \tag{5.6}$$

For instance, in the 1-dimensional case, $\partial[0, 1] = 1 - 0$ ([1]), while in two dimensions:

$$\partial\,([0, 1] \times [0, 1]) = -\,[0, 1] \times 1 + [0, 1] \times 0 + 1 \times [0, 1] - 0 \times [0, 1]. \tag{5.7}$$

We indicate the orientation on the boundary edges in Fig. A.1:

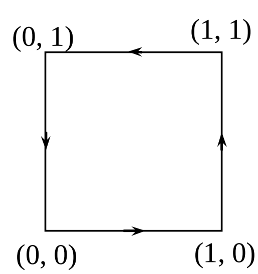

Figure A.1 – The boundary of an oriented square.

([1]) As a formal sum, we are treating the symbols 0 and 1 as abstract in their definitions, so the sum cannot be reduced to simply 1.

Note that if we take the boundary again then we will get:

∂^2 ([0, 1] × [0, 1])

$$= - (1, 1) + (0, 1) + (1, 0) - (0, 0) + (1, 1) - (1, 0) - (0, 1) + (0, 0) = 0.$$

In fact, this result generalizes to all other n, namely, the boundary of a boundary is always zero; i.e.:

$$\partial^2 = 0. \tag{5.8}$$

The boundary $\partial\sigma_n$ of an n-simplex σ_n is then the formal sum of the restrictions of σ_n to the faces of I^n :

$$\partial\sigma_n = \sum_{i=1}^{n} \sigma_n \,|\, F_i(1) - \sigma_n \,|\, F_i(0), \tag{5.9}$$

and since the faces of I^n look like I^{n-1}, so the restrictions of σ_n to each face become n–1-simplexes, one sees that $\partial\sigma_n$ will become an n–1-chain.

One can then extend by linearity to define the boundary of an n-chain c_n as in (5.4):

$$\partial c_n = \sum_a \lambda_a \, \partial\sigma_n(a). \tag{5.10}$$

There is an important theorem that relates to the integral of an n–1-form α over the boundary of an n-chain, namely:

$$\int_{\partial c_n} \alpha = \int_{c_n} d_\wedge \alpha. \tag{5.11}$$

Although this is usually just referred to as *Stokes's theorem*, it actually includes the fundamental theorem of calculus (n = 1), Green's theorem (n = 2), Stokes's and Gauss's theorems (n = 3). Furthermore, as one gets closer to the era in history when Stokes presumably proved his theorem, one finds an increasing amount of footnotes to the effect that Stokes was not the first to prove it, but he had learned of it from Lord Kelvin.

In order to get from (5.11) to Gauss's theorem (i.e., the divergence theorem), one must introduce a volume element in order to be able to define the Poincaré isomorphism #. If one expresses a as #**A** for some (unique) vector field **A** then:

$$\int_{\partial c_n} \#\mathbf{A} = \int_S d_\wedge \#\mathbf{A} = \int_S \#(\mathrm{div}\,\mathbf{A}). \tag{5.12}$$

If one also introduces a metric, so the Hodge star isomorphism * is well-defined then this can also be expressed in terms of * and the codifferential δ, but as we have pointed out above, the divergence is more intrinsic to volume elements and the # isomorphism than it is to the introduction of auxiliary metrics.

6. Lie derivative. – The concept of a Lie derivative is quite fundamental to continuum mechanics, since it amounts to differentiating geometric objects along the streamlines of a flow.

a. Systems of ordinary differential equations. – We would first like to point out that any vector field $\mathbf{X}(x) = X^i(x)\,\partial_i$ on $\mathbb{R}^n$ will define a system of n first-order ordinary differential equations by assuming that the vector field $\mathbf{X}$ amounts to the velocity vector field for a congruence of curves of the form $x^i(t)$:

$$\frac{dx^i}{dt} = X^i(x). \tag{6.1}$$

More precisely, this is a system of *autonomous* ordinary differential equations, since we are not assuming that $\mathbf{X}$ is also a function of t. (In fluid mechanics, that would correspond to steady flow.)

If a curve $x^i(t)$ satisfies the equations (6.1) then it will be an *integral curve* of the system, and the congruence of all of such integral curves will amount to a one-dimensional foliation of the differential system. Note that there is a difference between integrating the differential system (6.1) into a congruence of integral curves, which always exists, and actually specifying a parameterization of each integral curve, which is not always possible, except locally. That is, the congruence of integral curves is not always *hypersurface-normal* in the sense that there is some hypersurface for which all integral curves will intersect it transversally. Such a hypersurface would allow one to "synchronize" the parameters of the integral curves by defining a universal initial time point.

One way of specifying a common curve parameter for at least the integral curves that pass through some open subset U of the region in which the congruence (or $\mathbf{X}$ itself) is defined is to use the existence and uniqueness of "local flows" for the vector field $\mathbf{X}$, which amounts to the existence and uniqueness of local solutions to the differential system (6.1). The form of the existence and uniqueness solution that we shall use is that as long as $\mathbf{X}(x)$ is continuously differentiable, if one is given some open subset U that is transverse to $\mathbf{X}$ and an initial value t_0 of a parameter t that identifies U then there will always exist a (sufficiently-small) ε such that a one-parameter family of diffeomorphisms (onto) (1):

$$\Phi(t_0, t_0 + \Delta t) : U \to \mathbb{R}^n, \qquad x_0 \mapsto \Phi(t_0, t_0 + \Delta t)(x_0)$$

will exist for every $0 < \Delta t < \varepsilon$ such that if $x_0 = x(t_0)$ is the point along the integral curve $x(t)$ when $t = t_0$ then:

$$\Phi(t_0, t_0 + \Delta t)(x_0) = x(t_0 + \Delta t). \tag{6.2}$$

(1) A map $f : U \to \mathbb{R}^n$ is called a *diffeomorphism onto* iff it is one-to-one and continuously differentiable, and when one restricts to its image, the inverse map is also continuously differentiable.

Such a one-parameter family of diffeomorphisms is called a *local flow* for **X**, and the problems of extending U to a normal hypersurface (viz., an initial hypersurface) for the integral curves and extending ε to infinity are quite analytically involved ([1]). For instance, although one could conceivably patch together U's into a normal hypersurface, one could not necessarily use the same ε for all of them.

The functions $\Phi(t_0, t_0 + \Delta t)$ can be represented as an invertible matrix $\Phi^i_j(t_0, t_0 + \Delta t)$ that takes initial position vectors x_0^j at time t_0 to their time-evolutes at $t_0 + \Delta t$:

$$x^i(t_0 + \Delta t) = \Phi^i_j(t_0, t_0 + \Delta t)\, x_0^j\,. \tag{6.3}$$

Since taking the time derivative of this at t_0 will give:

$$\left.\frac{dx^i}{dt}\right|_{t=t_0} = X^i(x(t_0)) = \left.\frac{d\Phi^i_j}{dt}\right|_{t=t_0} x_0^j\,, \tag{6.4}$$

one calls the matrix $\Phi^i_j(t_0, t_0 + \Delta t)$ either the *time evolution* operator for the system or the *matrix of fundamental solutions*, since its columns will consist of the solutions to the system (6.1) when the initial point x_0^j is set equal to each of the canonical basis vectors for $\mathbb{R}^n$, namely, (1, 0, …, 0), (0, 1, 0, …, 0), …, (0, …, 0, 1).

A particularly useful special case in the present study is when **X** is a linear function on $\mathbb{R}^n$, and therefore by represented by a time-varying $n \times n$ real matrix:

$$\frac{dx^i}{dt} = X^i_j(t)\, x^j\,. \tag{6.5}$$

The time evolution matrix is defined everywhere that **X** is defined and for all Δt by simply exponentiating the time integral of the matrix $X^i_j(t)$:

$$\Phi^i_j(t_0, t_0 + \Delta t) = \exp \int_{t_0}^{t+\Delta t} X^i_j(\tau)\, d\tau\,. \tag{6.6}$$

In the event that $X^i_j(t)$ is not time-varying, the system (6.1) will be autonomous (or stationary) and the integral can be replaced with $\Delta t\, X^i_j$. Furthermore, the time evolution matrix will also become simply a function of Δt:

$$\Phi^i_j(\Delta t) = \exp\,(\Delta t\, X^i_j)\,. \tag{6.7}$$

([1]) Some good classic references on the existence and uniqueness of solutions to ordinary differential equations are Ince [**5**] and Coddington and Levinson [**6**], although the more modern notion of local flow is discussed in Arnol'd [**7**].

A very deep special case of **X** (x) is when it has zeros, since wherever **X** (x) = 0, one must have $dx^i / dt = 0$, which means that the zeroes of **X** will correspond to fixed points of the flow. Much of the structure of the flow of **X** can be obtained by examining the local stability of such fixed points, which will involve looking at the differential d **X**, and more to the point, the eigenvalues of its component matrix $\partial_j X^i$ at each fixed point.

b. Derivatives along a flow. – Assume that a differentiable function $f(x^i)$ is defined everywhere that **X** is defined, so one can speak of the variation of $f(t)$ along each integral curve $x^i(t)$, which one will get by composing the function f with the functions x^i :

$$f(t) = f(x^i(t)) = (f \cdot x^i)(t). \tag{6.8}$$

From the chain rule for differentiation, one will have:

$$\frac{df}{dt} = \frac{\partial f}{\partial x^i}\frac{dx^i}{dt} = X^i \partial_i f = \mathbf{X} f. \tag{6.9}$$

In other words, df / dt is the directional derivative of f in the direction of **X**, which is the immediate direction of the flow.

Although one could generalize this to any field $O(x)$ of geometric objects (vectors, tensors, differential forms) that is defined along the flow of **X**, the resulting directional derivative would not suffice to give a full picture of the variation of $O(x)$ along the integral curves. The main problem is that the objects $O(x(t))$ and $O(x(t + \Delta t))$ are not attached to the same point, so the spaces that they belong to [let us call them $E_{x(t)}$ and $E_{x(t+\Delta t)}$] must be identified with each other in some way before the definition:

$$\mathrm{L}_{\mathbf{X}} O \equiv \frac{dO}{dt} = \lim_{\Delta t \to 0} \frac{1}{\Delta t}[O(x(t+\Delta t) - O(x(t))] \tag{6.10}$$

will make sense.

For a vector field **Y**, the problem is in the fact that there is a difference between the composition **X** (**Y**f) and the composition **Y** (**X**f), as a directional derivative operator, namely:

$$[\mathbf{X}, \mathbf{Y}] = \mathbf{X}(\mathbf{Y}) - \mathbf{Y}(\mathbf{X}) = X^i \partial_i (Y^j \partial_j) - Y^i \partial_i (X^j \partial_j) = (X^i \partial_i Y^j - Y^i \partial_i X^j)\, \partial_j ,$$

which can be written in the form:

$$[\mathbf{X}, \mathbf{Y}] = (\mathbf{X}Y^i - \mathbf{Y}X^i)\, \partial_i . \tag{6.11}$$

Note that this calculation depends upon the basic assumption that:

$$\partial_i \partial_j = \partial_j \partial_i , \qquad \left(\text{i.e., } \frac{\partial^2}{\partial x^i \partial x^j} = \frac{\partial^2}{\partial x^j \partial x^i}\right) \tag{6.12}$$

which assumes that the functions that [**X**, **Y**] acts upon are twice continuously-differentiable.

One easily sees that [**X**, **Y**] does not have to vanish by the example of the infinitesimal rotations on $\mathbb{R}^3$:

$$\mathbf{X}_i = \tfrac{1}{2}\,\varepsilon_{ijk}\,(x^j\,\partial_k - x^k\,\partial_j), \tag{6.13}$$

which give:

$$[\mathbf{X}_i\,,\,\mathbf{X}_j] = \varepsilon_{ijk}\,\mathbf{X}_k\,. \tag{6.14}$$

In general, we set the *Lie derivative* of a function f and a vector field **Y** along the flow of the vector field **X** equal to:

$$\mathrm{L}_{\mathbf{X}} f = \mathbf{X}f, \qquad \mathrm{L}_{\mathbf{X}}\mathbf{Y} = [\mathbf{X}, \mathbf{Y}], \tag{6.15}$$

respectively, and extend to higher-rank contravariant tensors by imposing the demand that must act as a derivation on tensor products:

$$\mathrm{L}_{\mathbf{X}}\,(T_1 \otimes T_2) = (\mathrm{L}_{\mathbf{X}}\,T_1) \otimes T_2 + T_1 \otimes (\mathrm{L}_{\mathbf{X}}\,T_2), \tag{6.16}$$

which is then extended to indecomposable tensors by demanding linearity, as well.

In order to deal with the covariant tensors, when α is a 1-form, one defines:

$$\mathrm{L}_{\mathbf{X}}\,\alpha = d\,(\alpha(\mathbf{X})) + i_{\mathbf{X}}\,d_\wedge\alpha\,. \tag{6.17}$$

This is a special case of *Cartan's magic formula*, which applies to all exterior differential forms:

$$\mathrm{L}_{\mathbf{X}}\,\alpha = d_\wedge\,i_{\mathbf{X}}\alpha + i_{\mathbf{X}}\,d_\wedge\alpha\,, \tag{6.18}$$

namely, when α is a 1-form:

$$(\mathrm{L}_{\mathbf{X}}\,\alpha)(\mathbf{Y}) = \mathbf{Y}\,(\alpha(\mathbf{X})) + d_\wedge\alpha(\mathbf{X}, \mathbf{Y}). \tag{6.19}$$

If one uses the "intrinsic" formula for the exterior derivative of α:

$$d_\wedge\alpha(\mathbf{X}, \mathbf{Y}) = \mathbf{X}\,(\alpha(\mathbf{Y})) - \mathbf{Y}\,(\alpha(\mathbf{X})) - \alpha([\mathbf{X}, \mathbf{Y}]) \tag{6.20}$$

then

$$(\mathrm{L}_{\mathbf{X}}\,\alpha)(\mathbf{Y}) = \mathbf{X}\,(\alpha(\mathbf{Y})) - \alpha([\mathbf{X}, \mathbf{Y}]). \tag{6.21}$$

A particular enlightening special case of (6.18) is when one applies it to a volume element V. Since one must have $d_\wedge V = 0$, one will get:

$$\mathrm{L}_{\mathbf{X}}V = d_\wedge\,i_{\mathbf{X}}V = d_\wedge\,\#\mathbf{X} = \#(\mathrm{div}\,\mathbf{X}) = (\mathrm{div}\,\mathbf{X})\,V. \tag{6.22}$$

Hence, zero-divergence vector fields are distinguished by the fact that their flows are volume-preserving.

One can extend (6.17) to all covariant tensor fields by demanding that $\mathrm{L}_{\mathbf{X}}$ must be a linear derivation.

One can then combine the rules (6.15) for scalar fields and vector fields with the rule (6.17) for covector fields and then extend to all tensor fields by the remand that $L_{\mathbf{X}}$ must be a linear derivation.

References

1. K. Hoffman and R. Kunze, *Linear Algebra*, Prentice-Hall, NJ, 1962.
2. W. Greub, *Multilinear Algebra*, Springer, Berlin, 1967.
3. F. Warner, *Differentiable Manifolds and Lie Groups*, Scott Foresman, Glenview, IL, 1971.
4. H. Cartan, *Differential Forms*, Dover, NY, 2006.
5. E. L. Ince, *Ordinary Differential Equations*, Dover, Mineola, NY, 1956.
6. E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, NY, 1955.
7. V. I. Arnol'd:
   a. *Ordinary Differential Equations*, The MIT Press, MA, 1975.
   b. *Geometrical Methods in the Theory of Ordinary Differential Equations*, Springer, Berlin, 1983.

APPENDIX B

DIFFERENTIABLE MANIFOLDS

1. Topological spaces. – Since much of what one learns in a typical course in point-set topology is tailored to the various demands of many separate specialized branches of mathematics, such as differential topology, algebraic topology, functional analysis, and the topology of infinite-dimensional differentiable manifolds, we shall cull out only those notions that are relevant to the definition of a differentiable manifold.

a. A topology on a set. – The subject of topology had been around in the form of "analysis situs" for much of the late Nineteenth Century before Felix Haussdorff defined the more abstract concept of a topology on a set in his landmark work *Mengenlehre* (i.e., set theory), which first appeared around the time of World War I, although the 1927 second edition is the one that is still available [**1**]. Indeed, the methods of *analysis situs* (which means the same thing in Latin that *topos logos* does in Greek, namely, the "study of position") had much more in common with the modern concept of homology theory, such as the triangulation of spaces into elementary building blocks in order to describe the way that the space is "connected" and whether it has "holes" of various dimensions.

The basic idea of a topology τ on a set S ([1]) is that one defines a special class τ of subsets, which one calls *open* subsets that obey the basic axioms:

1. The set S and the empty set $\varnothing$ are elements of the class τ; i.e., they are open subsets.
2. The union of *any* family of open subsets will also be open.
3. The intersection of a *finite* family of open subsets will also be open.

Dually, one can define a class of *closed* subsets, which obey the axioms that come from the previous three by applying de Morgan's laws to the complements of open subsets, namely:

1. The set S and the empty set $\varnothing$ are closed subsets.
2. The intersection of *any* family of closed subsets will also be closed.
3. The union of a *finite* family of closed subsets will also be closed.

If one is given the open sets of a topology then one can define the closed sets to be the complements of open subsets, and vice versa. However, not all subsets will be either open or closed.

There are two topologies that always exist on any set S:

1. The *trivial* topology, which includes only the open subsets S and $\varnothing$.
2. The *discrete* topology, which includes all subsets.

([1]) A more modern reference than Haussdorff on point-set topology is Munkres [**2**].

Typically, those topologies are rarely of interest in physical models, although the discrete topology is usually given to finite groups in order to make then into (zero-dimensional) Lie groups, which we shall discuss in Appendix C.

Topologies are often defined by giving a *basis* for the topology. Namely, a basis for the topology τ is a sub-collection τ_b of the open subsets of τ such that and point of the set is contained in one of the subsets of τ_b, and any open subset in t can be expressed as the union of any number of open subsets in the basis or the intersection of a finite family of open subsets in the basis. For example, the real line $\mathbb{R}$ can be given a popular topology by using the open intervals (a, b) as a basis for the open subsets; here, one can include the possibilities that $a = -\infty$ or $b = +\infty$. The closed intervals $[a, b]$ would then be closed subsets of that topology, and the intervals of the forms $[a, b)$ or $(a, b]$ would be neither open nor closed.

When a set S is given a particular topology τ, one calls the pair (S, τ) a *topological space*. In general, a subset $A \subset S$ does not have to be either open or closed with respect to the topology τ. One can, however, associate A with an open subset and a closed subset. In the former case, one defines the *interior* of A to be the largest open subset $\overset{\circ}{A}$ that is a subset of A. (Here, "largest" means "with respect to the partial ordering of subset inclusion.") The *closure* of A is the smallest closed subset $\bar{A}$ that contains A. The *frontier* or *boundary* of A is the set difference $\partial A = \bar{A} - \overset{\circ}{A}$; that is, a boundary point is in the closure, but not the interior.

As an example of these constructions, take the real line with the topology that was just discussed. The interiors of the half-open intervals $[a, b)$ and $(a, b]$ would be (a, b), in either case, just as their closures would both be $[a, b]$, and their boundaries would be the set of endpoints $\{a, b\}$.

When one is given two topological spaces (S_1, τ_1) and (S_2, τ_2), one can define a topology on the Cartesian product $S_1 \times S_2$ of the two sets by using the Cartesian products $U_1 \times U_2$ of open subsets U_1 from τ_1 and U_2 from τ_2 as a basis for a topology that one calls the *product topology*. That construction can be iterated to a topology on the Cartesian product $S_1 \times \ldots \times S_n$ of a finite family of topological spaces. For instance, one can give $\mathbb{R}^n \equiv \mathbb{R} \times \ldots \times \mathbb{R}$ (n copies) a topology in that way by starting with the interval topology on $\mathbb{R}$.

Another common way of building up topological spaces from more elementary ones is the notion of a "quotient space." Suppose a set S has an equivalence relation ~ defined on it, namely, a relation $x \sim y$ between some pairs of elements that is:

1. Reflexive: $x \sim x$ for all x
2. Symmetric: $x \sim y$ iff $y \sim x$.
3. Transitive: If $x \sim y$ and $y \sim z$ then $x \sim z$.

The set $[x]$ of all $y \in S$ such that $y \sim x$ is called the *equivalence class* that x belongs to. If x and y are not related by ~ then $[x]$ and $[y]$ will be disjoint subsets. The set $S / \sim$ of all such $[x]$ is called the *quotient of S by the relation* ~. There is then a canonical projection $[.] : S \to S / \sim$, $x \mapsto [x]$ that one calls the *quotient map* that is associated with the relation.

Examples of equivalence relations that show up in topology mostly take the form of "identifications"; that is, one takes some set of points to be one equivalence class of "identified" points, while all other points are equivalent to only themselves. For instance, one can turn $[a, b]$ into a circle S^1 by identifying the points a and b into one equivalence class $[a] = [b] = \{a, b\}$ while any other equivalence class $[x] = \{x\}$ will contain only the single point x. That is essentially what one does with polar coordinates (r, θ) in a plane when one identifies $\theta = 0$ with $\theta = 2\pi$. One can also get S^1 from $\mathbb{R}$ by defining the equivalence relation $x \sim y$ iff there is some integer k such that $x - y = k$. One can then exhibit S^1 as $\mathbb{R} \to \mathbb{Z}$, and quotient map becomes $\mathbb{R} \to \mathbb{R} / \mathbb{Z}$, $x \mapsto [x]$. This has a close relationship to the concept of periodicity. One can also extend the aforementioned quotient constructions to produce the n-sphere S^n from a closed n-ball by identifying all of its boundary points to a single equivalence class and to product an n-torus T^n by defining the equivalence class on $\mathbb{R}^n$ that makes $x \sim y$ iff $x^i - y^i = k^i$ for each $i = 1, \ldots, n$, where k^i is an integer in each case. This example is relevant to the concerns of crystallography.

When S also has a topology t, one can define a topology on $S / \sim$ that one calls the *quotient topology*. A subset of $S / \sim$ is open in that topology iff its inverse image under the projection $[.]$ is an open subset of S. Once again, the most common quotient topological spaces that one tends to encounter are basically identification spaces, such as the identification of boundary points to a single equivalence class.

A *neighborhood* of a point x in a topological space (S, τ) is a subset N_x that contains an open subset U_x that contains x. U_x will then be itself a neighborhood of x that one refers to as an *open neighborhood*; one might also encounter closed neighborhoods.

A common way of obtaining neighborhoods is to first define a *metric* on S, which will be a function $d : S \times S \to \mathbb{R}$, $(x, y) \mapsto d(x, y)$ such that for all points x, y in S:

1. $d(x, y) \geq 0$.
2. $d(x, y) = 0$ iff $x = y$.
3. $d(x, y) = d(y, x)$.
4. $d(x, y) + d(y, z) \geq d(x, z)$.

The last axiom is called the "triangle inequality," since that is the way that things work for three points in $\mathbb{R}^2$ when the distance in question is defined by the Pythagorean law:

$$d(x, y) = \sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2}\,, \tag{1.1}$$

which defines the Euclidian metric. More generally, the Euclidian metric on $\mathbb{R}^n$ is defined by:

$$d(x, y) = \sqrt{(x^1 - y^1)^2 + \cdots + (x^n - y^n)^2}\,. \tag{1.2}$$

One can then define an *open d-ball* of radius r about a point x to be the set $B_x(r)$ of all y such that $d(x, y) < r$. If one defines the topology on the metric space (S, d) to have open d-balls for its basis then $B_x(r)$ will be an open neighborhood of x. Its closure will be the set of all y for which $d(x, y) \leq r$ and its boundary will be the set of all y such that $d(x, y) = r$. Such a topology is called a *metric topology*.

$\mathbb{R}$ can be given a metric topology when one defines the "distance" $d(x, y)$ from x to y to be:

$$d(x, y) = |\, y - x\,|. \tag{1.3}$$

The open ε-ball about a point x is then the open interval $(x - \varepsilon, x + \varepsilon)$, which is then an open subset of the interval topology. Conversely, any *finite* open interval (a, b) could be regarded as an open r-ball $(x - r, x + r)$ about the point x if one lets:

$$r = (b - a) / 2 \qquad \text{and} \qquad x = a + r = b - a. \tag{1.4}$$

Infinite open intervals can be expressed as infinite unions of finite (overlapping) open intervals. Hence, every open subset in a basis for the interval topology on $\mathbb{R}$ can also be regarded as an open subset in a basis for the metric topology, and *vice versa*.

$\mathbb{R}^n$ can be given a metric topology when one gives it the Euclidian metric. The open balls will then become balls in the geometrically-familiar sense of the term, and their boundaries will be n–1-spheres.

Note that the Minkowski scalar product $<.,.>$ on $\mathbb{R}^4$ does not define a metric in the present sense, since the equation $<\mathbf{x}, \mathbf{x}> = 0$ has more solutions than just $\mathbf{x} = 0$. Indeed, the set of all such $\mathbf{x}$ (viz., the light cone at the origin) plays an essential role in relativity. Hence, one should be careful when referring to the space-time "metric," since it is not a metric in the point-set topological sense.

b. Continuity of maps. – A map $f : S \to S'$, $x \mapsto f(x)$ from one topological space (S, τ) to another one (S', τ') is called *continuous* iff the inverse image (1) $f^{-1}(U)$ of every open subset U of S' is an open subset of S. This definition can also be phrased in terms of neighborhoods: f is continuous iff the inverse image of every neighborhood U_y of every point $y \in S'$ is a neighborhood of the point $x \in S$ such that $y = f(x)$. We illustrate that situation in Fig. A.1.

(1) That is, the set of all elements x in S such that $f(x)$ is an element of U.

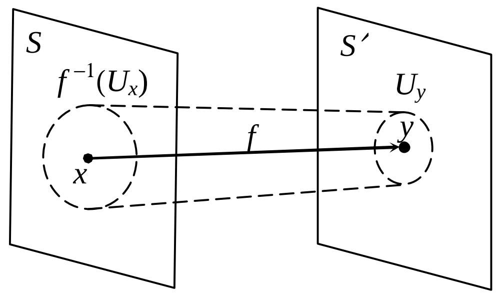

Figure A.1 - The continuity of a map between topological spaces.

In the case of the real line $\mathbb{R}$ with the interval (or metric) topology, one can look at an ε-ball about y and a δ-ball about x and recover the elementary calculus way of characterizing continuity by saying that "for every $\varepsilon > 0$, there is a $\delta > 0$ such that for all x such that $x - \delta < x < x + \delta$ one will have $y - \varepsilon < f(x) < y + \varepsilon$." Hopefully, one can see that the point-set topological way of describing continuity has a more intuitive appeal.

An even stronger requirement than continuity is "homeomorphism": $f : S \to S'$ is a *homeomorphism* iff it is continuous and invertible and its inverse is also continuous. Such a map between topological spaces will define a one-to-one correspondence between the points of the spaces, as well as the open subsets of the topologies. Hence, the two topological spaces will be equivalent as far as their points and open subsets are concerned.

We have already seen that when $\mathbb{R}$ it is given either the interval topology or the metric topology, an open subset of one topology will be an open subset of the other topology. Hence, the identity map, which takes every point of $\mathbb{R}$ to itself, will be a homeomorphism of the two topologies. Similarly, since every n-cube can be expressed as an infinite union of n-balls, and *vice versa*, the identity map on $\mathbb{R}^n$ will be a homeomorphism of the two topologies on that space. In other words, the two topologies on $\mathbb{R}^n$ are topologically equivalent, which is the soul of all homeomorphisms.

If one wishes to see how an open n-ball and an open n-cube are homeomorphic, one might inscribe a disc inside a square as an example and look at the association of points that one gets by radial contraction of a point on the square to a point on the circle or radial expansion of the point on the circle to a point on the square. If one thinks of the interior of the square as being filled up with a continuously-infinitude sequence of concentric squares and circles then that construction can be applied to each of them to give a homeomorphism of a square with a disc. We illustrate this in Fig. A.2:

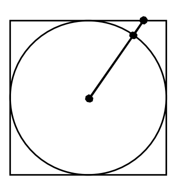

Figure A.2. – The homeomorphism of a square with a disc.

A weaker form of homeomorphism that will play a major role shortly is that of *local homeomorphism.* Namely, the map $f : S \to S'$ between the topological spaces S and S is a local homeomorphism iff every $x \in S$ has an open neighborhood U_x on which the restriction of f is a homeomorphism onto its image $f(U_x)$; one also says that $f : U_x \to S'$ is a *homeomorphism onto* in this case.

c. Topological properties. – A property of a topological space is said to be a *topological property* iff it is preserved by homeomorphisms. In many cases, topological properties are preserved by continuous maps, to begin with. The openness or closedness of a set would be examples of properties that are preserved by homeomorphisms, but not by continuous maps.

Some of the topological properties that one usually encounters in the study of point-set topology are the various separation axioms, connectedness, and compactness.

The most common separation axiom to impose upon a topological space is that it is "Haussdorff," which was actually one of that mathematician's original axioms for a topology to begin with, although nowadays it is usually treated as an extra axiom. A topological space (S, τ) is said to be a *Haussdorff space* iff for every two distinct points x, $y \in S$ there are two disjoint open neighborhoods U_x and U_y of x and y. One has that the inverse image of a Haussdorff space under a continuous map is Haussdorff, but not generally the direct image. That is because a map can take distinct points to the same point and disjoint subsets to intersecting ones, but the same thing is not true for the inverse image of a map. Hence, if $y, y' \in S'$ are distinct points in a Haussdorff space (S', τ') then any two points x, x' in S that map to y, y', resp., under a continuous map $f : S \to S'$ will be distinct. Similarly, if U_y, $U_{y'}$ are disjoint open neighborhoods of y, y' then their inverse images $f^{-1}(U_y)$, $f^{-1}(U_{y'})$, resp., under f will still be disjoint subsets. When f is a homeomorphism, the aforementioned property of f will also be shared by its inverse f^{-1}. Thus, any topological space that is homeomorphic to a Haussdorff space will be Haussdorff.

A topological space (S, τ) is *connected* iff it cannot be expressed as the union of disjoint open subsets. This is equivalent to saying that the only subsets that are both open and closed are S and $\varnothing$. Hence, $\mathbb{R}$ is connected, but the subset of all $x \in \mathbb{R}$ such that $|x| > 1$ is not. One the other hand, both $\mathbb{R}^n$ and the subset of all $x \in \mathbb{R}^n$ such that $d(x, 0) > 1$ will be connected when one uses the Euclidian metric.

In order to see that the image of a connected subset A under a continuous map $f : S \to S'$ is connected, suppose that $f(A)$ is disconnected; i.e., it is the union of disjoint open subsets U, V. The inverse images of U, V under f will also be disjoint open subsets of S whose union must be A, but that would imply that A would have to be disconnected, as well.

The fact that the image of a connected subset under a continuous map is continuous is the basis for the "intermediate-value theorem" in calculus, namely, that if $f : \mathbb{R} \to \mathbb{R}$ is continuous and $a < x < b$ then either $f(a) < f(x) < f(b)$ or $f(a) > f(x) > f(b)$, depending upon whether $f(a) < f(b)$ or $f(a) > f(b)$, resp. In other words, the image of the connected interval (a, b) will be either the connected interval $(f(a), f(b))$ or $(f(b), f(a))$.

Compactness of a topological space relates to how it can be covered by open subsets. An *open covering* of a topological space (S, τ) is a family of open subsets whose union contains S. Since that family might very well include an infinitude of open subsets, one says that S is *compact* iff every open covering of S contains a finite open subcovering; here, the term "subcovering" means a sub-family of open subsets in the first family such that its union also contains S. An important theorem of analysis is the *Heine-Borel* theorem that a subset of $\mathbb{R}^n$ (with the metric topology) is compact iff it is closed and bounded, in the sense that there is an n-ball of sufficiently large radius that contains it as a subset. Hence, the closed n-balls of finite radius are all compact, as are the closed n-cubes of finite edge dimension.

The way that one sees that the image of a compact space S under a continuous map f is compact is to define an arbitrary open covering of the image $f(S)$ and look at the corresponding covering of S by the inverse images of the open subsets of the covering of $f(S)$. If S is compact then that covering of S can be reduced to a finite open subcovering. One then selects the open subsets of the covering of $f(S)$ that correspond to the subsets of that finite subcovering and show that they also cover $f(S)$.

d. Topological manifolds. – A *topological manifold* is a topological space (M, τ) that is locally-homeomorphic to $\mathbb{R}^n$, with either its product or Euclidian metric topology. Hence, the points of M will always have neighborhoods that are topologically equivalent to $\mathbb{R}^n$.

Elementary examples of topological manifolds abound, such as open subsets of $\mathbb{R}^n$, open balls, n-dimensional spheres, projective spaces, tori, and such pathological examples as Möbius bands and Klein bottles. In fact, any open n-ball of finite radius in $\mathbb{R}^n$ is homeomorphic to $\mathbb{R}^n$, which shows that the Heine-Borel theorem really does depend upon the subset being closed in order for it to be compact, since an open n-ball of finite radius is bounded, but since it is homeomorphic to $\mathbb{R}^n$, which is not compact, an n-ball cannot be compact, either.

Two common ways of building up topological manifolds from simpler ones are by constructing product manifolds and quotient manifolds.

If M and N are both topological manifolds that are modeled on $\mathbb{R}^m$ and $\mathbb{R}^n$, resp., then one can first form the product space $M \times N$ as a topological space and then show that any point $(x, y) \in M \times N$ will have a neighborhood that is homeomorphic to $\mathbb{R}^m \times \mathbb{R}^n$. However, since x has a neighborhood U_x that is homeomorphic to $\mathbb{R}^m$, and y has a neighborhood V_y that is homeomorphic to $\mathbb{R}^n$, the product $U_x \times V_y$ will be homeomorphic to $\mathbb{R}^m \times \mathbb{R}^n$. This construction can be extended to a finite number of topological manifolds. Clearly, $\mathbb{R}^n$ itself is the product of n copies of the topological manifold $\mathbb{R}$. Similarly, an n-dimensional torus T^n will be the product of n copies of a circle S^1. If one removes a single point from $\mathbb{R}^n$ (such as the origin) then the resulting space will be homeomorphic to the product of $\mathbb{R}$ with the n–1-sphere S^{n-1}, and will thus be a topological manifold; this construction is the essence of polar coordinates.

A quotient manifold is somewhat harder to visualize. Basically, one defines an equivalence relation [.] on a topological manifold M, which defines a projection $[.] : M \to M / [.]$, $x \mapsto [x]$ from M to its quotient topological space, which consists of equivalence classes under the equivalence. One gives $M / [.]$ the quotient topology that was discussed above, and then examines the possibility that every point $[x]$ of $M / [.]$ will have a neighborhood that is homeomorphic to $\mathbb{R}^{n'}$, for some n' simply because every point x that projects onto $[x]$ will have a neighborhood that is homeomorphic to $\mathbb{R}^n$ for some n.

Common examples of quotient manifolds are the examples of quotient topological spaces that we gave above, such as spheres and tori, and projective spaces $\mathbb{R}P^n$. One can either define the latter spaces by the quotient map $\mathbb{R}^{n+1} - 0 \to \mathbb{R}P^n$, $\mathbf{v} \mapsto [\mathbf{v}]$, where $[\mathbf{v}]$ is the line through the origin that $\mathbf{v}$ belongs to, or the quotient map $S^n \to \mathbb{R}P^n$, $\mathbf{r} \mapsto [\mathbf{r}] = \{-\mathbf{r}, +\mathbf{r}\}$, which defines a point of $\mathbb{R}P^n$ to be a pair of antipodal points on the n-sphere. One sees that any line through the origin in $\mathbb{R}^{n+1}$ will intersect any n-sphere in $\mathbb{R}^{n+1}$ that is centered at the origin in precisely those two antipodal points.

A homeomorphism $x : U_p \subset M \to \mathbb{R}^n$, $x \mapsto (x^1(x), \ldots, x^n(x))$ can be called a *coordinate system* on the neighborhood U_p of the point p, and the individual functions $x^n(x)$ are then *coordinate functions*. When one has two such coordinate systems (U_p , x) and (V_p, y) on overlapping neighborhoods U_p , V_p of the point p, for any point p' in $U_p \cap V_p$, one can invert the restriction of the homeomorphism $x : U_p \to \mathbb{R}^n$, to $U_p \cap V_p$ and then compose that inverse with the restriction of $y : V_p \to \mathbb{R}^n$ to that same intersection and obtain a homeomorphism $y \cdot x^{-1} : \mathbb{R}^n \to \mathbb{R}^n$, $x^j \mapsto y^i(x^j)$ that is commonly called a *coordinate transformation*. In physics, one most commonly encounters such things in the form of systems of equations of the form:

$$y^i = y^i(x^j). \tag{1.5}$$

When a topological manifold M is not globally homeomorphic to $\mathbb{R}^n$, it cannot be covered by a single coordinate chart. However, quite often, it can be covered by a small number of charts. For instance, any sphere can be covered by two charts, which amount to overlapping open hemispheres. $\mathbb{R}P^n$ requires $n + 1$ overlapping charts, which are the images of the homogeneous coordinate charts $(x^0, \ldots, x^k, 1, x^{k+1}, \ldots, x^n)$ on $\mathbb{R}^{n+1} - 0$ for each $k = 0, \ldots, n$. One calls such charts *Plücker coordinates*.

The dimension of $\mathbb{R}^n$ that a topological manifold M is locally modeled on is called its *dimension*. All charts on M will have the same dimension n, and so will any topological manifold that is homeomorphic to M. All of this comes down to the idea that $\mathbb{R}^m$ is homeomorphic $\mathbb{R}^n$ to iff $m = n$.

2. Differentiable manifolds [3-5]. – Since physics is so deeply rooted in differential equations of various sorts, one will typically demand more of a coordinate transformation $y \cdot x^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ than mere invertibility and "bi-continuity." In particular, one generally expects that it will also be *bi-differentiable* to some level of differentiation; that is, that the map $y \cdot x^{-1}$ will also be k-times continuously-differentiable (if not smooth), along with its inverse map $x \cdot y^{-1}$. Such a map $f : \mathbb{R}^n \to \mathbb{R}^n$ that is invertible, k-times continuously-differentiable (i.e., C^k), and has a C^k inverse is called a C^k *diffeomorphism* of $\mathbb{R}^n$ with itself. In particular, when k is infinite, one calls it a *smooth diffeomorphism*. When two coordinate charts on overlapping neighborhoods of a point are related by a C^k (smooth, resp.) diffeomorphism of $\mathbb{R}^n$ with itself, one calls them C^k (smooth, resp.) coordinate charts. Notice that only pairs of charts can be described in that way, but not the individual charts, since it is the coordinate transformation that is C^k, not the coordinate functions.

a. Differential structures. – A *C^k (smooth, resp.) differentiable manifold* is defined to be a topological manifold M with a *C^k (smooth, resp.) differential structure*. A differentiable structure, in turn, is defined by a "maximal atlas of compatible C^k (smooth, resp.) coordinate charts." In most cases, it is simplest to consider a smooth atlas.

A *C^k atlas* of coordinate charts on M is a collection (typically infinite) of coordinate charts on M whose union covers M and whose coordinate transformations are all C^k. A coordinate chart (U, x^i) on M is *compatible* with that atlas if the coordinate transformations between (U, x^i) and all of the charts in the atlas that overlap (U, x^i) are C^k. The atlas of charts is called *maximal* (under subset inclusion) when any chart that is compatible with the atlas already belongs to it.

One finds that the obvious example of a differentiable manifold (which can, in fact, be given a smooth atlas) is $\mathbb{R}^n$ itself. Similarly, any open subset of $\mathbb{R}^n$ will define a differentiable manifold, and in fact, any open subset of a more general differentiable manifold. The examples of topological manifolds that were given above are also differentiable manifolds, such as spheres, tori, projective spaces. Examples of topological manifolds that are *not* differentiable manifolds are the various polytopes, such as cubes and tetrahedra.

b. Differentiable maps. – A *local representative* of a map $f : M \to N$ between differentiable manifolds (of some specified degree of differentiability, which will be tacitly assumed from now on) is a map $f_{xy} : \mathbb{R}^m \to \mathbb{R}^n$ such that there are coordinate charts (U, x^i) in M and (V, y^a) in N such that:

$$f = y^{-1} \cdot f_{xy} \cdot x\,. \tag{2.1}$$

That can be illustrated by a commutative diagram as in Fig. A.3:

$$\begin{tikzcd} M \arrow[r, "f"] & N \\ \cup & \cup \\ U \arrow[d, "x"'] & V \arrow[d, "y"] \\ \mathbb{R}^n \arrow[r, "f_{xy}"] & \mathbb{R}^n \end{tikzcd}$$

Figure A.3 – The local representative of a differentiable map.

f is said to be *differentiable* iff every local representative of f is differentiable. That has the effect of making the definition of differentiability independent of the choice of coordinate system.

The map f is called a *diffeomorphism* iff it is invertible and bi-differentiable. That will necessarily imply that M and N will be homeomorphic as topological manifolds, and will, *a fortiori*, have the same dimension. An open n-ball is not only homeomorphic to $\mathbb{R}^n$, but also diffeomorphic. However, as we pointed out before, although an n-cube is homeomorphic to an n-ball, it is not diffeomorphic, since the points along the vertices, edges, faces, etc., will not have consistent tangents at those points.

One can still define the differentiability of a map $f : S \to N$ from a subset S of M that is not open to a differentiable manifold N by taking advantage of the fact that differentiation is a "local" process, in the sense that if f and $f' : M \to N$ are differentiable maps that agree at a point $x \in M$ then their differential maps must agree at that point, as well. Hence, if one finds an open subset U that contains S and extends the map f to a differentiable map $\bar{f} : U \to N$ *in any manner* then one can say that its restriction to S is differentiable. This is commonly applied to the problem of defining a *differentiable singular k-simplex* $\sigma_k : I^k \to M$, in which $I^k \subset \mathbb{R}^k$ is a closed subset, namely, a closed k-cube.

Differential topology is typically concerned with the search for properties of differentiable manifolds that are common to diffeomorphic manifolds. Since all diffeomorphisms are also homeomorphisms, those properties will include the topological properties, such as connectedness and compactness.

Some of the common examples of differentiable maps that are not typically diffeomorphisms are differentiable curves in a differentiable manifold M, which are differentiable maps $x : \mathbb{R} \to M$, $t \mapsto x(t)$ and, more generally, *submanifolds* of M, which are differentiable maps $\sigma : S \to M$. The dimension of S does not have to be less than or equal to the dimension m of M, although often the most useful class of submanifolds are the *embedded* ones, which will be diffeomorphisms onto their images, so the dimension of S must be less than or equal to m, in that case.

c. Tangent and cotangent spaces. – One of the first things that one learns in the differential calculus of one real variable is that the derivative dy / dx of a differentiable function $y = y(x)$ at a point x_0 is the slope of the tangent line to the curve $(x, y(x))$ in $\mathbb{R}^2$ (which one calls the *graph* of y) at the point $(x_0, y(x_0))$. In multivariable calculus, the differential $dy = (\partial y / \partial x^1, \ldots, \partial y / \partial x^1)$ of a differentiable function $y(x^1, \ldots, x^n)$ on $\mathbb{R}^n$ at a point $x_0 = (x_0^1, \cdots, x_0^n)$ defines a tangent hyperplane to the graph of $y(x^1, \ldots, x^n, y(x^1, \ldots, x^n))$ in $\mathbb{R}^{n+1}$ at $(x_0, y(x_0))$.

One finds that the notion of a tangent space at a point can be adapted to the more general differentiable manifolds by taking advantage of the fact that if two differentiable curves $x(t)$ and $y(t)$ in M intersect at some point p then if they have the same tangent line in $\mathbb{R}^n$ at $x^i(p)$ with respect to some coordinate chart (U, x^i) about p, they will also have the same tangent line at $y^i(p)$ with respect to *any* coordinate chart (V, y^i) about p. Hence, the concept of tangency is independent of the choice of coordinate system.

One can then define a *tangent line* to a point $x \in M$ to be the set of all differentiable curves through x that have the same tangent line in $\mathbb{R}^n$ with respect to some, and therefore all, coordinate charts about x. In order to go from a tangent line to a curve to a tangent vector to a curve (namely, its velocity), one must specify a parameterization of the curve, since other parameterizations will give velocity vectors that all lie along the tangent line but have different lengths. Hence, a *tangent vector* $x \in M$ becomes the set of all differentiable curves through x that have the same velocity vector in some (and therefore all) coordinate charts about x.

Somewhat confusingly, it is traditional in differential topology and geometry to represent a tangent vector $\mathbf{X}$ at a point $x \in M$ by the directional derivative operator that acts upon differentiable functions that are defined in a neighborhood of x. Once again, one can do this by choosing a coordinate chart (U, x^i) about x and defining the directional derivative of a differentiable function f that is defined in some neighborhood of x in the direction of the tangent vector $\mathbf{X}$ to be:

$$\mathbf{X}f = X^i \frac{\partial f}{\partial x^i} \tag{2.2}$$

for that coordinate chart. When one changes to another chart (V, y^i) about x, so $y^i = y^i(x^j)$, one will have:

$$\mathbf{X}f = \bar{X}^i \frac{\partial f}{\partial y^i}, \tag{2.3}$$

in which $\bar{X}^i$ are the components of $\mathbf{X}$ in the new natural frame field on $\mathbb{R}^n$. From the chain rule of differentiation:

$$\frac{\partial f}{\partial y^i} = \frac{\partial x^j}{\partial y^i}\frac{\partial f}{\partial x^j}, \tag{2.4}$$

so one must set:

$$\bar{X}^i = \frac{\partial x^j}{\partial y^i} X^j \tag{2.5}$$

to be consistent.

The tangent vector $\mathbf{X}$ at a point $x \in M$, as a directional derivative, then becomes the linear, first-order differential operator on differentiable functions that are defined on neighborhoods of x whose local coordinate representatives are related by the contravariant transformation law (2.5). Hence, if α, β are scalars, and f, g are differentiable functions that are both defined on a neighborhood U_x of x then one must always have:

$$\mathbf{X}(\alpha f + \beta g) = \alpha \mathbf{X}f + \beta \mathbf{X}g, \qquad \mathbf{X}(f g) = (\mathbf{X}f)\, g + f\,(\mathbf{X}g). \tag{2.6}$$

The set T_xM of all tangent vectors at $x \in M$ is called the *tangent space to M at x*. It can be given the structure of an n-dimensional real vector space by saying that linear combinations $\alpha\mathbf{X} + \beta\mathbf{Y}$ of tangent vectors map to the corresponding linear combinations of vectors in $\mathbb{R}^n$ under some (hence, all) coordinate charts about x. Hence, T_xM is (by definition) linearly isomorphic to $\mathbb{R}^n$. One can specify that linear isomorphism by defining a *tangent n-frame* at x to be a set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of n linearly-independent tangent vectors $\mathbf{e}_i$ at x; i.e., a basis for T_xM. If one represents the tangent vector $\mathbf{X}$ in the form:

$$\mathbf{X} = X^i\, \mathbf{e}_i \tag{2.7}$$

then the linear isomorphism $T_xM \cong \mathbb{R}^n$ will take each such $\mathbf{X}$ to its component vector $(X^1, \ldots, X^n)$. Indeed, some people prefer to define tangent frames to be such isomorphisms, or more precisley, the inverse isomorphisms. Most commonly, one uses the *natural tangent frames*:

$$\mathbf{e}_i = \partial_i = \frac{\partial}{\partial x^i} \tag{2.8}$$

that are defined by the coordinate charts about x.

The dual space to T_xM consists of all linear functionals on the tangent vectors at x, and is referred to as the *cotangent space to M at x*; it is denoted by T_x^*M. Just as tangent vectors can be associated with directional derivatives of differentiable functions that are defined in a neighborhood of x, one can associate cotangent vectors with their differentials. Indeed, if $\mathbf{X} \in T_xM$ is a tangent vector at x then the differential df of a differentiable function f that is defined in a neighborhood of x will give:

$$df(\mathbf{X}) = \mathbf{X}f \tag{2.9}$$

when it is evaluated on $\mathbf{X}$.

The reciprocal coframe θ^i to a tangent frame $\mathbf{e}_j$ in T_xM will define a basis for T_x^*M so any tangent covector $\alpha \in T_x^*M$ can be expressed in the form:

$$\alpha = \alpha_i\, \theta^i. \tag{2.10}$$

Indeed, when α is evaluated on $\mathbf{X}$ the result will be:

$$\alpha(\mathbf{X}) = (\alpha_i\, \theta^i)\, (X^j\, \mathbf{e}_j) = \alpha_i\, X^j\, \theta^i\, (\mathbf{e}_j) = \alpha_i\, X^j\, \delta^i_j = \alpha_i\, X^i.$$

The tangent coframe that is reciprocal to the natural frame ∂_i is the *natural coframe* dx^i. Hence:

$$dx^i\, (\partial_j) = \delta^i_j\,. \tag{2.11}$$

A differentiable map $f : M \to N$ is associated with a linear map $df|_x : T_xM \to T_{f(x)}N$ at each $x \in M$ that one calls the *differential map to f at x*. As usual, it is defined in one (hence, all) pair of coordinate charts about x and $f(x)$. That is, if (U, x^i) is a coordinate chart about x and (V, y^a) is a coordinate chart about $f(x)$, such that one can express $f(x)$ in the coordinate form $y^a(x^i)$ then

$$df|_x = \left.\frac{\partial y^a}{\partial x^i}\right|_x \tag{2.12}$$

for those two coordinate systems.

For any other coordinate systems $(\bar{U}, \bar{x}^i)$ and $(\bar{V}, \bar{y}^a)$ about x and $f(x)$, respectively, the chain rule will give:

$$df|_x = \left.\frac{\partial \bar{y}^a}{\partial \bar{x}^i}\right|_x = \left.\frac{\partial \bar{y}^a}{\partial y^b}\right|_{f(x)} \left.\frac{\partial y^b}{\partial x^j}\right|_x \left.\frac{\partial x^j}{\partial \bar{x}^i}\right|_x . \tag{2.13}$$

Note that these three matrices are not all defined at the same point.

When f is a diffeomorphism, $df|_x$ will be invertible for all x. The Inverse Function Theorem gives a partial converse to that statement that if $df|_x$ is invertible at all x then there will be a neighborhood of each x on which f is a diffeomorphism. If $df|_x$ is a linear injection for all x then one will call f an *immersion*, while if it is a linear surjection (i.e., projection) then f will be a *submersion*. In the former case, the dimension of M must be less than that of N, while in the latter case, it must be greater than it.

A good bit of differential topology is taken up with seeing how much of the global behavior of a differentiable map can be predicted by information that is defined by the differential of that map at each point. Typically, one gets local information, at best, such as in the case of the Inverse Function Theorem.

One can take the disjoint union of all tangent spaces T_xM to all points $x \in M$ and arrive at a set $T(M)$ that one calls the *tangent bundle* to M. Similarly, the disjoint union of all cotangent spaces $T_x^* M$ over all x will give the *cotangent bundle* $T^*(M)$. One has canonical projections $T(M) \to M$, $\mathbf{X} \in T_xM \mapsto x$ and $T^*(M) \to M$, $\alpha \in T_x^* M \mapsto x$. These allow one to define topologies and differential structures on $T(M)$ and $T^*(M)$ that make them differentiable manifolds of dimension $2n$, but we shall not belabor those definitions, since we shall not actually be dealing with the general case very often in the main body of the present survey, because most of the quantum-mechanical references that are being reviewed never deal with the effects of going to space-time manifolds whose topologies are more pathological than Minkowski space.

We will say that when $f : M \to N$ is differentiable, there is a *tangent map* $Tf : T(M) \to T(N)$ that basically restricts to the differential map $df|_x : T_xM \to T_{f(x)}N$ at each x and a *cotangent map* $T^*f : T(N) \to T(M)$ that restricts to the "pull-back" map $f_x^* : T_{f(x)}^* N \to T_x^* M$, which takes any linear functional $\alpha \in T_{f(x)}^* N$ to the linear functional $f_x^* \alpha \in T_x^* M$ that will take any tangent vector $\mathbf{X}$ at x to the real number:

$$(f_x^* \alpha)(\mathbf{X}) = \alpha(df\big|_x(\mathbf{X})). \tag{2.14}$$

For the sake of consistency, the tangent vector $df\big|_x(\mathbf{X})$ is referred to as the "push-forward" of $\mathbf{X}$ by f, and is denoted by $f_*\mathbf{X}$. From (2.14), one can say that $f_x^* \alpha = \alpha \cdot f_*$, but that fact is largely incidental to the rest of this book.

d. Tensor fields on manifolds. – In order to go from tensor analysis on vector spaces, which we discussed in the previous Appendix, to tensor analysis on manifolds, one simply uses the tangent and cotangent spaces (viz., T_xM and $T_x^* M$) at each point x of a differentiable manifold M as the models for the vector spaces V and V^*, resp. The tensor product space:

$$T_l^k V = V^* \otimes \ldots \otimes V^* \otimes V \otimes \ldots, \otimes V \qquad (k \text{ copies of } V^*, \ l \text{ copies of } V)$$

will then get replaced with:

$$(T_l^k)_x M = T_x^* M \otimes \ldots \otimes T_x^* M \otimes T_xM \otimes \ldots \otimes T_xM.$$

The disjoint union of all $(T_l^k)_x M$ over all x then defines the *bundle of k-times covariant, l-times contravariant tensors on M*, which one can denote by $T_l^k M$. It, too, has a canonical projection $T_l^k M \to M$ that takes all tensors at x to x.

A *(k-times covariant, l-times contravariant) tensor field* on M is a "section" of the bundle $T_l^k M \to M$. Hence, it is a map $t : M \to T_l^k M$ that takes every $x \in M$ to a tensor $t(x)$ in $(T_l^k)_x M$. One can then locally convert any tensor field in this sense into a tensor field on a vector space by defining a coordinate chart (U, x^i) about x and using the natural frame field ∂_i and coframe field dx^i to obtain the local component functions of t with respect to that choice of local frame field:

$$t(x) = t^{j_1 \cdots j_l}_{i_1 \cdots i_k}(x)\, dx^{i_1} \otimes \ldots \otimes dx^{i_k} \otimes \partial_{j_1} \otimes \ldots \otimes \partial_{j_l} \,. \tag{2.15}$$

The degree of differentiability of the tensor field is then the same as the degree of differentiability of its component functions $t^{j_1 \cdots j_l}_{i_1 \cdots i_k}(x)$.

It is important to note that one does not need to define a coordinate chart in order to obtain component functions for a tensor field, but only a "local frame field." That is, one defines an open subset U and a frame field $\mathbf{e}_i(x)$ on U, which then associates each $x \in U$ with a tangent frame $\mathbf{e}_i(x)$ in T_xM. Although a coordinate chart will define a (natural) local frame field, a local frame field does not have to define a coordinate chart, since there are such things as anholonomic local frame fields.

In order to clarify that, it helps to define the notion of a (tangent) *vector field* on a differentiable manifold M. Such an object amounts to a section $\mathbf{X} : M \to T(M)$ of the projection $T(M) \to M$, which then takes each point x in M to a tangent vector $\mathbf{X}(x)$ in T_xM. Since we are representing tangent vectors by linear, first-order differential operators on differentiable functions on M – namely, directional derivatives – we can speak of composing the operators. Namely, if $\mathbf{X}(x)$ and $\mathbf{Y}(x)$ are vector fields on M then the composition of $\mathbf{Y}$ acting upon a function f, followed by $\mathbf{X}$ will be:

$$(\mathbf{XY})(x)\, f(x) = \mathbf{X}(x)\, [\mathbf{Y}(x)\, f(x)]. \tag{2.16}$$

Although $(\mathbf{XY})(x)$ will clearly be a linear second-order differential operator, and therefore, no longer a vector field, nonetheless, if one takes the difference:

$$[\mathbf{X}, \mathbf{Y}]\, f = (\mathbf{XY} - \mathbf{YX})\, f \tag{2.17}$$

then one will get a linear, first-order differential operator, which will be another vector field. That is due to the fact that if f is sufficiently differentiable then the mixed second partial derivatives will be symmetric in the coordinates and will therefore cancel.

One can see this most simply by looking at local component expressions for the vector field $\mathbf{X}$ and $\mathbf{Y}$ with respect to some local coordinate chart (U, x^i):

$$\mathbf{X}(x) = X^i(x)\, \partial_i\,, \qquad \mathbf{Y}(x) = Y^j(x)\, \partial_j\,. \tag{2.18}$$

That will make:

$$\begin{aligned}[\mathbf{X}, \mathbf{Y}] &= (X^i \partial_i)(Y^j \partial_j) - (Y^j \partial_j)(X^i \partial_i) \\ &= X^i (\partial_i Y^j)\, \partial_j + X^i Y^j \partial_i \partial_j - Y^j (\partial_j X^i)\, \partial_i - Y^j X^i \partial_j \partial_i\,, \\ &= X^i (\partial_i Y^j)\, \partial_j - Y^j (\partial_j X^i)\, \partial_i\,,\end{aligned}$$

which can be summarized by the formula:

$$[\mathbf{X}, \mathbf{Y}]^i = X^j \partial_j Y^j - Y^j \partial_j X^i\,. \tag{2.19}$$

The fact that [**X**, **Y**] is again a vector field eventually implies that the (infinite-dimensional) vector space $\mathfrak{X}(M)$ of all vector fields on M defines a "Lie algebra," which we shall discuss in more detail in the next Appendix.

Since a local frame field $\mathbf{e}_i\,(x)$ is composed of vector fields, one can take the Lie brackets of all pairs of them and produce local vector fields. Furthermore, since any tangent vector can be expressed in terms of that local frame field, one can express the Lie brackets in the form:

$$[\mathbf{e}_i, \mathbf{e}_j] = c_{ij}^k(x)\,\mathbf{e}_k\,. \tag{2.20}$$

When the $c_{ij}^k(x)$ (viz., the *structure functions* of the local frame field) all vanish, the frame field $\mathbf{e}_i(x)$ is called *holonomic*, and otherwise *anholonomic.* As we have mentioned before, natural frame fields are always holonomic, due to the symmetry of mixed second-order partial derivatives of sufficiently-differentiable functions with respect to coordinates.

e. Manifolds with other structures. – In order to go from differential topology to differential geometry – i.e., the geometry of curved space – one usually has to impose further structures upon the differentiable manifold that is one considers. Often those structures are defined by tensor fields on the manifold, and very often they are differential forms, in particular.

The most common extra structure that one imposes upon a differentiable manifold is a *metric tensor field* (which does not necessarily define a metric in the point-set topological sense). That would take the form of a symmetric, doubly-covariant tensor field $g\,(x)$ on M that is nondegenerate, moreover. Basically, that amounts to saying that its component matrix $g_{ij}\,(x)$ with respect to any local frame field $\mathbf{e}_i\,(x)$:

$$g = g_{ij}\,\theta^i\,\theta^j \tag{2.21}$$

will be invertible. Here, θ^i is the reciprocal coframe field to $\mathbf{e}_i$, and we have suppressed the explicit mention of the symmetrized tensor product symbol in $\theta^i\,\theta^j$.

What one gets from g is a scalar product on each tangent space T_xM. Namely, if $\mathbf{v}$ and $\mathbf{w}$ are tangent vectors in T_xM then their scalar product will be:

$$<\mathbf{v}, \mathbf{w}> = g\,(x)(\mathbf{v}, \mathbf{w}) = g_{ij}\,v^i\,w^i. \tag{2.22}$$

The local frame field $\mathbf{e}_i\,(x)$ will then be called *orthonormal* iff one has:

$$<\mathbf{e}_i, \mathbf{e}_j> = \mathrm{diag}\,[+\,1\ldots +\,1 - 1\ldots -\,1]. \tag{2.23}$$

When the diagonal elements are all positive, the scalar product will be the Euclidian one, and the metric will be called *Riemannian*. Otherwise, it will be *pseudo-Riemannian*, and the most important case for the theory of relativity is the *Lorentzian* case, in which one has either one positive and $n - 1$ negatives or the opposite combination, and the scalar product will then make each tangent space look like Minkowski space.

Since many of the other fundamental tensor fields that we shall encounter, such as a volume element, take the form of exterior differential forms on manifolds, we shall move on to that topic.

3. Differential forms on manifolds [3-6]. – Just as we could discuss exterior differential forms on vector spaces as special cases of tensor fields on vector spaces, we can carry out an analogous discussion of exterior differential forms on differentiable manifolds. That is because a k-vector field **A** on a differentiable manifold M is a completely-antisymmetric, completely-contravariant tensor field on M, while an exterior differential k-form is a completely-antisymmetric, completely covariant tensor field on M.

We shall denote the bundle of all k-vectors on M by $\Lambda_k M$ and the bundle of all k-forms by $\Lambda^k M$. The fiber of the former bundle at $x \in M$ will be $A_k(T_x M) = T_x M \wedge \ldots \wedge T_x M$ (k copies), while the fiber of the latter bundle at x will be $A^k(T_x M) = A_k(T_x^* M) = T_x^* M \wedge \ldots \wedge T_x^* M$. Hence, a k-vector field $\mathbf{A}(x)$ on M will be a section $\mathbf{A}: M \to \Lambda_k M$ of the projection $\Lambda_k M \to M$, which will then take each x to a k-vector in $A_k(T_x M)$ and a k-form $\alpha(x)$ will be a section $\alpha: M \to \Lambda^k M$ of the projection $\Lambda^k M \to M$ that takes x to a k-form in $A^k(T_x M)$.

When one has a local frame field $\mathbf{e}_i(x)$ on an open subset U in M, and its reciprocal coframe field θ^i, one can express both **A** and α in terms of local components:

$$\mathbf{A}(x) = \frac{1}{k!} A^{i_1 \cdots i_k}(x)\, \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_k}, \qquad \alpha(x) = \frac{1}{k!} \alpha_{i_1 \cdots i_k}(x)\, \theta^{i_1} \wedge \cdots \wedge \theta^{i_k}. \tag{3.1}$$

The exterior products of multi-vector fields and exterior forms can still be defined without making recourse to local frame fields since the tensor product of sections of the tensor bundles in question simply comes down to taking tensor products in the fibers, which does not depend upon components for its definition. If **A**, **B** are a k-vector field and a l-vector field, resp., while α, β are a k-form and an l-form resp., then one will still have:

$$\mathbf{A} \wedge \mathbf{B} = (-1)^{kl}\, \mathbf{B} \wedge \mathbf{A}, \qquad \alpha \wedge \beta = (-1)^{kl}\, \beta \wedge \alpha. \tag{3.2}$$

One can also define the interior product of a k-form by an l-vector field ($l < k$) or a k-vector field by an l-form, as before, by starting with the interior product of a decomposable k-form $\alpha = \alpha_1 \wedge \ldots \wedge \alpha_k$ by a vector field **v**:

$$\begin{aligned} i_{\mathbf{v}}\alpha &= i_{\mathbf{v}}(\alpha_1 \wedge \ldots \wedge \alpha_k) \\ &= \alpha_1(\mathbf{v})\, \alpha_2 \wedge \ldots \wedge \alpha_k - \alpha_2(\mathbf{v})\, \alpha_1 \wedge \alpha_3 \ldots \wedge \alpha_k - (-1)^k\, \alpha_k(\mathbf{v})\, \alpha_1 \wedge \ldots \wedge \alpha_{k-1}, \end{aligned}$$

extending this to indecomposable ones by linearity, and extending to l-vector fields by starting with decomposable ones:

$$i_{\mathbf{v}_1 \cdots \mathbf{v}_l}\alpha = i_{\mathbf{v}_l} \cdots i_{\mathbf{v}_1}\alpha, \tag{3.3}$$

and extending to indecomposable ones by linearity.

a. The exterior derivative. – The exterior derivative operator $d_\wedge : \Lambda^k M \to \Lambda^{k+1} M$ can still be defined uniquely by the requirements that:

1. $d_\wedge$ is linear.
2. $d_\wedge f = df$ when f is a 0-form (i.e., differentiable function).
3. It is an anti-derivation; hence, if α is a k-form and β is an l-form:

$$d_\wedge (\alpha \wedge \beta) = d_\wedge \alpha \wedge \beta + (-1)^k \alpha \wedge d_\wedge \beta. \tag{3.4}$$

4. Its square is zero:

$$d_\wedge d_\wedge = 0. \tag{3.5}$$

"Closed" and "exact" still mean the same thing for differential forms on manifolds. It is when one looks at the Poincaré lemma that one notices that topology now asserts itself. For a topologically-general differentiable manifold, one can say only that every closed form is *locally* exact, not that it is globally exact. In fact, the existence of k-forms that are closed, but not exact is the starting point for "de Rham cohomology," although we shall have no need for that here.

Another time that topology interferes with the basic constructions is when one needs a volume element on M. That would still amount to a global, non-zero n-form, but whether such things can even exist is also a matter of topology, namely, the orientability of $T(M)$.

b. Integration of differential forms. – As for the integration of exterior differential forms on a differentiable manifold M, one can best generalize the definitions that were given vector spaces by first generalizing the concept of a differentiable singular cubic n-simplex in $\mathbb{R}^n$ to a *differentiable singular cubic n-simplex* in an n-dimensional manifold M. That would simply be a differentiable map $\sigma_n : I^n \to M$, now. More specifically, in order for the map $\sigma_k(a)$ to preserve the dimension of I^k, one should also assume that it is an embedding. A *differentiable singular cubic n-chain in M* would then be a formal sum $c_n = \sum_a \lambda_a \sigma_n(a)$ of a finite number of differentiable singular cubic n-simplexes $\sigma_k(a)$.

If an n-form $\alpha = \alpha(x^1, \ldots, x^n)\, dx^1 \wedge \ldots \wedge dx^n$ is defined on the image of such an n-simplex σ_n then one can pull back α to give a k-form $\sigma_k^* \alpha$ on I^k, that one might represent in the form:

$$\sigma_k^* \alpha = \alpha(\bar{x}^1(x), \ldots, \bar{x}^1(x)) \det[\partial_j \bar{x}^j]\, dx^1 \wedge \cdots \wedge dx^n. \tag{3.6}$$

The integral:

$$\int_{\sigma_k} \alpha \equiv \int_{I^k} \sigma_k^* \alpha = \int_{I^k} \alpha(\bar{x}^i(x), \cdots, \bar{x}^k(x)) \det[\partial_j \bar{x}^i]\, dx^1 \wedge \cdots dx^k \tag{3.7}$$

can the be evaluated in the usual manner of multivariable calculus.

The extension to an integral over an n-chain c_n is then by linearity:

$$\int_{c_n} \alpha \equiv \sum_a \lambda_a \int_{\sigma_n(a)} \alpha . \tag{3.8}$$

Stokes's theorem basically generalizes to the statement that if an n–1-form α is defined on an n-chain c_k with a boundary $\partial c_n = \sum_a \lambda_a \, \partial\sigma_n(a)$ then:

$$\int_{\partial c_n} \alpha = \int_{c_n} d_\wedge \alpha . \tag{3.9}$$

In order to get from (3.9) to Gauss's theorem (i.e., the divergence theorem), one must introduce a volume element (hence, orientability) in order to take define Poincaré duality. One first expresses $\alpha = \#\mathbf{A}$ for some (unique) vector field $\mathbf{A}$. (3.9) will then take the form:

$$\int_{\partial c_n} \#\mathbf{A} = \int_S d_\wedge \#\mathbf{A} = \int_S \#(\operatorname{div}\mathbf{A}) . \tag{3.10}$$

More commonly, one sees the use of Hodge star isomorphism, which will imply both a volume element and a metric. Basically, one expresses α as $*\beta$ for some (unique) n–k+1-form β. (3.9) will then become:

$$\int_{\partial c_n} *\beta = \int_{c_n} d_\wedge * \beta = \pm \int_{c_n} *\delta\beta , \tag{3.11}$$

in which the sign will depend upon n and the signature type of the metric, as discussed in the previous Appendix. Of course, as was pointed out at that time, the introduction of a metric is not a natural construction for the divergence operator.

4. Lie derivatives on manifolds. – The concept of the Lie derivative $L_{\mathbf{X}}t$ of a tensor field t on a differentiable manifold M with respect to a vector field $\mathbf{X}$ is a straightforward adaptation of the previous discussion on vector spaces. One first defines the Lie derivative of a differentiable function to agree with the directional derivative:

$$L_{\mathbf{X}} f = \mathbf{X} f. \tag{4.1}$$

One then defines the Lie derivative of a vector field $\mathbf{Y}$ to be:

$$L_{\mathbf{X}} \mathbf{Y} = [\mathbf{X}, \mathbf{Y}] \tag{4.2}$$

and the Lie derivative of a covector field α to make:

$$(L_{\mathbf{X}}\, \alpha)(\mathbf{Y}) = \mathbf{X}\,(\alpha(\mathbf{Y})) - \alpha([\mathbf{X}, \mathbf{Y}]) \tag{4.3}$$

for any vector field $\mathbf{Y}$, which agrees with Cartan's magic formula for any k-form α:

$$L_{\mathbf{X}}\, \alpha = i_{\mathbf{X}}\, d_\wedge \alpha + d_\wedge\, i_{\mathbf{X}}\, \alpha \tag{4.4}$$

when $k = 1$. It also helps to use the "intrinsic" formula for the exterior derivative of a 1-form α:

$$d_\wedge\alpha(\mathbf{X}, \mathbf{Y}) = \mathbf{X}(\alpha(\mathbf{Y})) - \mathbf{Y}(\alpha(\mathbf{X})) - \alpha([\mathbf{X}, \mathbf{Y}]). \tag{4.5}$$

One then extends to any tensor field of arbitrary rank by demanding that $L_\mathbf{X}$ must be a linear derivation, so if t takes the decomposable form:

$$t = \alpha^1 \otimes \ldots \otimes \alpha^k \otimes \mathbf{Y}_1 \otimes \ldots \otimes \mathbf{Y}_l \tag{4.6}$$

then:

$$L_\mathbf{X}\, t =$$
$$L_\mathbf{X}\alpha^1 \otimes \ldots \otimes \alpha^k \otimes \mathbf{Y}_1 \otimes \ldots \otimes \mathbf{Y}_l + \ldots + \alpha^1 \otimes \ldots \otimes L_\mathbf{X}\,\alpha^k \otimes \mathbf{Y}_1 \otimes \ldots \otimes \mathbf{Y}_l +$$
$$\alpha^1 \otimes \ldots \otimes \alpha^k \otimes L_\mathbf{X}\mathbf{Y}_1 \otimes \ldots \otimes \mathbf{Y}_l + \ldots + \alpha^1 \otimes \ldots \otimes \alpha^k \otimes \mathbf{Y}_1 \otimes \ldots \otimes L_\mathbf{X}\mathbf{Y}_l .$$

One then extends to indecomposable tensor fields by linearity.

The two example of Lie derivatives that one encounters most frequently in continuum mechanics are Lie derivatives of the metric tensor field g and Lie derivatives of the volume element V.

In the first example, one gets the infinitesimal rate of strain in the flow of the velocity vector field $\mathbf{v}$:

$$\dot{e}_{ij} = L_\mathbf{v}\, g_{ij} = \partial_i v_j + \partial_i v_j . \tag{4.7}$$

When this vanishes, the vector field $\mathbf{v}$ is called a *Killing vector field*, and its flow will consist of isometries of the metric g.

In the second example, which we also discussed in the previous Appendix, one will get the (kinematical) compressibility of the flow of $\mathbf{v}$:

$$\lambda = L_\mathbf{v} V = i_\mathbf{v} d_\wedge V + d_\wedge\, i_\mathbf{v} V = d_\wedge \#\mathbf{v} = \#(\mathrm{div}\,\mathbf{v}) = (\mathrm{div}\,\mathbf{v})\, V. \tag{4.8}$$

When λ vanishes, the flow of $\mathbf{v}$ will be called *kinematically incompressible*. One then sees that this condition is equivalent to the vanishing of div $\mathbf{v}$, and the flow of $\mathbf{v}$ will consist of volume-preserving diffeomorphisms.

References

1. F. Haussdorff, *Mengenlehre*, 2nd ed., Walter De Gruyter and Co., Berlin, 1927; English translation by J. R. Aumann, *et al.*, *Set Theory*, Chelsea Pub. Co., NY, 1967.
2. J. Munkres, *Topology*, Prentice Hall, NJ, 2000.
3. F. Warner, *Differentiable Manifolds and Lie Groups,* Scott Foresman, Glenview, IL, 1971.
4. J. Milnor, *Topology from the Differentiable Viewpoint*, U. of Virginia Press, Charlottesville, 1965.
5. T. Frenkel, *The Geometry of Physics: an introduction*, Cambridge University Press, Cambridge, 1997.

6. H. Cartan, *Differential Forms*, Dover, NY, 2006.

APPENDIX C

Lie groups, Lie algebras, and their representations

Before we get into the business of doing the calculus of variations with physical fields, we must first introduce some basic notions from the theory of Lie groups. More to the point, it is the infinitesimal action of Lie groups on the points of space-time and the vectors in field space that enters into the discussion of variations of fields. An especially important class of group actions is defined by linear group actions on vector spaces, which are also known as "representations" of Lie groups.

The basic objective of this chapter is to attempt to distill out the definitions, theorems, and formulas from the vast body of literature on the subjects of Lie groups, Lie algebras, and their representations as they have been treated by both the pure mathematics community and the mainstream physics community. Indeed, there continues to exist a sizable gulf between the way that the topics are approached by the two communities, and that stems from the fact that the topics have so many applications to either domain of problems that any author on the subject will inevitably have some specific constellation of problems in mind. In the present case, that constellation would center around the problems of applying the calculus of variations to the various wave equations of quantum mechanics and their continuum-mechanical interpretations.

Some common references on the subject of this Appendix are [**1-5**]. In particular, the last one is oriented towards physics applications, while the first four are purely mathematical.

1. Basic concepts regarding Lie groups. – Only the most elementary notions from the theory of differentiable manifolds will be assumed in this chapter, since eventually the only manifolds that will be considered in this book will take the form of open subsets of real vector spaces. In particular, the topics that were covered in the previous Appendix should be more than sufficient for an understanding of this Appendix.

a. Basic definitions. – A *Lie group* is a set G that that has been given two structures that are compatible with each other, namely, a group structure and the structure of a differentiable manifold.

That is, G has a binary operation $G \times G \to G$, $(g, g') \mapsto gg'$ defined on it that is associative, possesses an identity element e, and every element g has a unique inverse element g^{-1}. Hence, for every $g, g_1, g_2, g_3 \in G$:

1. $(g_1 g_2) g_3 = g_1 (g_2 g_3)$.
2. $eg = ge = g$.
3. $gg^{-1} = g^{-1}g = e$.

(Sometimes, the axiom of "closure" is included with these – namely, that gg' always belongs to G – but we have included that in our definition of binary operation.)

Secondly, G has a topology defined on it, along with a maximal atlas of charts (U, x^i) that always make each open subset U homeomorphic to $\mathbb{R}^n$ in such a way that the coordinate changes will always be "sufficiently-differentiable" (typically, smooth) on the intersections of the open subsets. The *dimension* of a Lie group will then be its dimension as a differentiable manifold.

The compatibility requirement amounts to demanding that the group operations themselves must be differentiable. That is, the binary operation $G \times G \to G$ is a differentiable map, along with the inversion map $G \to G$, $g \mapsto g^{-1}$. As it turns out (see Chevalley [**1**]), one can always reduce a C^k atlas of charts on any Lie group to an analytic atlas so the group operations will become analytic, as well; this is especially useful when one is dealing with matrix exponentials.

One then extends the usual sort of group notions by adding a differentiability requirement to the definition. For instance, a *Lie subgroup* of G is a subgroup H of G such that the inclusion map $H \subset G$, $h \mapsto h$ is differentiable. A *Lie group homomorphism* is a group homomorphism $h : G \to G'$ that is also differentiable, and similarly for a *Lie group isomorphism*. Actually, it is sufficient to require continuity of the map, since there is a theorem that a continuous homomorphism of Lie groups must be analytic.

b. Examples of Lie groups. – Examples of Lie groups abound, so we shall concentrate on the ones that will be of interest to us in the cause of physical field theory, which will usually be groups whose elements act as transformations on the points of space-time or the vectors in some field space.

Perhaps the simplest Lie group that does not have dimension zero (which would be typical of discrete groups, such as $\mathbb{Z}$ or $\mathbb{Z}_n$) is the additive group $(\mathbb{R}, +)$ of real numbers. In fact, that group is isomorphic to the multiplicative group $(\mathbb{R}^*, \times)$ of positive real numbers by the exponential map. One can extend the dimension to the translation group $(\mathbb{R}^n, +)$ in the former case, although defining a multiplicative structure on $\mathbb{R}^n$ (i.e., an *algebra*, when the binary operation is bilinear) is not as simple as it is in the case of $(\mathbb{R}^*, \times)$.

Analogously, one can define the equivalence relation on points of the real line that two points $x, y \in \mathbb{R}$ are equivalent iff $y - x$ is an integer (or perhaps an integer multiple of 2π). As a topological space, the image T^1 of the projection $\mathbb{R} \to \mathbb{R} / \mathbb{Z}$, $x \mapsto [x]$ is homeomorphic to a circle S^1, and in fact, the addition on $\mathbb{R}$ projects to addition modulo $\mathbb{Z}$ on T^1 to define a compact, Abelian, one-dimensional Lie group that we shall just call the *circle group*. It is also isomorphic as a Lie group to the special orthogonal group $SO(2)$ for two-dimensional real Euclidian space and the unitary group $U(1)$ that is represented by the unit circle in the complex plane. One can extend the definition of T^1 to n dimensions by defining the *n-torus* $T^n = T^1 \times \ldots \times T^1$ to be the compact, Abelian, n-dimensional Lie group that is equivalently the image of the projection $\mathbb{R}^n \to \mathbb{R}^n / \mathbb{Z}^n$ or the product of n copies of the Lie group T^1.

Most of the Lie groups that we shall deal with in this book will be examples of *matrix Lie groups*. That is, they are defined by sets of invertible real or complex matrices for which the binary operation is matrix multiplication. Hence, they will be subgroups of the Lie groups $GL(n; \mathbb{R})$ or $GL(n; \mathbb{C})$, which are defined by all invertible $n \times n$ real or complex matrices, respectively. The identity element in either case is the $n \times n$ identity matrix I. The real (complex, resp.) dimension of those Lie groups is n^2. $GL(n; \mathbb{R})$ consists of two connected components, one of which contains the subgroup $GL^*(n; \mathbb{R})$ of all real $n \times n$ matrices with positive determinants, and the other of which is not a group (it has no identity element), although it is diffeomorphic to $GL^*(n; \mathbb{R})$ as a manifold; both manifolds still have dimension n^2. $GL(n; \mathbb{C})$, by contrast, is connected, since one can get from positive numbers to negative numbers without passing through zero when one considers curves in the complex plane, while removing the origin will disconnect $\mathbb{R}$.

In order to obtain most matrix Lie groups, one defines some set of algebraic conditions (which also makes them *algebraic groups*); that will usually reduce the dimension in the process. For instance, $SL(n; \mathbb{K})$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) is defined by the condition that the determinant must always be unity, which reduces the dimension by one. If the $\mathbb{K}$-vector space $\mathbb{K}^n$ has a (real or complex) orthogonal structure (i.e., a scalar product) defined on it then one can impose the restriction that the matrices must preserve that orthogonal structure and obtain the group $O(p, q; \mathbb{K})$ where the signature type of the scalar product is $(-1, \ldots, -1, +1, \ldots, +1)$ with p negative signs and q positive ones. For instance, $O(n)$ will default to the real orthogonal group of n-dimensional Euclidian space, so the symbol $\mathbb{K} = \mathbb{R}$ will be implicit; otherwise, we will write $O(n; \mathbb{C})$. If the matrix of the scalar product is η then the condition that a matrix A must be orthogonal amounts to:

$$A^{\mathrm{T}} \eta A = \eta. \tag{1.1}$$

This allows one to define an η-adjoint to A:

$$A^* = \eta A^{\mathrm{T}} \eta, \tag{1.2}$$

which makes:

$$A^{-1} = A^*. \tag{1.3}$$

In particular, for a Euclidian orthogonal matrix, $A^{-1} = A^{\mathrm{T}}$.

The $\mathbb{K}$-dimension of any orthogonal group for $\mathbb{K}^n$ is the same for all signature types. It will be $n (n - 1) / 2$, but that fact is easier to see when one considers the infinitesimal generators of one-parameter subgroups of orthogonal transformations, which we shall do later. Actually, it is unnecessary to specify the signature type of complex orthogonal spaces, since the fact that they take on complex values implies that one cannot distinguish negative subspaces from positive ones.

If one combines the orthogonality constraint with the constraint that the determinant of A must also be unity then one will arrive at the Lie group $SO(p, q; \mathbb{K})$, or *special orthogonal group for signature type* (p, q). In particular, $SO(3, 1)$ is the special Lorentz group and $SO(3)$ is the Lie group of orientation-preserving Euclidian rotations in three dimensions. The group $SO(3; \mathbb{C})$ is actually more physically interesting than one might first expect, since it is closely related to $SO(3, 1)$, although that is easier to show at the infinitesimal level, for which one gets an isomorphism of Lie algebras.

When $\mathbb{K} = \mathbb{C}$, and one gives $\mathbb{C}^n$ a Hermitian inner product, one can restrict $n\times n$ complex matrices to the ones that preserve that inner product and obtain the Lie group $U(n)$ of all *unitary* $n\times n$ matrices. For them:

$$A^{-1} = A^{\dagger} \equiv A^{\mathrm{T}*}, \tag{1.4}$$

in which $A^{\dagger}$ is referred to as the *Hermitian transpose of A*.

When one adds the constraint that the determinant must be unity, one will define the subgroup $SU(n)$ of *special unitary matrices* in dimension n. Although the fact that the elements of $SU(2)$ are complex 2×2 matrices might suggest that is it a complex manifold, actually, it is a real one. For one thing, its real dimension is three, which is an odd number, so it cannot possible admit a complex structure (i.e., an atlas of charts in $\mathbb{C}^m$ for some m with holomorphic coordinate changes). In fact, there is a 2-1 Lie group homomorphism $SU(2) \to SO(3)$ that is topologically the same thing as the map from any point on the three-sphere S^3 to the line that goes through it and the center, and that line will define a point in the real projective space $\mathbb{R}\mathrm{P}^3$. Similarly, there is a 2-1 Lie group homomorphism $SL(2; \mathbb{C}) \to SO_0(3, 1)$ that is topologically the complexification of that picture. Here the subscript "0" suggests that the special Lorentz transformations also preserve a "time orientation," which makes $SO_0(3, 1)$ the connected component of the identity in $O(3, 1)$; one calls it the *proper, orthochronous Lorentz group.* We shall return to the last two homomorphisms in the context of spin representations of Lie groups.

2. Fields as representations. – One of the fundamental sources of confusion between mathematicians and physicists in the name of field theory is their inconsistent use of the word "representation." We shall attempt to clarify the differing usages that mathematicians and physicists typically make.

a. Elementary representations of Lie groups. – To the mathematician, a *representation* of a group G (such as a Lie group, in particular) is a homomorphism $D : G \to GL(V)$, $g \mapsto D(g)$ that takes every element g in the group G to a corresponding element $D(g)$ in the group $GL(V)$ of invertible linear transformations of some vector space V. Hence, as a homomorphism, one must have:

$$D(gg') = D(g)\, D(g'),$$

in which the two group multiplications are both represented by the concatenation of symbols.

Any group homomorphism must take the identity element e in G to the identity transformation I in $GL(V)$, and the inverse image of I under D is always a subgroup $\ker(D)$ of G that one calls the *kernel* of D. Moreover, the map that takes all elements of G to I is always a homomorphism, although a trivial one. At the other extreme, $\ker(D) = e$ iff the map D is injective (i.e., one-to-one). In such a case, one calls the representation D *faithful*. In the case of Lie groups, for which the concept of dimension is well-defined, this can happen only when the dimension of G is less than or equal to that of $GL(V)$.

Similarly, the image $D(G)$ of G under D is always a subgroup of $GL(V)$. If D is surjective (i.e., onto) then the image of G under D will be all of $GL(V)$; for a Lie group, that will be possible only if the dimension of G is greater than or equal to that of $GL(V)$.

The homomorphism D will be both one-to-one and onto (or *bijective*) iff it is invertible, and in that case, one would call it an *isomorphism*; similarly, one would say that G is *isomorphic* to $GL(V)$ and notate that by $G \cong GL(V)$. An important isomorphism to know about is defined by the fact that $\ker(D)$ is always normal in G ([1]), so the coset space $G / \ker(D)$ will have a group structure, and $D(G) \cong G / \ker(D)$. As consequence, $D(G) \cong G$ iff $\ker(D) = e$ iff D is injective.

For example, the covering homomorphism $SU(2) \to SO(3)$ has $\mathbb{Z}_2 = \{I, -I\}$ for its kernel, and since it is surjective, $SU(2) / \{I, -I\} \cong SO(3)$. Hence, one can represent a three-dimensional Euclidian rotation as a pair of unitary 2×2 complex matrices with unity determinant that are each the negative of the other one.

There are more interesting examples of representations than the trivial representation of G in $GL(V)$. For instance, the determinant map $\det : GL(V) \to \mathbb{R}^*$, $A \mapsto \det(A)$, where $\mathbb{R}^*$ is the multiplicative group of non-zero real numbers, is a non-faithful representation of $GL(V)$ in the group of invertible linear transformations of the real line $\mathbb{R}$. The fact that it is a homomorphism follows from the product rule for determinants:

$$\det(AB) = \det(A)\, \det(B).$$

Any permutation group $S!$ of a set $S = \{O_1, \ldots, O_n\}$ of n objects can be represented in $GL(n; \mathbb{K})$, where is $\mathbb{K}$ is any suitable scalar field, such as $\mathbb{R}$ or $\mathbb{C}$. One simply associates each object O_i in S with the frame member $\mathbf{e}_i$ for some frame $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ in $\mathbb{K}^n$. Every permutation $\pi : S \to S$ (i.e., every invertible set map) will then correspond to a permutation of the frame members. The matrix of that permutation with respect to the initial frame will then be a corresponding permutation of the columns of the identity

([1]) This is easy to prove: If $k \in \ker D$ and $g \in G$ is arbitrary then $D(gkg^{-1}) = D(g)D(k)D(g^{-1}) = D(g)D(g^{-1}) = I$, so $gkg^{-1} \in \ker D$.

matrix for n dimensions. [If $GL(n; \mathbb{K})$ acts upon the dual space $\mathbb{K}^{n*}$ then the permutation will affect rows of the identity matrix.]

When one is dealing with linear transformations of V to begin with, the group G might be defined to be a subgroup of $GL(V)$, and one will then refer to the inclusion map $G \subset GL(V)$, $g \mapsto g$ as the *defining representation*. In particular, one can use that term for $GL(V)$ itself, as well as the subgroups of invertible linear transformations of its linear subspaces. If an orthogonal structure has been introduced then one can define the group of orthogonal transformations (e.g., rotations) for it, and if V is a complex vector space with a Hermitian inner product then one can deal with the defining representation of its unitary transformations.

An important consideration for any representation $D : G \to GL(V)$ is its reducibility. That is: Can one find a non-trivial representation of D in a subgroup $GL(V')$, where V' is a proper linear subspace of V? If so, one calls V' an *invariant subspace* for the representation. If no such proper subspace exists then one calls the representation D *irreducible*. When a representation is reducible, there is generally a decomposition of V into a direct sum $V_1 \oplus \ldots \oplus V_r$ of proper linear subspaces such that the image of D then becomes the direct product $GL(V_1) \times \ldots \times GL(V_r)$. The matrix $D(g)$ that represents the group element g will then decompose into block-diagonal form accordingly.

b. Weights and spins of representations. – Whenever one is dealing with a linear transformation T of a vector space to itself, one can always define the eigenvalues of T, although they might not belong to the given field of scalars; in the complex case, one also always define the eigenvectors of T, as well. Since one can say much about the structure of the transformation T when one knows all of its eigenvalues and eigenvectors, it is not surprising that the eigenvalues and eigenvectors of the linear transformations $D(g)$ of V that represent elements g of a group G will also play a fundamental roles, especially in the classification of irreducible representations of G. Indeed, to Nineteenth-Century algebraists, the equation:

$$\det[D(g) - \lambda I] = 0 \tag{2.1}$$

was usually referred to as the *fundamental equation* of the representation. (Usually the representation in question was that of left-translation, so one could extend the definition to algebras in the general sense.)

As one knows, if the eigenvectors actually exist in V then V can be decomposed into a direct sum $V(\lambda_1) \oplus \ldots \oplus V(\lambda_d)$ of subspaces that each correspond to a separate eigenvalue, and whose dimensions are equal to the multiplicities of the eigenvalues as roots of the characteristic equation. It is convenient for many purposes to have multiplicity one for each root, so the eigenspaces will each have dimension one; that will also make the matrix $D(g)$ *diagonalizable*, and it will have distinct diagonal elements. However, the existence of distinct roots of the characteristic polynomial is not necessary for the diagonalizability of the matrix, but only the existence of distinct roots for the "minimal polynomial." For instance, the identity matrix is clearly diagonalizable, although its characteristic polynomial is $(x - 1)^n$; its minimal polynomial, however, is $x - 1$.

One also knows that all elements $h' \in G$ that are conjugate to a given h will be represented by linear transformations with the same eigenvalues, since:

$$\begin{aligned}\det[D(ghg^{-1}) - \lambda I] &= \det[D(g)D(h)D(g^{-1}) - \lambda D(g)\, D(g^{-1})] \\ &= \det\{D(g)[D(h) - \lambda I]\, D(g^{-1})\} \\ &= \det[D(h) - \lambda I].\end{aligned}$$

When a Lie group G is compact, one can always find maximal compact Abelian subgroups, which one then calls *maximal torii* ([1]) (cf., e.g., Adams [**2**]). The elements $h \in T^k$, where T^k is a maximal torus, are especially fundamental, because the irreducible representations of G are in one-to-one correspondence with the irreducible representations of T^k, and the irreducible representations of $T^1 = S^1$ are all one-complex-dimensional (i.e., two-real dimensional). The eigenvalues of $D(h)$ in that case are called the *weights* of the representation D, and one can then decompose the vector space V of the representation D into eigenspaces of the weights. That is why complex vector spaces are the ones that one commonly uses for discussions of irreducible representations. Furthermore, since the elements of T^k commute with each other, so will the matrices $D(h)$ that represent them, and that will make the matrices simultaneously diagonalizable; i.e., there will be a frame on V for which all of the matrices $D(h)$ are diagonal.

When a Lie group G is not compact, maximal torii do not have to exist, and it becomes more convenient to consider irreducible representations of its Lie algebra. We shall defer the discussion of that to a later section of this chapter.

The (complex) dimension of the (finite-dimensional) complex vector space V in which the representation D is found has a dimension of the form $2s + 1$, where one refers to s as the *spin* of the representation. Hence, when that dimension is even, one must resort to half-odd-integer values for s. For instance, in the case of the defining representations of $SU(2)$ and $SL(2; \mathbb{C})$ in $\mathbb{C}^2$, $s = 1/2$. Since there is also a Hermitian conjugate representation in $\mathbb{C}^{2*}$, one distinguishes the two representations by the notations $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$, respectively; sometimes, the notations 2 and 2^* are also used.

The representations of $SU(2)$ are particularly straightforward, since it is compact and simple (i.e., it has no non-trivial normal subgroups). The maximal torii are all isomorphic to $U(1)$, which can be represented by the complex numbers of the form $e^{i\theta}$. Hence, in the case of the defining representation of $SU(2)$, the maximal torii will be matrices of the form:

$$\begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}.$$

Hence, the diagonal elements will define the weights of that representation, while the weight spaces (i.e., eigenspaces) that correspond to them will be spanned by the vectors $[1, 0]^{\mathrm{T}}$ and $[0, 1]^{\mathrm{T}}$, respectively. The higher-spin representations of $SU(2)$ will then

---

([1]) Actually, as Adams points out, there are compact, Abelian subgroups that do not take the form of T^k, such as the discrete subgroup of all matrices of the form diag[$\pm 1, \ldots, \pm 1$].

amount to tensor products of the defining representation and its Hermitian conjugate, which we will discuss in due course.

The representations of $SO(3)$ include the representations of $SU(2)$ by way of the isomorphism $SO(3) \cong SU(2) / \mathbb{Z}_2$. However, they also include the tensor products of the defining representation in $\mathbb{R}^3$ and its transpose in $\mathbb{R}^{3*}$. The latter two representations are then thought of as having spin 1, although that would really relate to the dimension of their complexifications. Interestingly, the complexification of $SO(3)$ to $SO(3, \mathbb{C})$ gives a Lie that is intimately related with the Lorentz group.

b. Representations as linear actions of groups. – The way that representations of groups enter into physical field theories is typically by way of the action of those groups on the points of space-time and the elements of the field space V (i.e., the vector space V in which the fields take their values). It is important to know that Sophus Lie was defining *transformation groups* (i.e., group actions on manifolds) before anyone actually formalized the present definition of a Lie group.

By definition, a *left action* of a Lie group G on a manifold M is a differentiable map $G \times M \to M$, $(g, x) \mapsto gx$ such that for every $x \in M$:

1) $ex = x$,
2) $(g_1 g_2)\, x = g_1(g_2\, x)$.

As a consequence:

3) $g^{-1}(g\, x) = x$.

A *right action* is a differentiable map $M \times G \to M$, $(x, g) \mapsto xg$ with analogous properties.

Typically, the distinction between left and right actions is meaningful only when one must define both at the same time. For instance, in order for the vector $\mathbf{v} = v^i\, \mathbf{e}_i$ to always represent the same element of a vector space independently of the choice of linear frame $\mathbf{e}_i$, when one left-multiplies the components v^i by an invertible matrix A^i_j, one must right-multiply the frame $\mathbf{e}_i$ by its inverse $\tilde{A}^i_j$.

If V is a vector space then the action $G \times V \to V$, $(g, \mathbf{v}) \mapsto g\mathbf{v}$ of a group G on a vector space V will be said to be *linear* iff for every $g \in G$ the *left-translation* map $L(g)$: $V \to V$, $\mathbf{v} \mapsto g\mathbf{v}$ is linear; $L(g)$ will be invertible as a result of the definition of the group action. Hence, $L(g)$ will belong to $GL(V)$, and the map $L : G \to GL(V)$, $g \mapsto L(g)$ will be a representation of G in $GL(V)$. The kernel of L consists of all elements of G that act trivially on V on the left (i.e., $g\mathbf{v} = \mathbf{v}$ for all $\mathbf{v}$). If $\ker(L) = e$ then the action is said to be *free*, and the representation L will be faithful.

The subgroup $G_\mathbf{v}$ of G that consists of all elements of G that fix the vector $\mathbf{v}$ is called the *isotropy subgroup* of the action at $\mathbf{v}$. When $G_\mathbf{v} = G$, one calls $\mathbf{v}$ a *fixed point* of the action, and at the other extreme, if $G_\mathbf{v} = e$ for every $\mathbf{v}$ then the action is called *effective.* An effective action will be free, since no element of G besides e will fix any point, but an action can be free without being effective, since there might be elements of G besides e that fix some points, but not others.

For a linear action of G on V, the *orbit* $G(\mathbf{v})$ of all vectors that $\mathbf{v}$ can go to under the action of G spans a linear subspace of V that is also an invariant subspace of the action. One then sees that the transformation group concept of an orbit is closely related to the representation concept of invariant subspace for a linear group action. The fact that they are not identical is due to the fact that a linear combination of vectors of the form $D(g_1)\mathbf{v}$ and $D(g_2)\mathbf{v}$ does not have to be of the form $D(g_3)\mathbf{v}$ for some $g_3 \in G$. For instance, an orbit might take the form of a circle that spans an invariant plane.

If the group action has only one orbit then one calls the action *transitive*. For any two elements $\mathbf{v}$, $\mathbf{v}' \in V$ there will be some element $g \in G$ such that $\mathbf{v}' = g\mathbf{v}$, but g is not generally unique; when g is always unique, one calls the action *simply transitive*.

Furthermore, if $\mathbf{v}$ and $\mathbf{v}'$ both belong to G(v) then $G_{\mathbf{v}} \cong G_{\mathbf{v}'}$, although the isomorphism will not be unique. As a differentiable manifold, one always has that $G(\mathbf{v})$ is diffeomorphic to $G / G_{\mathbf{v}}$, so if the isotropy subgroup $G_{\mathbf{v}}$ is the identity group then the orbit will be diffeomorphic to G itself. Hence, a simply-transitive action is both transitive and effective.

The most common examples of linear actions of groups on vector spaces that we shall be dealing with will be the defining representations of subspaces of either three-dimensional Euclidian space or four-dimensional Minkowski space and the spinorial representations of rotations and Lorentz transformations on $\mathbb{C}^2$ and $\mathbb{C}^4$.

c. Tensor representations of Lie groups. – In order to relate representations to the fields of physical field theory, one must first identify the vector space V that constitutes the field space. For elementary wave functions, that will typically be one of the real vector spaces $\mathbb{R}^n$, with $n = 1, 2, 3, 4$, or the complex vector spaces $\mathbb{C}^n$, also with $n = 1, 2, 3, 4$.

However, such elementary physical fields as electromagnetic and gravitational fields will involve going to "higher-spin" representations of physical symmetry groups, such as transformations of the points of space-time (viz., *dynamical* groups) or the elements of field space V (viz., *gauge transformations*). One must then consider tensor and spinor fields on space-time that take their values in vector spaces that can be decomposed into tensor products $V_1 \otimes \ldots \otimes V_r$ (r = *rank* of product) of other vector spaces; i.e., tensor and spinor fields. One must similarly consider representations of one's basic physical group G in the general linear group $GL(V_1 \otimes \ldots \otimes V_r)$.

A common source of confusion amongst physicists is due to the fact that the tensor product $V_1 \otimes \ldots \otimes V_r$ of any finite number of vector spaces is itself a *vector* space, in the sense that linear combinations of tensors of that rank are again tensors of that rank. Some physicists prefer to reserve the use of the word "vector" for what one could call "tensors of rank one," so to them the elements of a tensor space cannot also be vectors. Hopefully, that will not be a source of confusion in the present discussion.

We shall consider the simplest tensor product to begin with, namely, $V \otimes V$; the extension to higher-rank tensors will hopefully seem relatively straightforward.

When one changes from the frame $\mathbf{e}_i$ on V to another frame $\mathbf{f}_j$, one can represent that invertible linear transformation $T : V \to V$ by means of the matrix of components T^i_j of the new frame with respect to the initial one:

$$\mathbf{f}_j = T(\mathbf{e}_j) = \mathbf{e}_i T^i_j .$$

When one takes the tensor products of the new frame members, one can then say that:

$$\mathbf{f}_i \otimes \mathbf{f}_j = T(\mathbf{e}_i) \otimes T(\mathbf{e}_j) = \mathbf{e}_k T^k_i \otimes \mathbf{e}_l T^l_j = (\mathbf{e}_k \otimes \mathbf{e}_l) T^k_i \, T^l_j ,$$

due to the bilinearity of the tensor product.

Hence, one can say that there is a linear action of $GL(V)$ on $V \otimes V$ that is defined by the so-called *diagonal* representation of $GL(V)$ in $GL(V) \times GL(V)$, which takes every element T in $GL(V)$ to the element (T, T) in $GL(V) \times GL(V)$. Since (T, T) can then act linearly upon the elements of $V \otimes V$, one then has that $GL(V) \times GL(V)$ can be identified with a proper subgroup of $GL(V \otimes V)$. If one has a representation D of a group G in $GL(V)$ then one then calls the composition of group homomorphisms:

$$G \to GL(V) \to GL(V) \times GL(V) \subset GL(V \otimes V)$$

that takes g in G to $D(g)$ in $GL(V)$ to $(D(g), D(g))$ in $GL(V) \times GL(V)$ to the corresponding transformation of $V \otimes V$ the *tensor product of D with itself* or the *representation of G in the tensor product space $V \otimes V$*. One usually sees that representation take the form of the formula for the transformation of components under the transformation of the frame to which they are referred:

$$\bar{t}^{ij} = T^i_k \, T^j_l \, t^{kl} , \tag{2.2}$$

although the transformation T of components must be inverse (or *contragredient*) to the transformation of the frame in order for the tensor itself to be invariant.

One can also define the tensor product of two different representations $D_1, D_2 : G \to GL(V)$, but that construction will not be of much use to us in what follows. However, it does play an important role in chiral representations, for which one no longer uses only the diagonal subgroup of $G \times G$.

It is important to note the difference in dimensions between the groups $GL(V) \times GL(V)$ and $GL(V \otimes V)$: The former has dimension $2n^2$, while the latter has dimension $n^2 \times n^2 = n^4$; of course, the diagonal subgroup of $GL(V) \times GL(V)$ still has the same dimension as $GL(V)$, namely, n^2. Hence, there will generally be many more invertible linear transformations of $V \otimes V$ than the ones that come from tensoring a representation in $GL(V)$ with itself.

Even when the representation of G in $GL(V)$ is irreducible, its representation in $GL(V \otimes V)$ does not have to be. For instance, one has a direct sum decomposition:

$$V \otimes V = S_2 \oplus A_2 ,$$

in which $S_2 = S_2V$ is the vector space of symmetric, second-rank contravariant tensors, while $A_2 = A_2V$ is the vector space of antisymmetric ones.

One finds that symmetry type is preserved by the transformation (2.2):

$$\bar{t}^{ji} = T_k^j\, T_l^i\, t^{kl} = T_l^i T_k^j\, t^{kl} = \pm\, T_l^i T_k^j\, t^{lk} = \pm\, \bar{t}^{ij}\,.$$

Hence, S_2 and A_2 are invariant subspaces under the tensor product representation of any G in $GL(V)$, so that representation will be reducible.

In fact, if one has a scalar product defined on V then one can further reduce the representation in S_2 by the fact that a scalar product allows one to define a linear isomorphism $V \to V^*$, $\mathbf{v} \mapsto \mathbf{v}^*$, where $\mathbf{v}^*(\mathbf{w}) = <\mathbf{v}, \mathbf{w}>$ for any $\mathbf{w}$. One can then associate every element of $V \otimes V$ with an element of $V^* \otimes V$, which can be identified with the algebra of (not-necessarily-invertible) linear maps from V to itself, which then have matrices t_j^i with mixed components. One can then define a trace function on $V^* \otimes V$ and distinguish between the elements of $V^* \otimes V$ with trace zero and those with non-zero trace; that implies a corresponding decomposition of S_2 into $\mathbb{K} \oplus \overset{\circ}{S}_2$, where the elements of $\overset{\circ}{S}_2$ correspond to the matrices with zero trace. For a symmetric matrix s_j^i, this decomposition takes the form:

$$\frac{1}{n} t_k^k I + \overset{\circ}{t}{}^i_j\,,$$

in which $\overset{\circ}{t}{}^k_k = 0$. It is common to define:

$$\overset{\circ}{t}{}^i_j = t_j^i - \frac{1}{n} t_k^k\,,$$

but that decomposition is not unique, since one can add any matrix with zero trace to $\overset{\circ}{t}{}^i_j$ and still produce an acceptable decomposition of t_j^i into a traceless matrix and a trace part.

The decomposition $V \otimes V = \mathbb{K} \oplus \overset{\circ}{S}_2 \oplus A_2$ does, in fact, give a decomposition of the tensor product representation of G in $GL(V)$ into a direct sum of irreducible representations that one calls the *Clebsch-Gordan decomposition* of that representation.

Although it is customary to discuss the way that one can use the "Clebsch-Gordan coefficients" and "Young tableaux" in order to compute the dimensions of the invariant subspaces of a tensor product representation, nevertheless, we shall not actually need that machinery in what we shall do, so we shall pass over that subject. (See, however, [**5**].)

We need to make the essential comment that the tensor-product representations of the three-dimensional rotation group – or rather, $SU(2)$ – constitute the essence of the (non-relativistic) quantum theory of angular momentum. In particular, the Clebsch-Gordan decomposition relates to the theory of the addition of angular momenta.

d. Spinor representations of Lie groups. – The 2-1 homomorphisms that we mentioned above – namely, $SU(2) \to SO(3)$ and $SL(2; \mathbb{C}) \to SO_0(3, 1)$ – both play an important role in quantum mechanics. Indeed, it was only the discovery of electron spin that drew the attention of the physics mainstream to those homomorphisms in the first place when the former homomorphism was used by Pauli to include the spin of the electron in non-relativistic wave mechanics, and the latter was used by Dirac to discuss the relativistic theory.

In both cases, the main difference between the two Lie groups involved is topological: $SU(2)$ and $SL(2; \mathbb{C})$ are both *simply-connected* as topological spaces. Hence, any continuous loop in them can be continuously deformed to a point without leaving the manifold. The 2-1 character of the projection relates to the fact that the image Lie groups both have $\mathbb{Z}_2$ for their fundamental groups. The simply-connected Lie groups $SU(2)$ and $SL(2; \mathbb{C})$ then become *covering groups* for $SO(3)$ and $SO_0(3, 1)$, respectively, and the projections are *covering homomorphisms*.

It has also become customary amongst mathematicians to refer to the simply-connected covering group as the "spin" group that is associated with the multiply-connected one, and the representation of an element of $SO(3)$ or $SO_0(3, 1)$ by a pair of antipodal matrices in $SU(2)$ or $SL(2; \mathbb{C})$, resp., is the *spin representation* of those groups. That is then the basis for the construction of *spinor representations* of rotations or Lorentz transformations.

There are two basic differences between a tensor and a spinor:

1. The basic vector space is $\mathbb{C}^n$ ($n = 2$ or 4, in most cases), not $\mathbb{R}^n$.

2. The components transform by the matrices of $SU(2)$ or $SL(2; \mathbb{C})$, not $SO(3)$ or $SO(3, 1)$.

Before one gets to actual spinors, one first looks at how one represents the vector spaces $\mathbb{R}^3$ and $\mathbb{R}^4$ as complex 2×2 matrices and then how the rotations or Lorentz transformations act upon those vectors when they are represented in that way. We shall address both cases in parallel.

The complex vector space $M(2; \mathbb{C})$ of all 2×2 complex matrices has a complex dimension of four. Hence, a basis for it will consist of four linearly-independent matrices. It is customary to use the identity matrix, along with the three Pauli matrices for a basis:

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.3}$$

All of these matrices are Hermitian, and the last three have trace zero.

Any matrix A in $M(2; \mathbb{C})$ can then be expressed uniquely in the form:

$$A = A^{\mu}\,\sigma_{\mu}, \qquad \mu = 0, \ldots, 3 \tag{2.4}$$

for some set of four complex components A^{μ}. If one restricts oneself to only real components x^{μ} then the resulting set of matrices of the form:

$$[\boldsymbol{x}] = x^{\mu}\,\sigma_{\mu} = \begin{bmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 - ix^2 & x^0 - x^3 \end{bmatrix} \tag{2.5}$$

will define a real vector space that is linearly-isomorphic to $\mathbb{R}^4$. If one wishes to represent only spatial vectors then one sets $x^0 = 0$, so:

$$[\boldsymbol{x}] = x^{i}\,\sigma_{i} = \begin{bmatrix} x^3 & x^1 - ix^2 \\ x^1 - ix^2 & x^3 \end{bmatrix}, \qquad i = 1, 2, 3. \tag{2.6}$$

If one looks at the determinants in both cases then one will find that:

$$\det\,[\boldsymbol{x}] = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = \eta(\mathbf{x}, \mathbf{x}) \tag{2.7}$$

in the four-dimensional case and:

$$\det\,[\boldsymbol{x}] = -\,(x^1)^2 - (x^2)^2 - (x^3)^2 = -\,\delta(\mathbf{x}, \mathbf{x}) \tag{2.8}$$

in the spatial case.

Any matrix A in $M(2; \mathbb{C})$ acts upon other matrix, such as $[\boldsymbol{x}]$, by way of:

$$A[\boldsymbol{x}] = A^{\dagger}[\boldsymbol{x}]\,A. \tag{2.9}$$

Hence, since both the matrices of $SU(2)$ and $SL(2; \mathbb{C})$ live in $M(2; \mathbb{C})$, the former group acts upon spatial vectors that are presented in the form (2.6), while the latter group acts upon space-time vectors that are represented in the form (2.5). By the product rule for determinants:

$$\det\,(A^{\dagger}[\boldsymbol{x}]\,A) = \det A^{\dagger}\,\det\,[\mathbf{x}]\,\det A = \det\,[\mathbf{x}], \tag{2.10}$$

since $\det A = \det A^{\dagger} = 1$ for the matrices of both $SU(2)$ and $SL(2; C)$. Therefore, the actions preserve the Euclidian spatial and Minkowski space-time scalar product, respectively, and one is truly representing the action of rotations and Lorentz transformations on vectors that have been represented as 2×2 complex matrices. However, one should notice that since A is represented twice in (2.9) and $(-A)^{\dagger} = -A^{\dagger}$, it will not matter whether one uses the matrix A or $-A$, since the sign will square to positive in either case. Thus, despite the use of complex matrices, the action that we have defined is not a faithful representation of $SU(2)$ or $SL(2; \mathbb{C})$ in $M(2; \mathbb{C})$, but only a faithful representation of $SO(3)$ or $SO_0(3, 1)$, respectively.

The concept of an elementary spinor then emerges when one looks at the defining representations of the Lie groups in question, along with their Hermitian conjugate representations on the dual space. In both cases, the vector space of the defining representation will be $\mathbb{C}^2$, and its dual space will be $\mathbb{C}^{2*}$. Actually, for the relativistic case, it is more customary to consider the vector space $\mathbb{C}^4$ of *bi-spinors* and its dual, but it is important to realize that one can still represent $SL(2; \mathbb{C})$ in $\mathbb{C}^2$, and we shall return to that topic later on in our discussion of the relativistic Pauli equation.

We shall follow the tradition of Dirac [**6**] and use the "bra-ket" formalism for notating elements of $\mathbb{C}^n$ and its dual. A *ket*, which can be represented as a column vector:

$$| \psi > = \begin{bmatrix} \psi^1 \\ \vdots \\ \psi^n \end{bmatrix}, \tag{2.11}$$

is an element of $\mathbb{C}^n$, while a *bra*, which can be represented as a row vector:

$$< \psi | = [\psi_1 , \ldots, \psi_n], \tag{2.12}$$

is an element of $\mathbb{C}^{n*}$.

The way that one goes from one to the other is by way of the Hermitian conjugate:

$$| \psi >^\dagger = < \psi^* |, \tag{2.13}$$

and the effect of the bilinear pairing $<\psi \mid \psi' >$ is to reproduce the Hermitian inner product. (In quantum mechanics, one also typically has to integrate the real function $<\psi | \psi'>$ over all space, but for the present purposes that will not be necessary.)

A complex linear transformation $T : \mathbb{C}^n \to \mathbb{C}^n$ acts upon a ket $| \psi >$ to the left $T | \psi >$ and its Hermitian conjugate $T^\dagger$ acts upon a bra from the right $< \psi | T^\dagger$. Hence, it will be unitary iff

$$<\psi | T | \psi'> = <\psi | \psi'> \tag{2.14}$$

in any case.

Since the action of $SU(2)$ or $SL(2; \mathbb{C})$ by way of (2.9) is quadratic, while their action on elements of $\mathbb{C}^2$ or $\mathbb{C}^{2*}$ is linear, one sees that only the latter action is capable of resolving the difference between the two antipodal matrices that represent a rotation or Lorentz transformation. Hence, in a sense, spinors are "square roots" of vectors, although, not in the sense of tensor product, since the tensor product of any two spinors will be, by definition, decomposable, and not all second-rank tensors over $\mathbb{C}^2$ are decomposable.

A particular simple example of the Clebsch-Gordan decomposition is when one looks at the tensor product representation of SU (2) in $\mathbb{C}^2 \otimes \mathbb{C}^2$, which one denotes by $2 \otimes 2$ or $(\frac{1}{2}, 0) \otimes (0, \frac{1}{2})$. If one writes a basis for $\mathbb{C}^2$ in the form of the vectors $u = [1, 0]^{\mathrm{T}}$ and $d = [0, 1]^{\mathrm{T}}$ then a basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$ will be defined by $\{u \otimes u, u \otimes d, d \otimes u, d \otimes d\}$. However, since this representation of SU (2) is not irreducible, one can first polarize the elements of $\mathbb{C}^2 \otimes \mathbb{C}^2$ into symmetric and antisymmetric ones, which will then give the Clebsch-Gordan decomposition of the representation in the form $\mathbb{C}^2 \otimes \mathbb{C}^2 = S_2(\mathbb{C}^2) \oplus A_2(\mathbb{C}^2)$, where the three-dimensional complex vector space S_2 $(\mathbb{C}^2)$ consists of the symmetric second-rank tensors over $\mathbb{C}^2$, and the one-dimensional vector space A_2 $(\mathbb{C}^2)$ consists of the antisymmetric ones. These representations are irreducible, and one sometimes notates the decomposition in the form $2 \otimes 2 = 3 \oplus 1$. A basis for $S_2(\mathbb{C}^2)$ can be defined by $\{u \otimes u, u \odot d, d \otimes d\}$ and a basis for A_2 $(\mathbb{C}^2)$ can be defined by $\{u \wedge d\}$, in which $\wedge$ means the exterior product and $\odot$ is the symmetric product:

$$u \odot d = \tfrac{1}{2}(u \otimes d - d \otimes u). \tag{2.15}$$

We shall deal with the Dirac bi-spinors, which are based in $\mathbb{C}^4$, in the last chapter, since they require special treatment. Similarly, our discussion of non-relativistic, spinning, wave mechanics will enlarge upon the discussion of Pauli spinors, which are based in $\mathbb{C}^2$.

e. Representations of groups on field space. – We shall now attempt to apply the abstract mathematical notations that we just discussed to the case of physical field theories.

Since we will not be addressing topological issues in what follows, for us, a *field* will usually take the form of a "sufficiently-differentiable" (or simply smooth) map $\Psi : S \to \mathbb{C}^r$, in which S is an open subset of $\mathbb{R}^n$, which will generally be either Euclidian space $E^3 = (\mathbb{R}^3, \delta_{ij})$, Galilean space-time $\mathbb{R} \times E^3$, or Minkowski space $\mathfrak{M}^4 = (\mathbb{R}^4, \eta_{\mu\nu})$. Typically, S will take the form of the set of points on which Ψ is non-zero, so it closure will be the *support* of Ψ. Furthermore, even in the four-dimensional case, S will usually have a cylindrical topology, in the sense that $S = \mathbb{R} \times \Sigma^3$, where Σ^3 is a three-dimensional differentiable manifold with compact closure. The latter constraint tends to be an unavoidable consequence of the demands of time evolution when one neglects the possibility of topology-changing processes, such as the formation of vortex pairs.

It is often useful to consider the *graph* of the field Ψ, which is the subset of $S \times \mathbb{C}^r$ that consists of all points of the form $(x, \Psi(x))$; when no coordinate system has been specified, we use the notation x for any point of S, and not just its spatial projection. Hence, the graph of Ψ can also be regarded as a *section* of the projection $S \times \mathbb{C}^r \to S$, $(x, z^a) \mapsto x$, so it is an injective map $\Psi : S \to S \times \mathbb{C}^r$, $x \mapsto (x, \Psi(x))$. (Had we chosen to go the topological route, $S \times \mathbb{C}^r$ would become a "local trivialization" of a "complex vector bundle" of "rank r" over S, so Ψ would become a section of that vector bundle.)

We define the set $\Gamma(S, \mathbb{C}^r)$ of all sections of the projection $S \times \mathbb{C}^r \to S$ (or equivalently, the space of all functions $\Psi : S \to \mathbb{C}^r$). It is an infinite-dimensional complex vector space, although in many cases, its complex structure can prove to be distracting, since the set S is assumed to be a *real*, *finite-dimensional* differentiable manifold, so the complex structure on $\Gamma(S, \mathbb{C}^r)$ is confined to the field space $\mathbb{C}^r$. Hence, when differentiating Ψ, one will be taking real derivatives, not complex ones, and any complex operations, such as complex scalar multiplication and conjugation, can refer only to the field space.

The action of a Lie group G on sections can then be defined by an action of G on $\Gamma(S, \mathbb{C}^r)$, which, in turn can be defined by the action of G on $S \times \mathbb{C}^r$. Typically, we shall consider only the simplest kind of action, which involves G acting upon S independently of $\mathbb{C}^r$, which we think of as the "horizontal" part of the action, and acting upon $\mathbb{C}^r$ in two different ways, both of which are assumed to be linear actions. Firstly, it acts independently of its action on S, and secondly, the Lie algebra $\mathfrak{g}$ acts upon $\mathbb{C}^r$ by way of the differential map $d\Psi$, which we shall discuss below. There will also be a corresponding action upon $\Psi^*(x)$ by means of the complex conjugates of the linear transformations on $\mathbb{C}^r$ that represent each element of G, and a representation of $\mathfrak{g}$ in $\mathfrak{gl}(\mathbb{C}^r)$ by way of $d\Psi^*$.

3. Representations of Lie algebras. – For the purposes of the calculus of variations, it is generally more useful to know how the infinitesimal generators of finite transformations act upon points of space-time and elements of field space. That is because in many cases the variations of the field in question will be generated by such infinitesimal generators of (one-parameter subgroups of) finite transformations. Hence, one will be dealing with elements of the Lie algebra $\mathfrak{g}$ of some Lie group G.

a. Basic definitions regarding Lie algebras. – A *Lie algebra* is a vector space $\mathfrak{g}$ that is given a bilinear pairing $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, $(a, b) \mapsto [a, b]$ that has the properties:

1. *Antisymmetry:* $[a, b] = -[b, a]$,
2. *Jacobi identity:* $[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$.

The first condition says that the algebra product – namely, the *Lie bracket* $[a, b]$ – is not generally commutative, since if it were then one would also need to have:

$$[a, b] = 0$$

in any event. In such a case, one would call the Lie algebra *Abelian.* Another consequence is that one must always have $[a, a] = 0$ ([1]).

The second condition says that the algebra product is typically not associative, either, since one will generally have:

$$[a, [b, c]] = [[a, b], c] - [b, [c, a]].$$

b. The Lie algebra of a Lie group. – Any differentiable manifold is associated with an infinite-dimensional Lie algebra in the form of the tangent vector fields on it, but in the case of Lie groups, one can identify sub-algebras of that infinite-dimensional algebra that have the same dimension as the group itself. In one case, one considers the action of the group G on itself by left-translation, and in the other, by right-translation. As a result, one can say that vector fields on G that are invariant under left-translation are *left-invariant*, and analogously for *right-invariant* vector fields. Both of them define isomorphic Lie algebras, but since any right or left invariant vector field can be specified by its vector at any point of G, one usually specifies the vector field by a vector at the identity e and uses the tangent space T_eG to represent the Lie algebra $\mathfrak{g}$ of G. One defines the Lie bracket $[a, b]$ of tangent vectors a, b at e to be the tangent vector at e that corresponds to the Lie bracket $[\tilde{a}, \tilde{b}]$ of the left (or right) invariant vector fields $\tilde{a}$, $\tilde{b}$ that the elements a, b generate.

Since a tangent vector at a point can be regarded as an equivalence class of differentiable curves through that point that differentiate to the same vector in any coordinate chart, one can obtain an element of $\mathfrak{g} = T_eG$ by passing a differentiable curve through e and differentiating it at e. For instance, if $\gamma(s)$ is such a curve through e (say, with $\gamma(0) = e$) then one can obtain an element $a \in \mathfrak{g}$ by way of:

$$a = \left.\frac{d\gamma}{ds}\right|_{s=0} .$$

Conversely, one can start with a and obtain a differentiable curve $A(s)$ by extending α to a left (or right)-invariant vector field on G and looking at the integral curve to it that

([1]) If one wishes to quibble about the best way to characterize antisymmetry, one must include the possibility that the field of scalars is $\mathbb{Z}_2$. Hence, in that case, although $[a, a] = 0$ for all a would imply that $[a, b] + [b, a] = 0$ for all a, b, the converse would not have to be true, since $[a, a] + [a, a] = 2[a, a] = 0$ for any a, regardless of whether $[a, a]$ is or is not zero. However, we shall have no use for such exotic pathologies.

goes through e. Such a curve will then be a *one-parameter subgroup* of G. The curve is easiest to obtain for Lie groups that consist of invertible matrices, since one can define the exponential of any matrix a:

$$\exp(a) = \sum_{n=0}^{\infty}\frac{1}{n!}a^n . \tag{3.1}$$

The one-parameter subgroup that a generates is the image of the line through the origin of T_eG and a under the exponential map:

$$A(s) = \exp(sa). \tag{3.2}$$

One then refers to a as the *infinitesimal generator* of $A(s)$; more casually, the elements of $\mathfrak{g}$ are the "infinitesimal transformations" that are associated with elements of G.

A useful construction on any group, and especially the orthogonal and unitary groups that one encounters in mechanics and gauge field theories is the *Cartan-Killing form.* If $\mathfrak{g}$ is a Lie algebra, and $a, b \in \mathfrak{g}$ are any two elements, then one can define their Cartan-Killing form to be the symmetric bilinear functional:

$$\langle a, b\rangle = \mathrm{Tr}\, \mathrm{ad}(a)\, \mathrm{ad}(b),$$

where $\mathrm{ad}(a) : \mathfrak{g} \to \mathfrak{g}$ is the linear map that takes any $c \in \mathfrak{g}$ to $[a, c]$, and similarly for $ad(b)$.

When $\mathfrak{g}$ is a matrix Lie algebra to begin with, and $A, B \in \mathfrak{g}$ are any two of its matrices, one will have:

$$\langle A, B\rangle = \mathrm{Tr}(AB).$$

The Cartan-Killing form will define a scalar product on $\mathfrak{g}$ when $\mathfrak{g}$ is semi-simple (i.e., contains no proper Abelian ideals), and it will be definite when the underlying Lie group is compact.

In particular, for $\mathfrak{so}(3)$, one has the basis:

$$I_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad I_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \qquad I_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{3.3}$$

which can be expressed in the single defining equation:

$$[I^i_j]_k = \varepsilon_{ijk} . \tag{3.4}$$

One will then have:

$$\langle I_a, I_b\rangle = -2\, \delta_{ab} . \tag{3.5}$$

Hence, as an orthogonal space, $\mathfrak{so}(3)$ is essentially three-dimensional, real, Euclidian space.

In the case of $\mathfrak{so}(3, 1)$, if one uses the basis:

$$J_1 = \left[\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right], \quad J_2 = \left[\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right], \quad J_3 = \left[\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

$$K_1 = \left[\begin{array}{c|ccc} 0 & 1 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \quad K_2 = \left[\begin{array}{c|ccc} 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \quad K_3 = \left[\begin{array}{c|ccc} 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right]$$

then the matrix of the Cartan-Killing form will take the block-diagonal form:

$$\left[\begin{array}{c|c} -2\delta_{ab} & 0 \\ \hline 0 & 2\delta_{ab} \end{array}\right].$$

This is also the signature type of the scalar product that one defines on the six-dimensional real vector spaces of 2-forms or bivectors on $\mathbb{R}^4$ when one introduces the scalar product:

$$<F, G> = (F \wedge G)(\mathbf{V}),$$

in the case of 2-forms, for instance. Here, $\mathbf{V} \in \Lambda^4(\mathbb{R}^4)$ is a choice of volume element.

c. Representations of Lie algebras. – A *representation* of a Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism $\mathfrak{D}: \mathfrak{g} \to \mathfrak{gl}(V)$, where $\mathfrak{gl}(V)$ is the Lie algebra of all (not-necessarily invertible) linear maps from V to itself. Hence, $\mathfrak{D}$ must be a linear map of the underlying vector spaces, while one must always have:

$$[\mathfrak{D}(a), \mathfrak{D}(b)] = [a, b].$$

Such a representation is called *faithful* iff the linear map $\mathfrak{D}$ is injective. That is true iff the kernel of $\mathfrak{D}$ vanishes, and ker($\mathfrak{D}$) is always a Lie subalgebra of $\mathfrak{g}$, since if a and b both map to zero under $\mathfrak{D}$, so will $[\mathfrak{D}(a), \mathfrak{D}(b)]$, by the linearity of $\mathfrak{D}$ and the bilinearity of the bracket. Moreover, if $\mathfrak{D}(k) = 0$ then $[\mathfrak{D}(a), \mathfrak{D}(k)] = 0$ for any $\mathfrak{D}(a) \in \mathfrak{g}$, so ker $\mathfrak{D}$

will also be an *ideal* in $\mathfrak{g}$, and the difference vector space $\mathfrak{g} - \ker \mathfrak{D}$ ([1]) will be a Lie algebra in its own right. Similarly, the image of $\mathfrak{D}$ is always a Lie subalgebra of $\mathfrak{gl}(V)$, since if $\mathfrak{D}(a)$ and $\mathfrak{D}(b)$ are elements of $\mathfrak{D}(\mathfrak{g})$ then $[\mathfrak{D}(a), \mathfrak{D}(b)]$ must be the image of $[a, b]$.

If one has a representation $D : G \to GL(V)$ of a Lie group G then one can obtain a representation $\mathfrak{D}: \mathfrak{g} \to \mathfrak{gl}(V)$ of the Lie algebra $\mathfrak{g}$ in the Lie algebra $\mathfrak{gl}(V)$ by differentiating D at e. One way of doing that is to start with a curve $\gamma(s)$ through $e \in G$ whose tangent vector at e is a, map $\gamma(s)$ to a curve $\mathfrak{D}(\gamma(s))$ in $GL(V)$, which must also go through $I \in GL(V)$, and differentiate the image curve at I in order to obtain the tangent vector $\mathfrak{D}(a)$ in $T_I GL(V) \cong \mathfrak{gl}(V)$.

When one thinks of the representation of G in $GL(V)$ as a linear action of G on V, one will find that the differential of that representation will associate elements $a \in \mathfrak{g}$ with vector fields $\tilde{a}$ on V that one calls *fundamental vector fields.* In particular, the fundamental vector fields on V will define a Lie subalgebra of the Lie algebra $\mathfrak{X}(V)$ of vector fields on V.

Basically, one starts with a, generates a one-parameter subgroup $\exp(sa)$, lets it act upon a vector $\mathbf{v} \in V$ to produce a differentiable curve $\mathbf{v}(s) = \exp(sa)\mathbf{v}$ in V and differentiates that curve at $\mathbf{v}$ (i.e., $s = 0$):

$$\tilde{a}(\mathbf{v}) = \left.\frac{d}{ds}\right|_{s=0} \exp(sa)\mathbf{v}. \tag{3.6}$$

If $(x^1, \ldots, x^n)$ is a coordinate system for V, so $\{\partial_1, \ldots, \partial_n\}$ is the natural frame field that it defines then any vector field on V can be represented in the form:

$$X(\mathbf{v}) = X^i(\mathbf{v})\, \partial_i \tag{3.7}$$

for a unique set of smooth functions $X^i(\mathbf{v})$, and that includes the fundamental vector fields for a group action on V. All that one needs to do is to choose a basis $\{\varepsilon_1, \ldots, \varepsilon_r\}$ for $\mathfrak{g}$ and map those basis elements to corresponding fundamental vector fields $\{\tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_r\}$ on V, and all of the other fundamental vector fields on V will have constant components with respect to that basis.

One example of a representation of a Lie algebra by fundamental vector fields for the action of its Lie group on a vector space V is the representation of the Lie algebra of infinitesimal translations of V by vector fields of the form:

$$\tilde{\varepsilon}(\mathbf{v}) = \varepsilon^i\, \partial_i, \tag{3.8}$$

([1]) I.e., the vector space of all translates of $\ker \mathfrak{D}$ in $\mathfrak{g}$; its vectors are then equivalence classes of vectors in $\mathfrak{g}$ that differ by an element of $\ker \mathfrak{D}$. The algebraic operations on the equivalence classes are the projections of the operations on the elements of $\mathfrak{g}$ under the projection $\mathfrak{g} \to \mathfrak{g} - \ker \mathfrak{D}$.

in which the components ε^i are constants.

Another example is the representation of the infinitesimal rotations of $\mathfrak{so}(3)$ by fundamental vector fields on $\mathbb{R}^3$ that one gets by starting with an infinitesimal rotation $\omega \in \mathfrak{so}(3)$, which we regard as consisting of real, anti-symmetric 3×3 matrices, exponentiating it to a one-parameter subgroup $R(s) = \exp(s\omega)$, letting the matrices act upon the points of $\mathbb{R}^3$:

$$\bar{x}^i(s) = R^i_j(s) x^j ,$$

and differentiating those curves $\bar{x}^i(s)$ at $s = 0$:

$$\tilde{\omega}^i(x) = \left.\frac{d}{ds}\right|_{s=0} R^i_j(s) x^j = \omega^i_j x^j . \tag{3.9}$$

In particular, if one uses the basis for $\mathfrak{so}(3)$ that is defined by the I_i in (3.3) then the corresponding fundamental vector fields will be:

$$\tilde{I}_k(x) = [I^i_j]_k \; x^j \, \partial_i = \varepsilon_{ijk} \, x^j \, \partial_i \, ; \tag{3.10}$$

i.e.:

$$\tilde{I}_1(x, y, z) = -z\frac{\partial}{\partial y} + y\frac{\partial}{\partial z}, \quad \tilde{I}_2(x, y, z) = z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}, \quad \tilde{I}_3(x, y, z) = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}, \tag{3.11}$$

respectively.

d. Weights of representations of Lie algebras. – As mentioned above, for non-compact Lie groups, such as the Lorentz group or $SL(2; \mathbb{C})$, it becomes more convenient to deal with the representations of their Lie algebras, instead of the Lie groups themselves. Fortunately, for the most purposes in physics – in particular, variational field theory – it is really the representations of the Lie algebra that are most useful. Of course, if one has a representation $D : G \to GL(V)$ of a Lie group G then one can get a representation of its Lie algebra $\mathfrak{D}: \mathfrak{g} \to \mathfrak{gl}(V)$ by differentiating D at the identity. The opposite process is harder to define, since more than one Lie group can have the same Lie algebra.

The infinitesimal analogue of a maximal torus is a *Cartan subalgebra*, namely, a maximal, Abelian subalgebra $\mathfrak{h} \subset \mathfrak{g}$. In fact, if a Lie group has a maximal torus then the tangent space to it at the identity will define a Cartan subalgebra. The dimension of a Cartan subalgebra is called the *rank* of the Lie algebra.

Once again, when one wishes to enumerate the irreducible representations of a Lie algebra $\mathfrak{D} : \mathfrak{g} \to \mathfrak{gl}(V)$, it is sufficient to enumerate the irreducible representations of any of its Cartan subalgebras $\mathfrak{h}$. If one restricts $\mathfrak{D}$ to $\mathfrak{D} : \mathfrak{h} \to \mathfrak{gl}(V)$, then since the elements a

$\in \mathfrak{h}$ all commute with each other, so will their representatives $\mathfrak{D}(a)$ in $\mathfrak{gl}(V)$; therefore, the matrices $\mathfrak{D}(a)$ will be simultaneously diagonalizable, and the diagonal elements will be the weights of the representation, namely, the eigenvalues of the matrix $\mathfrak{D}(a)$. Similarly, the vector space V will decompose into a direct sum of weight spaces.

In the case of Lie algebras, one finds that the eigenvector equation:

$$\mathfrak{D}(a)\,\mathbf{v} = \lambda(a)\,\mathbf{v} \tag{3.12}$$

makes $\lambda(a)$ into a linear functional on a Cartan subalgebra $\mathfrak{h}$. Hence, one can think of the weights of a representation as something that belong to the dual space $\mathfrak{h}^*$.

If we return to the case of $SU(2)$ then we will find that since its maximal torii were all circles, the Cartan subalgebras of $\mathfrak{su}(2)$ will all be lines through the origin. Since the elements of $\mathfrak{su}(2)$ are anti-Hermitian 2×2 complex matrices, it is common to use a basis for $\mathfrak{su}(2)$ that amounts to the Pauli matrices σ_i times $\pm\, i$; we shall use $-\, i$:

$$\tau_i = -\, i\sigma_i\,. \tag{3.13}$$

One can think of this process as a higher-dimensional analogue of the process of taking a real angle θ and multiplying it by i to produce an element of the Lie algebra of $U(1)$, which will be the imaginary line.

Whereas the elements of a maximal torus in $SU(2)$ took the form of diag$[e^{i\theta}, e^{-i\theta}]$, by differentiation, its corresponding Cartan subalgebra $\mathfrak{h}$ would take the form diag$[i\theta, -\, i\theta]$, although since $\mathfrak{h}$ is a one-dimensional linear subspace of $\mathfrak{g}$, one could just as well use diag$[i, -\, i] = i\,\sigma_3 = -\,\tau_3$.

Any of the three matrices τ_i can be used as the basis (i.e., the generator h) for a Cartan subalgebra, and we choose h to be τ_3. We can then define:

$$e_\pm = \tau_1 \pm i\tau_2\,, \tag{3.14}$$

which will give the commutation relations:

$$[h, e_\pm] = \pm\, e_\pm\,, \qquad [e_+\,, e_-] = h\,. \tag{3.15}$$

The elements $e_\pm$ can be regarded as "raising" and "lowering" operators in regard to the weights of a given representation $\mathfrak{D} : \mathfrak{su}(2) \to \mathrm{gl}(V)$, since if $\mathbf{v} \in V$ is an eigenvector of $\mathfrak{D}(h)$ of weight a then $\mathfrak{D}(e_\pm)\mathbf{v}$ will be eigenvectors of $\mathfrak{D}(h)$ of weights $a \pm 1$, respectively. Hence, $\mathfrak{D}(e_\pm)$ will raise or lower the weight of the representation of h in $\mathfrak{gl}(V)$ by one unit.

Tensor products of representations of Lie algebras are defined analogously to the tensor products of Lie groups, and can also be obtained by differentiating the latter.

Now that we have introduced fundamental vector fields for the action of a group G on a differentiable manifold M, which defines a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{X}(M)$, one can take advantage of the fact that tangent vectors on M are usually represented as linear, first-order differential operators on the infinite-dimensional vector space $C^\infty(M)$ of smooth functions on M, namely, as directional derivative operators. Similarly, tangent vector fields can also be regarded as such:

$$X f = X^i \frac{\partial f}{\partial x^i}. \tag{3.16}$$

Hence, one also regard this as a representation $\mathfrak{D}$: $\mathfrak{g} \to \mathfrak{gl}(C^\infty)$ in the (infinite-dimensional) Lie algebra of linear transformations of $C^\infty(M)$. For the purposes of quantum mechanics, one usually has the functions take their values in $\mathbb{C}$, so one can think of the eigenvalues of the vector field X, when regarded as a differential operator, as relating to the weights of that representation, at least when one restricts $\mathfrak{D}$ to a Cartan subalgebra of $\mathfrak{g}$.

This type of representation is most commonly used in the representation of linear and angular momentum as Hermitian operators on wave functions, which is part of the process of "canonical quantization." For instance, the linear momentum covector p whose components are p_i , which belongs to the Abelian Lie algebra $\mathbb{R}^3$, will be represented by the triple of linear, first-order differential operators:

$$P_i = \frac{\hbar}{i}\frac{\partial}{\partial x^i}, \tag{3.17}$$

which also commute:

$$[P_i , P_j] = 0. \tag{3.18}$$

The differential operators P_i are $\hbar / i$ times the fundamental vector fields that are associated with a basis for the Lie algebra $\mathbb{R}^3$.

The eigenfunctions ψ of these operators are then solutions of:

$$P_i \psi = \frac{\hbar}{i}\frac{\partial \psi}{\partial x^i} = p_i \psi, \tag{3.19}$$

namely:

$$\psi_i = e^{ip_i x_i / \hbar} \qquad \text{(no sum over } i\text{)}. \tag{3.20}$$

Similarly, the components of an angular momentum covector ω whose components are ω_i , which then belongs to the Lie algebra $\mathfrak{so}(3) \cong \mathfrak{su}(2)$, go to the differential operators:

$$L_i = (\mathbf{x} \times \mathbf{P})_i = \frac{\hbar}{i}\varepsilon_{ijk} x_j \frac{\partial}{\partial x^j}, \tag{3.21}$$

which are $\hbar / i$ times the fundamental vector fields associated with a basis for $\mathfrak{so}(3)$.

One then finds that:

$$[L_i, L_j] = \varepsilon_{ijk} L_k . \tag{3.22}$$

Due to the symmetries of the sphere, it is usually more convenient to represent the components of ω by means of partial derivatives with respect to spherical coordinates, although we shall not go into that here, but refer the curious to any good book on the quantum theory of angular momentum [**7**].

We shall deal with the specifics of the representations of the Lorentz group in the chapter on relativistic, spinning particles.

References

1. C. Chevalley, *Theory of Lie Groups*, Princeton Univ. Press, Princeton, 1946.
2. J. Adams, *Lectures on Lie Groups,* U. of Chicago Press, Chicago, 1969.
3. V. Varadarajan, *Lie Groups and Lie Algebras, and their Representations,* Prentice-Hall, Englewood, NJ, 1974.
4. S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces,* Academic Press, NY, 1978.
5. R. Gilmore, *Lie Groups, Lie Algebras, and Some of Their Applications*, Dover, NY, 2006.
6. P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford University Press, Oxford, 1982 (first edition, 1930).
7. M. E. Rose, *Elementary Theory of Angular Momentum*, Wiley, NY, 1957.

APPENDIX D

THE MECHANICS OF POINTS AND RIGID BODIES

Although it must be presumed of the reader that they are familiar with the usual topics in classical mechanics that are taught to advanced undergraduate and beginning graduate students of physics, nonetheless, since the presentation in this book attempts to use some of the modern methods of mathematical physics, such as differential forms, moving frames, and Lie groups, we shall include an appendix that might serve to apply those most elementary notions to physics at an elementary level. Hence, it is also presently assumed that the basic terminology and methodology of the previous three Appendixes is reasonably familiar, as well.

Furthermore, since some of the basic notions from the geometry of jets and jet manifolds will be employed occasionally, it is also helpful to see that the concept of a "jet" has a natural interpretation in terms of the kinematical state of a moving body, whether point-like, extended, but rigid, or extended and deformable.

The topics in this appendix will be non-relativistic in character, while the relativistic forms of them will be discussed in more detail in the relevant places in the text.

There are two approaches to mechanics in general: One can start with the motion of point-like matter and then build up one's model of the motion of extended matter by treating it as an infinitude of points (e.g., a congruence of curves) or one can treat extended matter as the more realistic manifestation and regard point-like matter as a simplifying approximation that allows one to use finite-dimensional mathematical methods, instead of infinite-dimensional ones. It is probably best to not choose up sides at a time like that, but to regard either extended, continuous matter or point-like matter as an approximation to the other, since – for example – even the atomic ions of a crystal lattice or the molecules of a gas will still resolve to more complex subsystems upon closer inspection, so treating them as lattice points or point masses in space is clearly an approximation in its own right. Hence, in this appendix, we shall discuss point-like matter and its next-most elementary extension to rigid bodies, while discussing continuous matter in the main body of the text.

1. Point mechanics. – In this section, we shall first discuss the kinematics of a point that moves in space, and then the dynamics of a massive point that is subject to the action of forces.

a. The kinematics of moving points. – If a point is assumed to move in a space M, which will be an n-dimensional differentiable manifold that is usually referred to as the *configuration manifold,* then its motion will best be modeled by a "sufficiently-differentiable" curve $x : [t_0, t_1] \to M$, $t \mapsto x(t)$. Here, "sufficiently-differentiable" means that $x(t)$ is as many times continuously-differentiable as is required by the demands of the problem at hand, which we shall call C^k generically. If one does not wish to deal with analytical details, such as the fact that differentiating a C^k function will generally produce a C^{k-1} function, which will belong a larger space of functions that include the C^k functions as a proper subspace, then one can demand that all functions should be smooth; i.e., continuous derivatives will exist for every k.

The first derivative of $x(t)$ with respect to t for each value of t is called the *velocity* of the curve at $x(t)$:

$$\mathbf{v}(t) = \left.\frac{dx}{dt}\right|_t . \tag{4.1}$$

When then tangent vector $\mathbf{v}(t)$ is not zero, it will generate a line $[\mathbf{v}(t)]$ in the tangent space $T_{x(t)}M$ that is, in fact, the tangent line to the curve at that point. When $\mathbf{v}(t) = 0$, one thinks of the point $x(t)$ as a *fixed point* of the curve.

If one changes the parameterization of the curve $x(t)$ by a diffeomorphism onto of the time interval $[t_0, t_1] \to \mathbb{R}$, $t \mapsto \bar{t}$ then, from the chain rule, the new velocity of the curve will be:

$$\bar{\mathbf{v}}(\bar{t}) = \left.\frac{dx}{d\bar{t}}\right|_{\bar{t}} = \left.\frac{dt}{d\bar{t}}\right|_t \mathbf{v}(\bar{t}(t)) . \tag{4.2}$$

Hence, the reparameterized curve will consist of the same points with the same tangent lines, but the magnitudes of the velocity vectors (i.e., speeds) will have been rescaled by the non-zero factor $dt/d\bar{t}$. That means that fixed points will remain fixed, and non-fixed points will remain non-fixed.

One of the subtleties that physics had to confront with the emergence of general relativity was the fact that unless one has some way of comparing tangent vectors at finitely-separated points, such as a unique way of translating them from one point to the other, the actual definition of the acceleration of a C^2 curve in a differentiable manifold M is more debatable than its definition when the curve is in $\mathbb{R}^n$. Of course, one should remember that when $x(t)$ and $x(t + \Delta t)$ are finitely-separated points of M, one cannot actually define $\Delta x(t) = x(t + \Delta t) - x(t)$ without some sort of translation, either, so in order to define acceleration, one might imitate what one did in order to define velocity; viz., look at the expression in a local coordinate chart about a chosen point, and then define an equivalence class of curves at that point.

Hence, let $x(t)$ be the chosen point, while (U, x^i) is a C^k coordinate chart about $x(t)$. By composing $x(t)$ with the coordinate functions x^i, one will get a curve $x^i(t)$ in $\mathbb{R}^n$. The derivative of $x^i(t)$ with respect to t will then be just an elementary exercise in calculus, and it will produce functions $v^i(t)$ that represent the components of $\mathbf{v}(t)$ with respect to the natural frame $\{\partial_i, i = 1, \ldots, n\}$ on $\mathbb{R}^n$:

$$v^i(t) = \frac{dx^i}{dt} . \tag{4.3}$$

As discussed in App. B, one then defines the abstract tangent vector **v** (t) at x (t) to be the equivalence class of all differentiable curves through x (t) that have the same values for $v^i(t)$ in some (and therefore all) coordinate charts about x (t).

One can repeat that process and define the components a^i (t) of the second derivative of x^i (t) with respect to time:

$$a^i(t) = \frac{dv^i}{dt} = \frac{d^2x^i}{dt^2}. \tag{4.4}$$

One can then define the abstract acceleration vector $\mathbf{a}(t)$ to be something in the second tangent space $T_{\mathbf{v}(t)}T_{x(t)}M$ that represents the equivalence class of all curves $\mathbf{v}(t)$ through $\mathbf{v}(t)$ that have the same derivatives in some (and therefore all) coordinate chart about $x(t)$, or better yet, the equivalence class of all curves $x(t)$ through the point $x(t)$ that have the same first and second derivatives, since one then does not have to clarify what one means by the curve $\mathbf{v}(t)$ in the tangent space $T_{x(t)}M$.

If one puts the information about the curve together into a $(3n+1)$-tuple of real numbers $(t, x^i(t), v^i(t), a^i(t))$ then one will see that it can serve as a good definition of the kinematical state of the motion that is defined by the curve at time t, at least, up to second-order. It also coincides with the definition of the *2-jet* j_t^2x of that curve at time t, namely, the equivalence class of all C^2 curves through $x(t)$ that have the same values for $x^i(t)$, $v^i(t)$, and $a^i(t)$ at t in some (and therefore any) coordinate system about $x(t)$.

In fact, what we have really defined is a section of the projection $J^2(\mathbb{R}, M) \to \mathbb{R}$, $j_t^2x \mapsto t$, in which we have defined $J^2(\mathbb{R}, M)$ to be the set of all 2-jets of C^2 curves in M. It is also a differentiable manifold of dimension $3n + 1$, and a typical set of coordinate functions looks like (t, x^i, v^i, a^i). That projection is called the *source projection*, while the projection onto M that takes j_t^2x to x is called the *target projection*, and the projection onto $J^0 = \mathbb{R} \times M$ that takes j_t^2x to (t, x) is called the *contact projection.*

Furthermore, when one has:

$$v^i(t) = \frac{dx^i}{dt}, \qquad a^i(t) = \frac{d^2x^i}{dt^2} \tag{4.5}$$

for all t, one calls the section $s : \mathbb{R} \to J^2(\mathbb{R}, M)$, $t \mapsto (t, x^i(t), v^i(t), a^i(t))$ *integrable.* If one introduces the concept of the *2-jet prolongation* of the curve $x(t)$, namely:

$$j^2x(t) = (t, x^i(t), \dot{x}(t), \ddot{x}(t)), \tag{4.6}$$

in which the dot means differentiation with respect to t, then one can say that the condition for integrability of a section is that:

$$s = j^2x. \tag{4.7}$$

b. The dynamics of moving points. In order to explain the motion of a point that one observes kinematically, one must first attribute some physical properties to the point and the space that it inhabits that will allow one to associate a dynamical state with the kinematical state. At the infinitesimal level, this relationship is simply the one that exists between the elements of the dual V^* to a vector space V and the vectors of the latter space.

The main physical property that one attributes to the point itself is its mass $m(t)$, which is a positive real function of time that is supposed to account for its inertia (which is never defined rigorously), as well as its attraction to other masses in the rest of space. One can then rescale the velocity vector $\mathbf{v}$ by the mass to get the *linear momentum vector field* along the curve $x(t)$ in space M that is followed by the point:

$$\mathbf{p}(t) = m(t)\,\mathbf{v}(t), \tag{4.8}$$

and when one expresses this in terms of the components of the vector fields with respect to some local frame field, one will get:

$$p^i(t) = m(t)\,v^i(t). \tag{4.9}$$

Actually, there are times when one might wish to generalize the transformation that is associated with $m(t)$ from a dilatation to a shear, such as when the mass behaves differently in different directions. An example of this is in the Abraham-Lorentz-Poincaré model of the electron, which had a "transverse mass" in addition to the "longitudinal" one, so one might replace the scalar m with a symmetric mass matrix m whose eigenvalues would then represent the mass in the principal directions. However, we shall refrain from introducing that extra degree of generality here, although the concept of "transverse momentum" will factor crucially in our study of the Dirac electron.

In order to make a dual object out of linear momentum, one can either use the spatial metric g to convert the vector field $\mathbf{p}$ to a linear momentum 1-form:

$$p = i_{\mathbf{p}}\, g \qquad\qquad (p_i = g_{ij}\, p^j) \tag{4.10}$$

or convert the velocity vector field $\mathbf{v}$ to the *covelocity 1-form:*

$$v = i_{\mathbf{v}}\, g \qquad\qquad (v_i = g_{ij}\, v^j) \tag{4.11}$$

and multiply by $m(t)$.

Under the canonical bilinear pairing of linear functionals with vectors that represents the evaluation of a linear functional on a vector, one will get:

$$p\,(\mathbf{v}) = m\,v\,(\mathbf{v}) = m\,v^2, \tag{4.12}$$

which is twice the *kinetic energy* of the moving mass.

One can also define a force 1-form $F(t)$ along the curve $x(t)$, which can either represent forces that that are applied to the mass only along its path, such as impacts and friction, or forces that will be applied no matter where the path goes, such as gravitation and electromagnetism. One finds that the concept of force is dual to that of infinitesimal displacement $\delta\mathbf{x}$, which will be regarded as a vector field along $x(t)$, in that the evaluation of a force one-form F on an infinitesimal – or *virtual* – displacement $\delta\mathbf{x}$ will give the *virtual work* that is performed by F over that displacement:

$$\delta W = F(\delta\mathbf{x}) = F_i\, \delta x^i. \tag{4.13}$$

When one integrates the 1-form F along a curve segment $x(t)$ for $t \in [t_0, t_1]$, the resulting number:

$$W[x(t)] = \int_{x(t)} F = \int_{x(0)}^{x(1)} F_i(t)\, dx^i = \int_{t_0}^{t_1} F_i(t) v^i(t)\, dt \tag{4.14}$$

will be called the *work that is done by F along that curve segment.* The scalar $F(\mathbf{v}) = F_i v^i$ that appears in the integral represents the *power* that is being added to or dissipated by the mass in its motion as a result of its interaction with the environment.

When the force 1-form F is defined over a region of M that includes the trace of $x(t)$, one can imagine how the work done along a curve segment will change when one deforms the curve segment itself. In particular, one can look at curve segments that connect the same two points $x(0)$ and $x(1)$. When two curve segments $x(t)$ and $\bar{x}(t)$ connect the same two points, one can concatenate $x(t)$ with the time-reverse of $\bar{x}(t)$ and obtain a closed loop that is based upon the initial point $x(0)$. If the integral of F around the loop vanishes then its value along the time-reverse of $\bar{x}(t)$ will be minus the value of the integral along $x(t)$; hence, it will be the same on both curves when time goes forward for both. As a result, if that were true for any curves that connect $x(0)$ with $x(1)$ then one could unambiguously associate a number $U(x(1))$ with $x(1)$ that equals the work that is done by F along any curve that connects $x(0)$ to $x(1)$. Clearly, $U(x(0)) = 0$. One finds that the work done around any loop in M vanishes iff $F = - dU$ for some potential function U that is defined where F is defined. The function U is then defined only up to a constant that amounts to choosing the value of $U(x)$ to be zero at some specified point. Such a force is then called *conservative.*

A weaker condition is to demand that the work done by F along any two homotopic ([1]) curves must be the same, which is equivalent to saying that F is closed: $d_\wedge F = 0$. If M is simply-connected then the two conditions are equivalent, but when M is multiply-connected, it is possible for there to be closed loops that are not homotopic to a single path between two points, because there might be a "hole" that gets in the way. For instance, the plane minus a point has that property.

If the kinematical state of a moving point is described by a first-order jet section $s(t) = (t, x(t), \mathbf{v}(t))$ then the *dynamical state* is described by a 1-form on that section:

$$\phi = F_i\, dx^i + p_i\, dv^i\,. \tag{4.15}$$

([1]) That is, there exists a continuous, one-parameter sequence of curve segments that starts with one curve segment and ends with the other.

Since the component functions are functions on $J^1(\mathbb{R}, M)$, their functional dependency can subsume the definition of mechanical constitutive laws:

$$F_i = F_i\,(t, x^i, v^i), \qquad p_i = p_i\,(t, x^i, v^i). \tag{4.16}$$

Note that although F_i can be defined globally on space-time (t, x^i), such as gravitation or electromagnetic forces, it might also be defined only for points on the curve $x^i(t)$, and any v^i, such as viscous drag. Similarly, linear momentum is typically defined to be independent of x^i, but dependent upon v^i, and possibly t, when the mass varies in time. That is, the functional form of F_i can be as it is in (4.16) or $F_i\,(t, v^i)$, while typically, p_i must have the form $p_i\,(t, v^i)$, such as $m(t)\,v^i$. (Of course, this is only true for point matter, in general.)

The evaluation of ϕ on a virtual displacement of the kinematical state:

$$\delta s = \delta x^i(t)\frac{\partial}{\partial x^i} + \delta v^i(t)\frac{\partial}{\partial v^i} \tag{4.17}$$

will give the function of t:

$$\phi\,(\delta s) = F_i\,\delta x^i + p_i\,\delta v^i\,, \tag{4.18}$$

which is the total *virtual work* that is done by the forces and the kinetic sources as a result of the virtual displacement.

In order to get to the equations of motion that are associated with a choice of f, one can either postulate the balance of linear momentum:

$$F = \frac{dp}{dt}\,, \tag{4.19}$$

which is essentially Newton's second law, or first look at what happens to (4.18) when one assumes that the virtual displacement δs is *integrable;* i.e.:

$$\delta s = j^1(\delta x) = \delta x^i(t)\frac{\partial}{\partial x^i} + \frac{d(\delta x^i)}{dt}\frac{\partial}{\partial v^i}\,, \tag{4.20}$$

namely, one will get:

$$\phi\,(\delta s) = (F_i - \dot{p}_i)\,\delta x^i + \frac{d}{dt}(p_i\,\delta x^i)\,, \tag{4.21}$$

after an application of the product rule for differentiation. We shall think of the vector field $j^1(\delta x)$ as the *first prolongation* of the vector field:

$$\delta x\,(t) = \delta x^i(t)\frac{\partial}{\partial x^i}\,. \tag{4.22}$$

One sees that one can get back to (4.19) by assuming that:

$$\phi(\delta s) = 0 \qquad (\text{mod "}d / dt\text{"}) \tag{4.23}$$

for all integrable δs, by which, we mean that $\phi(\delta s)$ differs from 0 by a time derivative. This assumption is best explained in the language of the calculus of variations, which will be discussed in its own chapter. For now, we shall say only that it is essentially another form of d'Alembert's principle.

2. The theory of moments. – Let us first clarify that the moments that we shall be discussing refer to things that usually take the form of cross products of a "radius vector field" with the vectors of translational mechanics (e.g., velocity, acceleration, momentum, force), not to the moments of a distribution, such as the total mass, center-of-mass, moment of inertia, etc., which will be discussed in the context of extended matter.

a. Kinematical moments. – When one has chosen a point O in an affine space A^n (n = 2, 3), one can speak of the *moment* of a kinematical state (t, x, $\mathbf{v}$, $\mathbf{a}$), which will be (t, $\mathbf{r}$, $\mathbf{r} \wedge \mathbf{v}$, $\mathbf{r} \wedge \mathbf{a}$), in which $\mathbf{r} = x - O$ is the displacement vector that takes the origin O to the point x. The bivector:

$$\varpi = \frac{1}{r^2}\,\mathbf{r} \wedge \mathbf{v} \tag{5.1}$$

is called the *orbital angular velocity*, while the bivector:

$$\alpha = \frac{1}{r^2}\,\mathbf{r} \wedge \mathbf{a} \tag{5.2}$$

is the *orbital angular acceleration*.

As long as one has a scalar product on A^n or $\mathbb{R}^n$, one can always associate a bivector or 2-form with a unique element of the Lie algebra that pertains to the choice of scalar product. The association is easiest to explain in terms of component matrices, since if $\{\mathbf{e}_i, i = 1, \ldots, n\}$ is any basis for $\mathbb{R}^n$ then one can express a bivector $\mathbf{B}$ as $\frac{1}{2}B^{ij}\,\mathbf{e}_i \wedge \mathbf{e}_j$. Since the component matrix B^{ij} is anti-symmetric, the matrix $B_i^{\ j} = g_{ik}\,B^{kj}$ will always belong to the Lie algebra of infinitesimal orthogonal transformations for the chosen scalar product. For the Euclidian plane, that Lie algebra will be $\mathfrak{so}(2)$, for Euclidian space, it will be $\mathfrak{so}(3)$, and for Minkowski space, it will be $\mathfrak{so}(3, 1)$.

Hence, ϖ and α can be associated with the matrices:

$$\varpi^i_j = \frac{1}{r^2}(x_j\,v^i - x^i\,v_j), \qquad \alpha^i_j = \frac{1}{r^2}(x_j\,a^i - x^i\,a_j), \tag{5.3}$$

respectively.

A basic property of ϖ^i_j is that:

$$\varpi^i_j x^j = \frac{1}{r^2}(r^2\,v^i - x^i\,v_j\,x^j) = (\mathbf{v} - \langle\mathbf{v}, \hat{\mathbf{r}}\rangle\hat{\mathbf{r}})^i = (\mathbf{v}_\perp)^i\,; \tag{5.4}$$

that is, it gives the components of the projection of **v** in the direction perpendicular to **r**. (We have defined $\hat{\mathbf{r}}$ to be the unit vector that points in the direction of **r**.)

Another is that:

$$\varpi^i_j v^j = \frac{1}{r^2}(x_j v^j v^i - x^i v^2) = \frac{v^2}{r}(<\hat{\mathbf{r}}, \hat{\mathbf{v}}> \hat{\mathbf{v}} - \hat{\mathbf{r}})^i = -\frac{v^2}{r}(\hat{\mathbf{r}}_\perp)^i ; \tag{5.5}$$

that is, it gives the projection of the centripetal acceleration in the direction that is perpendicular to **v**.

In the case of circular motion, for which $v_r = <\mathbf{v}, \hat{\mathbf{r}}> = <\hat{\mathbf{r}}, \hat{\mathbf{v}}> = 0$, $\varpi^i_j x^j$ will reproduce the tangential velocity, while $\varpi^i_j v^j$ will reproduce the centripetal acceleration.

From (5.4), one can reconstruct the velocity vector **v** from:

$$\mathbf{v} = <\mathbf{v}, \hat{\mathbf{r}}>\hat{\mathbf{r}} + \mathbf{v}_\perp = v_r\, \hat{\mathbf{r}} + \varpi\, \mathbf{r} \qquad (v_r \equiv <\mathbf{v}, \hat{\mathbf{r}}>), \tag{5.6}$$

or, in components:

$$v^i = \frac{v_r}{r} x^i + \varpi^i_j x^j = \left(\frac{v_r}{r}\delta^i_j + \varpi^i_j\right) x^j . \tag{5.7}$$

Statements that are analogous to the foregoing ones can be made for the matrix α^i_j, except that velocity gets replaced with acceleration. That is:

$$\alpha^i_j x^j = (\mathbf{a}_\perp)^i, \qquad \mathbf{a} = a_r\, \hat{\mathbf{r}} + \alpha^i_j v^j\, \mathbf{e}_i \qquad (a_r = <\mathbf{a}, \hat{\mathbf{r}}>). \tag{5.8}$$

b. Dynamical moments. – When one takes the moment of linear momentum **p**, one will get the *orbital angular momentum:*

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}. \tag{5.9}$$

In fact, in a lot of older literature, this is referred to as the "moment of momentum."

The moment of the force **F** is referred to as the *torque* by physicists and the "(force) moment" by mechanical engineers:

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}. \tag{5.10}$$

If one takes the moments of both sides of Newton's second law, when it is expressed in the form $\mathbf{F} = d\mathbf{p} / dt$, then one will get:

$$\mathbf{r} \times \mathbf{F} = \mathbf{r} \times \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) - \frac{d\mathbf{r}}{dt} \times \mathbf{p} ,$$

and as long as the linear momentum has the conventional "convective" form ($\mathbf{p} = m\, d\mathbf{r} / dt$), one can put this into the form:

$$\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt} . \tag{5.11}$$

Therefore, Newton's second law for the balance of linear momentum is equivalent to a corresponding law for the balance of orbital angular momentum. Similarly, the law of conservation of linear momentum will become the law of conservation of orbital angular momentum. That is, in the absences of external torques acting upon the mass, its orbital angular momentum (about some reference point O) will remain constant in time.

3. Rotation of a rigid-body about a fixed point in space. – Although the full non-relativistic theory of the motion of rigid bodies involves the group of rigid motions, nonetheless, the only groups that will get represented in the field spaces of the quantum wave functions that are considered in this book will be the groups of rotations and Lorentz transformations. Hence, we shall discuss the non-relativistic motions of a rigid body in space in which only the rotations play a role, which would be when a point of the rigid body is fixed in space.

a. The rigidity constraint. – A rigid body is one step up from point-like matter in the sense that it represents an orthonormal frame at each point of a curve in E^n. It also represents an approximation to extended matter in which one regards the amount of deformation that the material object experiences during its motion as negligible in comparison to the rigid motions of its points. That is equivalent to saying that the distance between any two points of the object stays the same during its motion, so one will be dealing with isometries. As a result of that approximation, any orthonormal frame that is attached to any point of the body will move in the same way as any other orthonormal frame, and the only ambiguity in the definition of a rigid body as an orthonormal frame $\mathbf{e}_i(t)$ along a curve $x(t)$ will be in the choice of some representative point whose motion will be described by $x(t)$ and the choice of a representative frame at each t that will serve to describe the angular attitude of the body. Typically, one reduces a rigid body to an orthonormal frame at its center of mass and chooses that frame either arbitrarily or by using the principal frame of the moment of inertia of the mass distribution about an axis through that center of mass.

We shall now describe the kinematical state of a rigid body in terms of both a moving frame along a curve and a curve in the rotation group.

b. Kinematics of rotations about a fixed point in space. – If the rigid body is associated with an oriented, orthonormal 3-frame $\mathbf{e}_i(0)$ at some point O in three-dimensional Euclidian space then any other oriented, orthonormal 3-frame $\mathbf{e}_i$ at O can be associated with a unique proper rotation in $SO(3)$. In particular, if the rigid body is in a state of continuous motion, so that at each time t, one has an oriented, orthonormal frame $\mathbf{e}_i(t)$ at O, then there will be a unique matrix $R^i_j(t)$ in $SO(3)$ that is defined by:

$$\mathbf{e}_i(t) = \mathbf{e}_j(0)\, R^j_i(t)\,. \tag{6.1}$$

Hence, the motion of the rigid body that is described by $\mathbf{e}_i(t)$ can just as well be described by the continuous curve $R^i_j(t)$ in $SO(3)$.

Once again, when the kinematical state of the motion is referred to a fixed frame, such as $\mathbf{e}_j$ (0), it will be referred to as *inertial*, while it will be *co-moving* when it is referred to the moving frame $\mathbf{e}_i\,(t)$.

If one assumes successive levels of differentiability for the curve $R^i_j(t)$ then one can define inertial and co-moving angular velocities and accelerations by successive differentiation of (6.1) with respect to t and expressing the resulting kinematical object in terms of either $\mathbf{e}_i$ (0) or $\mathbf{e}_i$ (t), resp. The first two differentiations give:

$$\dot{\mathbf{e}}_i = \mathbf{e}_j\,(0)\;\dot{R}^j_i\,, \qquad \ddot{\mathbf{e}}_i = \mathbf{e}_j\,(0)\;\ddot{R}^j_i\,, \tag{6.2}$$

so one can think of the second-order kinematical state of the motion as being described in an inertial frame by either (t, $\mathbf{e}_i\,(t)$, $\dot{\mathbf{e}}_i(t)$, $\ddot{\mathbf{e}}_i(t)$) or (t, $R^j_i(t)$, $\dot{R}^j_i(t)$, $\ddot{R}^j_i(t)$). The former is an integrable section of the source projection $J^2(\mathbb{R}, SO(\mathbb{R}^3)) \to \mathbb{R}$, $(t, \mathbf{e}_i, \dot{\mathbf{e}}_i, \ddot{\mathbf{e}}_i) \mapsto t$, while the latter is an integrable section of the source projection $J^2(\mathbb{R}, SO(3)) \to \mathbb{R}$, $(t, R^j_i, \dot{R}^j_i, \ddot{R}^j_i) \mapsto t$. Here, we are using $SO(\mathbb{R}^3)$ as the notation for the manifold of oriented, orthonormal frame in three-dimensional Euclidian space.

In order to see what the second-order kinematical state is with respect to the moving frame $\mathbf{e}_i\,(t)$, one needs to first substitute:

$$\mathbf{e}_i\,(0) = \mathbf{e}_j\,(t)\,\tilde{R}^j_i(t)\,, \tag{6.3}$$

in which the tilde refers to the matrix inverse, in the first of (6.2) to get:

$$\dot{\mathbf{e}}_i = \mathbf{e}_j\;\omega^j_i\,, \tag{6.4}$$

in which we have defined the *angular velocity* of the moving frame $\mathbf{e}_i\,(t)$ with respect to the inertial frame $\mathbf{e}_i$ (0) to be:

$$\omega^j_i\,(t) = \tilde{R}^j_k\dot{R}^k_i \qquad \text{or} \qquad \omega = R^{-1}\dot{R}\,. \tag{6.5}$$

If one differentiates the basic relationship $R^{-1}(t)\;R(t) = I$ then one will find the useful corollary to the definition that one also has:

$$\omega = -\;\dot{R}^{-1}R\,. \tag{6.6}$$

Furthermore, since ω is an element of the Lie algebra $\mathfrak{so}(3)$, and therefore the infinitesimal generator of a one-parameter subgroup of the Lie group $SO(3)$ that agrees with $R(t)$ at time t, $-\,\omega$ will be the infinitesimal generator of a one-parameter subgroup that agrees with $R^{-1}(t)$ at time t. Hence, one will also have:

$$\omega = \dot{R}R^{-1} = -\;R\dot{R}^{-1}\,. \tag{6.7}$$

In particular, both $R^{-1}\dot{R}$ and $\dot{R}R^{-1}$ will yield the same angular velocity matrix.

Since we have already defined one kind of angular velocity by way of the moment of velocity about a reference point O, we should clarify how the current definition would relate to the previous one. Basically, all that one has to do to make an angular velocity matrix from an orbital angular velocity that is defined by a position vector $\mathbf{r}$ and a velocity vector $\mathbf{v}$ is to note that the unit vectors $\hat{\mathbf{r}}$ and $\hat{\mathbf{v}}_\perp$ define an instantaneous oriented, orthonormal 2-frame $\{\hat{\mathbf{r}}, \hat{\mathbf{v}}_\perp\}$ in the instantaneous plane of motion, which is spanned by $\mathbf{r}$ and $\mathbf{v}$ at time t, unless they are collinear. Hence, one can complete the 2-frame $\{\hat{\mathbf{r}}, \hat{\mathbf{v}}_\perp\}$ at O to an oriented, orthonormal 3-frame $\{\hat{\mathbf{r}}, \hat{\mathbf{v}}_\perp, \mathbf{n}\}$ by setting the *normal* unit vector $\mathbf{n}$ equal to:

$$\mathbf{n} = \hat{\mathbf{r}} \times \hat{\mathbf{v}}_\perp = \frac{1}{v_\perp}\hat{\mathbf{r}}\times\mathbf{v} = \frac{r}{v_\perp}\boldsymbol{\varpi} = \frac{1}{\varpi}\boldsymbol{\varpi}, \tag{6.8}$$

which is then collinear with $\boldsymbol{\omega}$. One might refer to the 3-frame $\{\hat{\mathbf{r}}, \hat{\mathbf{v}}_\perp, \mathbf{n}\}$ as the *tracking frame* for the motion of the point that is described by $\mathbf{r}(t)$. Let us introduce the notation $\{\mathbf{f}_i(t), i = 1, 2, 3\}$ for that frame, with the members identified in the same sequence.

Hence, if one chooses the inertial frame to be the oriented orthonormal 3-frame $\mathbf{f}_i(0)$ that the tracking frame defines at time $t = 0$ then one can regard any of the time-evolutes $\mathbf{f}_i(t)$ of the initial frame as spatial rotations of the initial one, so there will be a corresponding rotation matrix $R_i^j(t)$ at every t such that:

$$\mathbf{f}_i(t) = \mathbf{f}_j(0)\, R_i^j(t). \tag{6.9}$$

That rotation matrix will then allow one to define an angular velocity in the sense that we are currently discussing.

If one wishes to go in the opposite direction – i.e., to start with the angular velocity of a moving frame with respect to an inertial one and define an velocity moment – then one sees that unless one can also specify the time evolution of an $r(t)$ and its derivatives, one can typically only define the motion of a point on a sphere of some chosen, but arbitrary, constant radius r_0. That is, the angular velocity matrix will produce only the tangential component of a velocity when it is applied to a position vector $\mathbf{r}$ that points in the direction of $\mathbf{f}_1(t)$.

If one substitutes (6.3) in the second of (6.2) then one will get:

$$\ddot{\mathbf{e}}_i = \mathbf{e}_j\, \alpha_i^j, \tag{6.10}$$

in which we have introduced the *angular acceleration* of the moving frame with respect to the inertial frame:

$$\alpha_i^j = \tilde{R}_k^j \ddot{R}_i^k. \tag{6.11}$$

If one starts from the definition of ω in (6.5) and differentiates with respect to t then after some straightforward algebra, one can also say that:

$$\alpha = \dot{\omega} + \omega\omega. \tag{6.12}$$

Thus, one once again has that $\alpha \neq \dot{\omega}$, in general.

Hence, the basic second-order kinematical state of the rigid body that rotates about a fixed point in space can be defined by a section s (t) of the source projection $J^2(\mathbb{R}, SO(3)) \to \mathbb{R}$, in the form $(t, I, \omega(t), \alpha(t))$ when one refers it to the moving frame. However, one sees from (6.12) that the relationship between ω and α is not a simple differentiation with respect to time, so the section is no longer integrable. Thus, rotational kinematics gives a very tangible example of what the non-integrability of a section of the aforementioned source projection can imply.

For the purposes of dynamics, it will be sufficient for us to consider first-order kinematical states, which will then take the form $s : \mathbb{R} \to J^1(\mathbb{R}, SO(3))$, $t \mapsto s(t)$, with:

$$s(t) = (t, R(t), \dot{R}(t)) \quad \text{or} \quad (t, I, \omega(t)), \tag{6.13}$$

depending upon whether one describes in the inertial frame or the moving one, resp.

A virtual displacement of the kinematical state s (t) will be a vector field δs (t) on the image of s (t) in $J^1(\mathbb{R}, SO(3))$. Hence, in terms of the natural frame for a local coordinate chart (t, R^i_j, ρ^i_j) on $J^1(\mathbb{R}, SO(3))$, one will have ([1]):

$$\delta s(t) = \delta R^i_j(t)\frac{\partial}{\partial R^i_j} + \delta \dot{R}^i_j(t)\frac{\partial}{\partial \rho^i_j} = \delta I^i_j(t)\,\mathbf{I}^j_i + \delta\omega^i_j(t)\,\boldsymbol{\eta}^j_i\,. \tag{6.14}$$

In order to define the new frame field $\{\mathbf{I}^j_i, \boldsymbol{\eta}^j_i\}$ and the components $\delta I^i_j(t)$, $\delta\omega^i_j(t)$ with respect to it, we start by varying the definition of ω:

$$\delta\omega = \delta(R^{-1}\dot{R}) = \delta R^{-1}\dot{R} + R^{-1}\delta\dot{R} = -\,R^{-1}\delta R\,R^{-1}\dot{R} + R^{-1}\delta\dot{R}\,,$$

which becomes:

$$\delta\omega = -\,R^{-1}\delta R\,\omega + R^{-1}\delta\dot{R}\,. \tag{6.15}$$

With the definition:

$$\delta I = R^{-1}\,\delta R, \tag{6.16}$$

(6.14) will become:

$$\delta s(t) = (R^{-1}\delta R)^i_j\,\mathbf{I}^j_i + (R^{-1}\delta\dot{R} - R^{-1}\delta R\,\omega)^i_j(t)\,\boldsymbol{\eta}^j_i\,,$$

([1]) We shall not consider the variations of the kinematical state that are associated with a reparameterization of the time line.

and after reorganizing this in terms of δR and $\delta\dot{R}$, one will get:

$$\delta s\,(t) = (R^{-1}\delta R)^i_j\,\mathbf{I}^j_i - R^{-1}\delta R\,\omega)^i_j(t)\,\boldsymbol{\eta}^j_i + (R^{-1}\delta\dot{R})^i_j\,\boldsymbol{\eta}^j_i\,,$$

and upon temporarily reverting to the use of the internal indices, one can put this into the form:

$$\delta s\,(t) = \delta R^i_j\,[(\mathbf{I}^j_i - \omega^j_l\,\boldsymbol{\eta}^l_k)\,\tilde{R}^k_j] + \delta\dot{R}^i_j\,(\boldsymbol{\eta}^j_k\tilde{R}^k_i)\,.$$

Upon comparing this with the inertial expression for δs, we can make the identifications:

$$\frac{\partial}{\partial R^i_j} = (\mathbf{I}^j_i - \omega^j_l\,\boldsymbol{\eta}^l_k)\,\tilde{R}^k_j\,, \qquad \frac{\partial}{\partial \rho^i_j} = \boldsymbol{\eta}^j_k\tilde{R}^k_i\,. \tag{6.17}$$

This can be solved for $\mathbf{I}^j_i$ and $\boldsymbol{\eta}^j_k$ directly:

$$\mathbf{I}^j_i = \left(\frac{\partial}{\partial R^i_j} + \omega^j_l\,\frac{\partial}{\partial \rho^k_k}\right)R^i_k\,, \qquad \boldsymbol{\eta}^j_i = \frac{\partial}{\partial \rho^k_j}\,R^i_k\,. \tag{6.18}$$

An essential relationship between dw and dI can be derived by first differentiating $\delta R = R\ \delta I$ with respect to time and then varying $\dot{R} = R\omega$, and then equating the result:

$$\frac{d}{dt}\delta R = \dot{R}\ \delta I + R\ \delta\dot{I}\,, \qquad \delta\dot{R} = \delta R\ \omega + R\ \delta\omega,$$

namely:

$$\delta\omega = \delta\dot{I} + [\omega, \delta I]. \tag{6.19}$$

b. Dynamics – Dynamics is dual to kinematics (at least, infinitesimally) by way of the bilinear pairing that evaluates a force 1-form on a virtual displacement to produce a an increment of virtual work. That duality can also be expressed in terms of the duality between elements of a Lie algebra $\mathfrak{g}$ and its dual vector space $\mathfrak{g}^*$. In the present case, $\mathfrak{g}$ is $\mathfrak{so}(3)$, so $\mathfrak{g}^*$ is $\mathfrak{so}(3)^*$. When elements of both vector spaces are represented by real, antisymmetric 3×3 matrices, one can express that bilinear pairing by means of the trace of the product of the matrices. That is, if the matrix A represents an element of $\mathfrak{so}(3)^*$ and $\mathbf{B}$ represents an element of $\mathfrak{so}(3)$ then the evaluation of A on $\mathbf{B}$ will be:

$$A\,(\mathbf{B}) = \mathrm{Tr}\,A\,\mathbf{B} = A^i_jB^j_i\,. \tag{6.20}$$

If one represents the elements of $\mathfrak{so}(3)$ and its dual as three-component vectors in $\mathbb{R}^3$ (given the vector product) and its dual $\mathbb{R}^{3*}$ then the evaluation of an element $\alpha \in \mathfrak{so}(3)^*$ on an element $\mathbf{b} \in \mathfrak{so}(3)$ is simply:

$$\alpha(\mathbf{b}) = \alpha^i\, b_i\,. \tag{6.21}$$

Now, there are two basic kinematical elements to address: the infinitesimal rotation δR^i_j or $\delta\theta^i_j = \tilde{R}^i_k\, \delta R^k_j$ and the angular velocity $\dot{R}^i_j$ or $\omega^i_j = \tilde{R}^i_k\, \dot{R}^k_j$, in which the first expression relates to the inertial frame, while the second one relates to the moving frame. The corresponding dual objects under the bilinear pairing of virtual work are the torque and angular momentum, respectively.

In inertial form, the fundamental 1-form for rotational dynamics takes the form:

$$\phi = \tau^j_i\, dR^i_j + L^j_i\, d\rho^i_j\,. \tag{6.22}$$

When this 1-form is evaluated on a virtual displacement of the kinematical state:

$$\delta s = \delta R^i_j \frac{\partial}{\partial R^i_j} + \delta\rho^i_j \frac{\partial}{\partial \rho^i_j} \tag{6.23}$$

that will produce:

$$\delta W = \phi(\delta s) = \tau^j_i\, \delta R^i_j + L^j_i\, \delta\rho^i_j\,, \tag{6.24}$$

and if the virtual displacement is integrable [so $\delta\rho^i_j = d(\delta R^i_j)/dt$] then one will have:

$$\delta W = \tau^j_i\, \delta R^i_j + L^j_i \frac{d}{dt}(\delta R^i_j) = (\tau^j_i - \frac{dL^j_i}{dt})\, \delta R^i_j + \frac{d}{dt}(L^j_i\, \delta R^i_j)\,. \tag{6.25}$$

If the variation δR^i_j vanishes at the endpoint of a curve segment then the vanishing of the virtual work for all allowable virtual displacements will imply the equations of motion:

$$\tau^j_i = \frac{dL^j_i}{dt}\,. \tag{6.26}$$

This is the inertial form of Newton's second law for rotational motion; i.e., the balance of angular momentum.

In order to get the non-inertial form, one starts with the fundamental 1-form in the form:

$$\phi = \bar{\tau}^j_i\, I^i_j + \bar{L}^j_i\, \eta^i_j\,, \tag{6.27}$$

in which the coframe field $\{ I^i_j,\ \eta^i_j \}$ is reciprocal to the frame field $\{\mathbf{I}^i_j, \boldsymbol{\eta}^i_j\}$ and uses a virtual displacement in the form :

$$\delta s = \delta I^i_j \, \mathbf{I}^j_i + \delta\omega^i_j \, \boldsymbol{\eta}^j_i \tag{6.28}$$

in order to give:

$$\delta W = \bar{\tau}_i^{\,j} \delta I^i_j + \bar{L}^j_i \, \delta\omega^i_j . \tag{6.29}$$

If the virtual displacement is integrable, in the sense that it obeys (6.19):

$$\delta\omega^i_j = \nabla(\delta I^i_j) = \frac{d}{dt}\delta I^i_j + [\omega, \delta I]^i_j \tag{6.30}$$

then

$$\begin{aligned}
\delta W &= \bar{\tau}_i^{\,j} \delta I^i_j + \bar{L}^j_i \, \nabla(\delta I^i_j) \\
&= \bar{\tau}_i^{\,j} \delta I^i_j + \bar{L}^j_i \left[\frac{d}{dt}(\delta I^i_j) + [\omega, \delta I]^i_j \right] \\
&= (\bar{\tau}_i^{\,j} + [\bar{L}, \omega]^j_i) \, \delta I^i_j + \bar{L}^j_i \, \frac{d}{dt}(\delta I^i_j) \\
&= (\bar{\tau}_i^{\,j} - [\omega, \bar{L}]^j_i - \frac{d\bar{L}^j_i}{dt}) \, \delta I^i_j + \frac{d}{dt}(\bar{L}^j_i \delta I^i_j) ,
\end{aligned}$$

in which the antisymmetry of δI^i_j has been used, along with the product rule for differentiation.

The equations of motion, namely, the balance of angular momentum, will take the form:

$$\bar{\tau}_i^{\,j} = \nabla \bar{L}^j_i = \frac{d\bar{L}^j_i}{dt} + [\omega, \bar{L}]^j_i , \tag{6.31}$$

or, more concisely:

$$\tau = \frac{dL}{dt} + [\omega, L], \tag{6.32}$$

if we drop the overbars.

The corresponding conservation law will be that in the absence of external moments, angular momentum will satisfy:

$$0 = \frac{dL}{dt} + [\omega, L] \tag{6.33}$$

relative to the rotating frame.

If $L = I\,(\omega)$ with I constant in time then (6.32) will give Euler's equation in the form:

$$\tau = I \, \frac{d\omega}{dt} + [\omega, I(\omega)]. \tag{6.34}$$

This equation shows that there are two ways by which a rotational motion of ω can come about:

1. Precession: This would result from an applied external torque τ.

2. Nutation: This would result from a deviation from sphericality in the mass distribution, which would make [ω, I (ω)] non-vanishing. [The spherical case would make I (ω) proportional to ω.]

MULTIPLICATION TABLES FOR THE CLIFFORD ALGEBRA OF MINKOWSKI SPACE

The basis for $\mathcal{C}(4, \eta)$ that will be used to define the multiplication tables for that Clifford algebra (i.e., the structure constants a^A_{BC}, b^A_{BC}, c^A_{BC}) will be the one that was defined in IX.2.*a*:

$$\begin{aligned}
E_0 &= 1,\\
E_{\mu+1} &= \mathbf{e}_\mu \quad (\mu = 0, \ldots, 3),\\
E_{4+i} &= \mathbf{e}_0\,\mathbf{e}_i \ (i = 1, 2, 3),\ E_8 = \mathbf{e}_1\,\mathbf{e}_2,\ E_9 = \mathbf{e}_3\,\mathbf{e}_1,\ E_{10} = \mathbf{e}_2\,\mathbf{e}_3,\\
E_{11} &= \mathbf{e}_0\,\mathbf{e}_1\,\mathbf{e}_2,\ E_{12} = \mathbf{e}_0\,\mathbf{e}_3\,\mathbf{e}_1,\ E_{13} = \mathbf{e}_0\,\mathbf{e}_2\,\mathbf{e}_3,\ E_{14} = \mathbf{e}_1\,\mathbf{e}_2\,\mathbf{e}_3,\\
E_{15} &= \mathbf{e}_0\,\mathbf{e}_1\,\mathbf{e}_2\,\mathbf{e}_3.
\end{aligned}$$

A matrix entry equals the row in which it appears times the column. For brevity, the basis element E_A will be referred to simply by the value of A. The appearance of a + sign implies that the element is the same as its matrix transpose, while a – sign implies that it is the negative of that element.

Table D.1 – Multiplication table for $\mathcal{C}(4, \eta)$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	+	0	5	6	7	2	3	4	13	12	11	8	9	10	15	14
2	+	–	–0	8	9	1	–11	12	–3	4	14	6	–7	–15	–10	13
3	+	–	–	–0	10	11	1	–13	2	14	–4	–5	–15	7	–9	12
4	+	–	–	–	–0	–12	13	1	14	–2	3	–15	5	–6	–8	11
5	+	–	–	+	+	0	–8	9	–6	7	15	3	–4	–14	–13	10
6	+	–	+	–	+	–	0	–10	5	15	–7	–2	–14	4	12	–9
7	+	–	+	+	–	–	–	0	15	–5	6	–14	2	3	–11	8
8	+	+	–	–	+	–	–	+	–0	–10	9	–1	–13	12	–4	–7
9	+	+	–	+	–	–	+	–	–	–0	–8	–13	–1	–11	–3	–6
10	+	+	+	–	–	+	–	–	–	–	–0	12	–11	–1	–2	–5
11	+	+	–	+	+	–	+	+	+	–	–	–0	10	–13	–7	–4
12	+	+	+	–	+	+	–	+	–	+	–	–	–0	8	–6	–3
13	+	+	+	+	–	+	+	–	–	–	+	–	–	–0	–5	–2
14	+	–	+	+	+	–	–	–	+	+	+	–	–	–	0	–1
15	+	–	–	–	–	+	+	+	+	+	+	–	–	–	–	–0

Blank cells in the following tables contain 0 (not be confused with $0 \equiv E_0 = 1$). Since the symmetric product {.,.} and antisymmetric product [.,.] need to be divided by two in the polarization of the algebra product, the table entries will be multiplied by two in those respective cases (e.g., 11 refers to $2E_{11}$).

Table D.2. – Symmetric product for $\mathcal{C}(4,\ \eta)$ (times 2).

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	+	0							13	12	11	8	9	10		
2	+		–0				–11	12			14		–7	–15	–10	
3	+			–0		11		–13		14		–5		7	–9	
4	+				–0	–12	13		14			–15	5		–8	
5	+			+	+	0					15		–4	–14		10
6	+		+		+		0			15		–2		4		–9
7	+		+	+				0	15			–14	2			8
8	+	+			+			+	–0			–1			–4	–7
9	+	+		+			+			–0			–1		–3	–6
10	+	+	+			+					–0			–1	–2	–5
11	+	+		+	+		+	+	+			–0				
12	+	+	+		+	+		+		+			–0			
13	+	+	+	+		+	+				+			–0		
14	+		+	+	+				+	+	+				0	
15	+					+	+	+	+	+	+					–0

Table D.3. – Antisymmetric product for $\mathcal{C}(4,\ \eta)$ (times 2).

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0																
1			5	6	7	2	3	4							15	14
2		–		8	9	1			–3	4		6				13
3		–	–		10		1		2		–4		–15			12
4		–	–	–				1		–2	3			–6		11
5		–	–				–8	9	–6	7		3			–13	
6		–		–		–		–10	5		–7		–14		12	
7		–			–	–	–			–5	6			3	–11	
8			–	–		–	–			–10	9		–13	12		
9			–		–	–		–	–		–8	–13		–11		
10				–	–		–	–	–	–		12	–11			
11			–			–				–	–		10	–13	–7	–4
12				–			–		–		–	–		8	–6	–3
13					–			–	–	–		–	–		–5	–2
14		–				–	–	–				–	–	–		–1
15		–	–	–	–							–	–	–	–	