Geometric Algebra for Special and General Relativity

by

Joseph Wilson

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Master of Science in Mathematics

Victoria University of Wellington
February 18, 2022

Abstract

This thesis is an inquiry into *geometric algebra* for the study of relativistic physics. It is divided into two parts: the first is on geometric algebra abstractly and its application special relativity; the second concerns the extension to general relativity and curved spacetime.

{TO DO: ...}

Contents

Part I.

Special Relativity and Geometric Algebra

Chapter 1.

Introduction

{TO DO: Intro sucks.}

The Special Theory of Relativity is a model of *spacetime* — the geometry in which physical events take place. Spacetime comprises the Euclidean dimensions of space and time, but only in a way relative to each observer moving through it: there exists no single 'universal' ruler or clock. Instead, two observers in relative motion find their respective clocks and rulers are found to disagree, according to the Lorentz transformation laws. The insight of special relativity is that one should focus not on the observer-dependent notions of space and time, but on the Lorentzian geometry of spacetime itself.

Seven years after Albert Einstein introduced this theory,[1] he succeeded in formulating a relativistic picture which included gravity. In this General Theory of Relativity, gravitation is identified with the curvature of spacetime over astronomical distances. Both theories coincide locally (i.e., when confined to sufficiently small extents of spacetime, over which the effects of curvature are negligible). In part I, we will focus on special relativity, leaving considerations of curvature to part II.

The study of local spacetime geometry amounts to the study of its intrinsic symmetries.[2] These symmetries form the Poincaré group, and consist of spacetime translations and Lorentz transformations, the latter being the extension of the rotation group for Euclidean space to relativistic rotations of spacetime. The standard matrix representation of the Lorentz group, $\mathrm{SO}^+(1,3)$, is the connected component of the orthog-

[1] Einstein's paper [1] was published in 1905, the so-called *Annus Mirabilis* or "miracle year" during which he also published on the photoelectric effect, Brownian motion and the mass-energy equivalence. Each of the four papers was a monumental contribution to modern physics.

[2] This insight is part of Felix Klein's Erlangen programme of 1872 [2], wherein geometries (Euclidean, hyperbolic, projective, etc.) are studied in terms of their symmetry groups and their invariants.

onal group

$$O(1,3) = \left\{ \Lambda \in GL(\mathbb{R}^4) \,\middle|\, \Lambda^{\mathsf{T}} \eta \Lambda = \eta \right\}$$

with respect to the bilinear form $\eta = \pm\mathrm{diag}(-1, +1, +1, +1)$. The rudimentary tools of matrix algebra are sufficient for an analysis the Lorentz group, and are familiar to any physicist. However, they are not always the most suitable tool available for problems of relativity.

The last century has seen many other mathematical objects be applied to the study of generalised rotation groups such as $SO^+(1,3)$ or the $\mathbb{R}^3$ rotation group SO(3). Among these tools is the *geometric algebra*, invented[3] by William Clifford in 1878 [4] building upon the work of Hamilton and Grassmann, which constitute the main theme of this thesis.

[3] Clifford algebra (an alias) was independently discovered by Rudolf Lipschitz two years later [3]. Lipschitz was the first to use them to the study the orthogonal groups.

Geometric algebra remains largely unknown in the physics community, despite arguably being far superior for algebraic descriptions of rotations than traditional matrix techniques. To appreciate this, we ought to glean the history that led to the relative obscurity of Clifford algebras.

The quest for an optimal formalism for rotations

Mathematics has seen the invention of a variety of vector formalisms since the 1800s, and the question of which is best suited to physics has a long and contentious history. Complex numbers had been long known[4] to be useful descriptions of planar rotations. William Hamilton's efforts to extend the same ideas into three dimensions by inventing a "multiplication of triples" bore fruition in 1843 when the quaternion algebra $\mathbb{H}$, defined by

[4] Since Wessel, Argand and Gauss in the 1700s [5].

$$\hat{\boldsymbol{\imath}}^2 = \hat{\boldsymbol{\jmath}}^2 = \hat{\boldsymbol{k}}^2 = \hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}} = -1,$$

famously came to him in revelation. In following decades, William Gibbs developed the competing vector calculus of $\mathbb{R}^3$ with the usual vector cross and dot products. The ensuing vector algebra "war" of 1890–1945 saw Hamilton's prized[5] quaternion algebra pitted against Gibbs' easier-to-visualise vector calculus, with Gibbs' calculus eventually dominating

[5] Hamilton had a dedicated following in the time: the *Quaternion Society* existed from 1895 to 1913.

because of their relatively easier learning curve. Today, quaternions are generally regarded as an old-fashioned mathematical curiosity.

Despite this, various authors, in appreciating quaternions' elegant handling of $\mathbb{R}^3$ rotations, have tried coercing them into Minkowski space $\mathbb{R}^{1,3}$ for application to special relativity [6–8]. This has been done in various ways, usually by complexifying $\mathbb{H}$ into an eight-dimensional algebra $\mathbb{C} \otimes \mathbb{H}$ and then restricting the number of degrees of freedom as seen fit [9, 10]. However, it is fair to say that quaternionic formulations of special relativity never gained notable traction.

The superior vector formalism for physics

Today, relativists are most familiar with tensor calculus, differential forms and the Dirac γ-matrix formalism, and have relatively little to do with quaternions or derived algebras.[6] Arguably, this outcome of history is unfortunate: matrix descriptions of rotations cannot match the efficiency of quaternions, yet quaternions remain 'peculiar' to many and are intrinsically tied to three dimensions.

[6] See [5, 11] for more historical discussion of quaternions and their adoption in physics.

In this respect, geometric algebra is a perfect middle-ground. Its rotor formulation of rotations is algebraically efficient like the quaternions, but is not specific to $\mathbb{R}^3$ — indeed, geometric algebra is general to any dimension or metric signature. Furthermore, objects like vectors, bivectors and k-vectors (familiar from exterior differential calculus) are first-class objects in the geometric algebra, yet obey identical rotor transformation laws. Unlike exterior calculus, multivectors are often invertible, making algebraic manipulation easy.

In quantum theory, Dirac's γ-matrix formalism is simply a matrix representation of a geometric algebra (see section 3.2.4). Although some physicists come away from quantum theory with the impression that Clifford algebra is something *inherently quantum*, this is a misconception: geometric algebra is applicable to vast areas of geometry and physics, classical and quantum, and from elementary levels.[7]

[7] See [12] for discussion of diverse applications of geometric algebra.

Part I of this thesis introduces geometric algebra with emphasis on its relation to other common formalisms in physics. The principle focus is then on its applications to special relativity, where Lorentz transformations are described as rotors in the geometric algebra. In chapter 5, this leads to a novel technique for composing Lorentz transformations in terms of rotor generators, also explicated in [13]. Later, the focus is on calculus and the extension of the ideas of part I to manifolds and general relativity in part II. {TO DO: Summarise part II better.}

Chapter 2.

Preliminary Theory

Many of the tools we will develop take place in various associative algebras. As well as the geometric algebra of spacetime, we will encounter tensors, exterior forms, quaternions, and other structures in this category. Instead of defining each algebra axiomatically as needed, it is easier to develop the general theory of associative algebras and then define each special case. This enables the use of the same tools and the same terminology thoughout.

Therefore, this section is an overview of the abstract theory of associative algebras, which more generally belongs to *ring theory*.[8] *Algebras*, *quotients* and *gradings* are defined, as well as *tensors*, *multivectors* and *exterior forms*. Most definitions in this chapter can be readily generalised by replacing the field $\mathbb{F}$ with a ring. The excitable reader may skip this chapter and refer back as needed.

[8] A RING is a field without the requiring commutativity nor existence of multiplicative inverses.

2.1. Associative Algebras

Throughout, $\mathbb{F}$ denotes the underlying field of some vector space. (Eventually, $\mathbb{F}$ will always be taken to be $\mathbb{R}$, but we may begin in generality.)

Definition 1. *An* ASSOCIATIVE ALGEBRA A *is a vector space equipped with a product* $\circledast : A \times A \to A$ *which is associative and bilinear.*

Associativity means $(\boldsymbol{u} \circledast \boldsymbol{v}) \circledast \boldsymbol{w} = \boldsymbol{u} \circledast (\boldsymbol{v} \circledast \boldsymbol{w})$, while bilinearity means the product is:

- compatible with scalars: $(\lambda \boldsymbol{u}) \circledast \boldsymbol{v} = \boldsymbol{u} \circledast (\lambda \boldsymbol{v}) = \lambda(\boldsymbol{u} \circledast \boldsymbol{v})$ for $\lambda \in \mathbb{F}$; and
- distributive over addition: $(\boldsymbol{u}+\boldsymbol{v}) \circledast \boldsymbol{w} = \boldsymbol{u} \circledast \boldsymbol{w} + \boldsymbol{v} \circledast \boldsymbol{w}$, and similarly for $\boldsymbol{u} \circledast (\boldsymbol{v} + \boldsymbol{w})$.

This definition can be generalised by relaxing associativity or by letting $\mathbb{F}$ be a ring. However, we will use "algebra" exclusively to mean an associative algebra over a field (usually $\mathbb{R}$).

Examples. Any ring forms an associative algebra when considered as a one-dimensional vector space. The complex numbers can be viewed as a real 2-dimensional algebra by defining $\circledast$ to be complex multiplication; $(x_1, y_1) \circledast (x_2, y_2) \coloneqq (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2)$.

The free tensor algebra

The most general (associative) algebra containing a given vector space V is the TENSOR ALGEBRA $V^{\otimes}$. The tensor product $\otimes$ satisfies exactly the relations of definition 1 with no others. Thus, the tensor algebra associative, bilinear and *free* in the sense that no further information is required in its definition.

As a vector space, the tensor algebra is equal to the infinite direct sum

$$V^{\otimes} \cong \bigoplus_{k=0}^{\infty} V^{\otimes k} \equiv \mathbb{F} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots \tag{2.1}$$

where each $V^{\otimes k}$ is the subspace of TENSORS OF GRADE k.

2.1.1. Quotient algebras

Owing to the maximal generality of the free tensor algebra, any other associative algebras may be constructed as a *quotient* of $V^{\otimes}$. In order for a quotient $V^{\otimes}/\sim$ to itself form an algebra, the equivalence relation $\sim$ must preserve the associative algebra structure.

Definition 2. *A* CONGRUENCE *on an algebra A is an equivalence relation*

$\sim$ which is compatible with the algebraic relations, so that if $a \sim a'$ and $b \sim b'$ then $a + b \sim a' + b'$ and $a \circledast b \sim a' \circledast b'$.

The quotient of an algebra by a congruence naturally has the structure of an algebra, and so is called a QUOTIENT ALGEBRA.

Lemma 1. *The QUOTIENT $A/\sim$ of an algebra A by a congruence $\sim$, consisting of equivalence classes $[a] \in A/\sim$ as elements, forms an algebra with the naturally inherited operations $[a]+[b] \coloneqq [a+b]$ and $[a]\circledast[b] \coloneqq [a\circledast b]$.*

Proof. The fact that the operations $+$ and $\circledast$ of the quotient are well-defined follows from the structure-preserving properties of the congruence. Addition is well-defined if $[a] + [b]$ does not depend on the choice of representatives: if $a' \in [a]$ then $[a'] + [b]$ should be $[a] + [b]$. By congruence, we have from $a \sim a'$ so that $[a + b] = [a' + b]$ and indeed $[a] + [b] = [a'] + [b]$. Likewise for $\circledast$. □

Instead of presenting an equivalence relation, it is often easier to define a congruence by specifying the set of elements which are equivalent to zero, from which all other equivalences follow from the algebra axioms. Such a set of all 'zeroed' elements is called an ideal.

Definition 3. *A (TWO-SIDED) IDEAL of an algebra A is a subset $I \subseteq A$ which is closed under addition and invariant under multiplication, so that*

- *if $a, b \in I$ then $a + b \in I$; and*
- *if $r \in A$ and $a \in I$ then $r \circledast a \in I \ni a \circledast r$.*

We will use the notation $\{\!\{A\}\!\}$ to mean the ideal generated by setting $a \sim 0$ for all a in A. For example, $\{\!\{a\}\!\} = \operatorname{span}\{r \circledast a \circledast r' \mid r, r' \in A\}$ is the ideal consisting of sums and products involving the specified element a, and $\{\!\{\boldsymbol{u} \otimes \boldsymbol{u} \mid \boldsymbol{u} \in V\}\!\}$, or simply $\{\!\{\boldsymbol{u} \otimes \boldsymbol{u}\}\!\}$, is the ideal in $V^{\otimes}$ consisting of sums of terms of the form $a \otimes \boldsymbol{u} \otimes \boldsymbol{u} \otimes b$ for vectors $\boldsymbol{u}$ and arbitrary $a, b \in V^{\otimes}$.

Lemma 2. *An ideal uniquely defines a congruence, and vice versa, by the identification of I as the set of zero elements, $a \in I \iff a \sim 0$.*

Proof. The set $I \coloneqq \{a \mid a \sim 0\}$ is indeed an ideal because it is closed under addition (for $a, b \in I$ we have $\implies a + b \sim 0 + 0 = 0$ so $a + b \in I$) and invariant under multiplication (for any $a \in I$ and $r \in A$, we have $r \circledast a \sim r \circledast 0 = 0 = 0 \circledast r \sim a \circledast r$). Conversely, let $a \sim a'$ and $b \sim b'$. Since $\sim$ respects addition:

$$\left.\begin{aligned} a - a' \in I \\ b - b' \in I \end{aligned}\right\} \implies (a + b) - (a' + b') \in I \iff a + b \sim a' + b',$$

and multiplication:

$$\left.\begin{aligned} (a - a') \circledast b \in I \\ a' \circledast (b - b') \in I \end{aligned}\right\} \implies a \circledast b - a' \circledast b' \in I \iff a \circledast b \sim a' \circledast b',$$

the equivalence defined by $a \sim b \iff a - b \in I$ is a congruence. □

The equivalence of ideals and congruences is a general feature of abstract algebra.[9] Furthermore, both can be given in terms of a homomorphism between algebras,[10] and this is often the most convenient way to define a quotient.

[9] E.g., in group theory, ideals are *normal subgroups* and define congruences, which are equivalence relations satisfying $gag^{-1} \sim \mathrm{id}$ whenever $a \sim \mathrm{id}$.

[10] A *homomorphism* is a structure-preserving map; in the case of algebras, a linear map $\Psi : A \to A'$ which satisfies $\Psi(a \circledast b) = \Psi(a) \circledast' \Psi(b)$.

Theorem 1 (first isomorphism theorem). *If $\Psi : A \to B$ is a homomorphism, between algebras, then*

1. *the relation $a \sim b$ defined by $\Psi(a) = \Psi(b)$ is a congruence;*
2. *the kernel $I \coloneqq \ker \Psi$ is an ideal; and*
3. *the quotients $A/{\sim} \equiv A/I \cong \Psi(A)$ are all isomorphic.*

Proof. We assume A and B associative algebras. (For a proof in universal algebra, see [14, §15].)

To verify item 1, suppose that $\Psi(a) = \Psi(a')$ and $\Psi(b) = \Psi(b')$ and note that $\Psi(a + a') = \Psi(b + b')$ by linearity and $\Psi(a \circledast b) = \Psi(a' \circledast b')$ from

$\Psi(a \circledast b) = \Psi(a) \circledast \Psi(b)$, so the congruence properties of definition 2 are satisfied.

For item 2, note that $\ker \Psi$ is a vector subspace, and that $a \in \ker \Psi$ implies $a \circledast r \in \ker \Psi$ for any $r \in A$ since $\Psi(a \circledast r) = \Psi(a) \circledast \Psi(r) = 0$. Thus, $\ker \Psi$ is an ideal by definition 3.

The first equivalence in item 3 follows from lemma 2. For an isomorphism $\Phi : A/\ker \Psi \to \Omega(A)$, pick $\Phi([a]) = \Psi(a)$. This is well-defined because the choice of representative of the equivalence class $[a]$ does not matter; $a \sim a'$ if and only if $\Psi(a) = \Psi(a')$ by definition of $\sim$, which simultaneously shows that Φ is injective. Surjectivity follows since any element of $\Psi(A)$ is of the form $\Psi(a)$ which is the image of $[a]$. □

With the free tensor algebra and theorem 1 in hand, we are able to describe any associative algebra as a quotient of the form $V^{\otimes}/I$.

Definition 4. *The* DIMENSION $\dim A$ *of a quotient algebra* $A = V^{\otimes}/I$ *is its dimension as a vector space. The* BASE DIMENSION *of* A *is the dimension of the underlying vector space* V*.*

Algebras of finite base dimension may be infinite-dimensional, as is the case for the tensor algebra itself (which is a quotient by the trivial ideal).

2.1.2. Graded algebras

Associative algebras may possess another layer of useful structure: a grading. An example grading for the tensor algebra has already been exhibited in eq. (2.1). Gradings generalise the *degree* or *rank* of tensors or forms, and the notion of *parity* (even/oddness) for functions or polynomials.

Informally, an algebra's grading provides a labelling for some of its elements, such that labels are combined simply (usually by addition) under the algebra's multiplication.

Definition 5. *An algebra A is R-GRADED for $(R, +)$ a monoid*[11] *if there exists a decomposition*

$$A = \bigoplus_{k \in R} A_k$$

such that $A_i \circledast A_j \subseteq A_{i+j}$, i.e., $a \in A_i, b \in A_j \implies a \circledast b \in A_{i+j}$.

[11] A MONOID is a group without the requirement of inverses; i.e., a set with an associative binary operation for which there is an identity element.

The monoid is usually taken to additive over $\mathbb{N}$ or $\mathbb{Z}$, possibly modulo some integer. For instance, the tensor algebra $V^{\otimes}$ is $\mathbb{N}$-graded, since if $a \in V^{\otimes p}$ and $b \in V^{\otimes q}$ then $a \otimes b \in V^{\otimes p+q}$. Indeed, $V^{\otimes}$ is also $\mathbb{Z}$-graded if for $k < 0$ we understand $V^{\otimes k} := \{\mathbf{0}\}$ to be the trivial vector space. The tensor algebra is also $\mathbb{Z}_p$-graded, where $\mathbb{Z}_p \equiv \mathbb{Z}/p\mathbb{Z}$ is addition modulo any $p > 0$, since the decomposition

$$V^{\otimes} = \bigoplus_{k=0}^{p-1} Z_k \quad \text{where} \quad Z_k = \bigoplus_{n=0}^{\infty} V^{\otimes k+np} = V^{\otimes k} \oplus V^{\otimes (k+p)} \oplus \cdots$$

satisfies $Z_i \otimes Z_j \subseteq Z_k$ when $k \equiv i + j \mod p$. In particular, $V^{\otimes}$ is $\mathbb{Z}_2$-graded,[12] and its elements admit a notion of *parity*: elements of $Z_0 = \mathbb{F} \otimes V^{\otimes 2} \otimes \cdots$ are even, while elements of $Z_1 = V \otimes V^{\otimes 3} \otimes \cdots$ are odd, with parity is respected by $\otimes$ as it is for the integers.

[12] Algebras which are $\mathbb{Z}_2$-graded are sometimes called *superalgebras*, with the prefix ‘super-’ originating from supersymmetry theory.

Importantly, just as not all functions $f : \mathbb{R} \to \mathbb{R}$ are even or odd, not all elements of a $\mathbb{Z}_2$-graded algebra are even or odd. More generally not all elements of a graded algebra belong to a single graded subspace.

Definition 6. *If $A = \bigoplus_{k \in R} A_k$ is an R-graded algebra, then an element $a \in A$ is HOMOGENEOUS if it belongs to some A_k, in which case it is said to be a k-VECTOR.*[13] *If $a \in A_{k_1} \oplus \cdots \oplus A_{k_n}$ is inhomogeneous, we may call it a $\{k_1, \ldots, k_n\}$-multivector.*

[13] We use the aliases *bivector* for 2-vector, *trivector* for 3-vector, *scalar* for 0-vector, etc.

All elements of a graded algebra are either inhomogeneous or a k-vector for some k, and every non-zero k-vector is a sum of one or more irreducible terms, called *blades*, in the following sense.

Definition 7. *A k-BLADE is a k-vector $a \in A_k$ of the form $a = \boldsymbol{u}_1 \circledast \cdots \circledast \boldsymbol{u}_k$ where each $\boldsymbol{u}_i \in A_1$ is a 1-vector.*

Note that not all k-vectors are blades. For example, in the $\mathbb{Z}$-graded tensor algebra, the bivector $\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 + \boldsymbol{e}_3 \otimes \boldsymbol{e}_4 \in \mathbb{R}^4 \otimes \mathbb{R}^4$ where $\{\boldsymbol{e}_i\}$ is the standard basis of $\mathbb{R}^4$, cannot be factored into a blade of the form $\boldsymbol{u} \otimes \boldsymbol{v}$ for any $\boldsymbol{u}, \boldsymbol{v} \in V$.

{TO DO: Does this even make sense for a general graded algebra??}

Graded quotient algebras

A grading structure may or may not be inherited by a quotient — in particular, not all quotients of $V^{\otimes}$ inherit its $\mathbb{Z}$-grading. When reasoning about quotients of graded algebras, the following fact is useful.

Lemma 3. *Quotients commute with direct sums, so if*

$$A = \bigoplus_{k \in R} A_k \quad \textit{and} \quad I = \bigoplus_{k \in R} I_k \quad \textit{then} \quad A/I = \bigoplus_{k \in R} (A_k/I_k)$$

where R is some index set.

Proof. It is sufficient to prove the case for direct sums of length two. We then seek an isomorphism $\Phi : (A \oplus B)/(I \oplus J) \to (A/I) \oplus (B/J)$. Elements of the domain are equivalence classes of pairs $[(a, b)]$ with respect to the ideal $I \oplus J$. The direct sum ideal $I \oplus J$ corresponds to the congruence defined by $(a, b) \sim (a', b') \iff a \sim a'$ and $b \sim b'$. Therefore, the assignment $\Phi = [(a, b)] \mapsto ([a], [b])$ is well-defined. Injectivity and surjectivity follow immediately. □

The general non-preservation of gradings motivates strengthening the notion of an ideal:

Definition 8. *An ideal I of an R-graded algebra $A = \bigoplus_{k \in R} A_k$ is* HOMOGENEOUS *if $I = \bigoplus_{k \in R} I_k$ where $I_k = I \cap A_k$.*

Not all ideals are homogeneous.[14] The additional requirement that an ideal be homogeneous ensures that the associated equivalence relation,

[14] For example, the ideal $I = \{\!\{\boldsymbol{e}_1 + \boldsymbol{e}_2 \otimes \boldsymbol{e}_3\}\!\}$ is distinct from $\bigoplus_{k=0}^{\infty} (I \cap V^{\otimes k}) = \{\!\{\boldsymbol{e}_1, \boldsymbol{e}_2 \otimes \boldsymbol{e}_3\}\!\}$ because the former does not contain $\operatorname{span}\{\boldsymbol{e}_1\}$, while the latter does.

as well as respecting the basic algebraic relations of definition 2, also preserves the grading structure. And so, we have a 'graded' analogue to lemma 1:

Theorem 2. *If A is an R-graded algebra and I a homogeneous ideal, then the quotient A/I is also R-graded.*

Proof. By lemma 3 and the homogeneity of I, we have $A/I = \bigoplus_{k\in R}(A_k/I_k)$. Elements of A_k/I_k are equivalence classes $[a_k]$ where the representative is of grade k. Thus, $(A_p/I_p) \circledast (A_q/I_q) \subseteq A_{p+q}/I_{p+q}$ since $[a_p] \circledast [a_q] = [a_p \circledast a_q] = [b]$ for some $b \in A_{p+q}$. Hence, A/I is R-graded. □

2.2. The Wedge Product: Multivectors

Perhaps the simplest (yet most useful) nontrivial quotient of the tensor algebra is the *exterior algebra*, first popularised by Hermann Grassmann in 1844.

Definition 9. *The* EXTERIOR ALGEBRA *over a vector space V is*

$$\textstyle\bigwedge V := V^{\otimes} \big/ \{\!\{\boldsymbol{u} \otimes \boldsymbol{u}\}\!\} .$$

The product in $\bigwedge V$ is called the WEDGE PRODUCT*, denoted $\wedge$.*

$\{\!\{\boldsymbol{u} \otimes \boldsymbol{u}\}\!\} \equiv \{\!\{\boldsymbol{u} \otimes \boldsymbol{u} \mid \boldsymbol{u} \in V\}\!\}$ *is the ideal defined by $\boldsymbol{u} \otimes \boldsymbol{u} \sim 0$ for any vectors $\boldsymbol{u} \in V$.*

The wedge product is also called the *exterior*, *alternating* or *antisymmetric* product. The property suggested by its various names is easily seen by expanding the square of a sum:

$$(\boldsymbol{u} + \boldsymbol{v}) \wedge (\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{u} \wedge \boldsymbol{u} + \boldsymbol{u} \wedge \boldsymbol{v} + \boldsymbol{v} \wedge \boldsymbol{u} + \boldsymbol{v} \wedge \boldsymbol{v}.$$

Since all terms of the form $\boldsymbol{w} \wedge \boldsymbol{w} = 0$ are definitionally zero, we have

$$\boldsymbol{u} \wedge \boldsymbol{v} = -\boldsymbol{v} \wedge \boldsymbol{u}$$

for all vectors $\boldsymbol{u}, \boldsymbol{v} \in V$. By associativity, it follows that $\boldsymbol{v}_1 \wedge \boldsymbol{v}_2 \wedge \cdots \wedge \boldsymbol{v}_k$ vanishes exactly when the $\boldsymbol{v}_i$ are linearly dependent.[15]

15

Proof. Blades of the form $a = \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_k$ vanish when two or more vectors are repeated. If $\{\boldsymbol{u}_i\}$ is linearly dependent, then any one $\boldsymbol{u}_i$ can be written in terms of the others, and

The ideal $\{\!\{\boldsymbol{u} \otimes \boldsymbol{u}\}\!\}$ is homogeneous with respect to the $\mathbb{Z}$-grading of the parent tensor algebra,[16] and hence $\bigwedge V$ is itself $\mathbb{Z}$-graded (by theorem 2). In particular, the decomposition into fixed-grade subspaces

$$\bigwedge V = \bigoplus_{k=0}^{\dim V} \bigwedge^k V \quad \text{where} \quad \bigwedge^k V = \operatorname{span}\{\boldsymbol{v}_1 \wedge \boldsymbol{v}_2 \wedge \cdots \wedge \boldsymbol{v}_k \mid \boldsymbol{v}_i \in V\},$$

is respected by the wedge product, i.e., $(\bigwedge^p V) \wedge (\bigwedge^q V) \subseteq \bigwedge^{p+q} V$. Definitions 6 and 7 carry over directly into $\bigwedge V$, so elements of $\bigwedge^k V$ are k-vectors, and elements of the form $\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_k$ are k-blades.

[16] This follows because $\{\!\{\boldsymbol{u} \otimes \boldsymbol{u}\}\!\}$ is generated by grade 2 elements $\boldsymbol{u} \otimes \boldsymbol{u} \in V^{\otimes 2}$.

By counting the number of possible linearly independent sets of k vectors in $\dim V$ dimensions, it follows that in base dimension $\dim V = n$,

$$\dim \bigwedge^k V = \binom{n}{k}, \quad \text{and hence} \quad \dim \bigwedge V = 2^n.$$

In particular, note that $\dim \bigwedge^k V = \dim \bigwedge^{n-k} V$. Elements of the one-dimensional subspace $\bigwedge^n V$ are called PSEUDOSCALARS.[17]

[17] The prefix ‘pseudo’ means $k \mapsto n-k$. Hence, a pseudovector is an $(n-1)$-vector, etc.

Blades have direct geometric interpretations. The bivector $\boldsymbol{u} \wedge \boldsymbol{v}$ is interpreted as the directed planar area spanned by the parallelogram with sides $\boldsymbol{u}$ and $\boldsymbol{v}$. (Note that blades have no ‘shape’; only directed magnitude.) Similarly, higher-grade elements represent directed (hyper)volume elements spanned by parallelepipeds (see fig. 2.1). In fact, any k-blade may be viewed as a subspace of V with an oriented scalar magnitude:

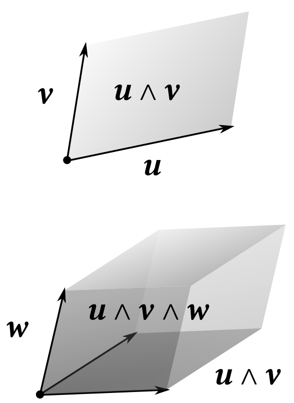


Figure 2.1.: Bivectors and trivectors have orientations induced by the order of the wedge product.

Definition 10. *The* SPAN *of a non-zero k-blade $b = \boldsymbol{u}_i \wedge \cdots \wedge \boldsymbol{u}_i$ is the k-dimensional subspace $\operatorname{span}\{b\} = \operatorname{span}\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$. Define the span of zero to be the trivial subspace.*

Notably, a blade’s span is independent of the particular $\wedge$-decomposition of the blade into vectors. (E.g., if $\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_k = \boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{v}_k$ are two such decompositions, then $\operatorname{span}\{\boldsymbol{u}_i\} = \operatorname{span}\{\boldsymbol{v}_i\}$.)

2.2.1. As antisymmetric tensors

The exterior algebra may equivalently be viewed as the space of antisymmetric tensors equipped with an antisymmetrising product. Consider the map

$$\mathrm{Sym}^{\pm}(\boldsymbol{u}_1 \otimes \cdots \otimes \boldsymbol{u}_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (\pm 1)^{\sigma} \boldsymbol{u}_{\sigma(1)} \otimes \cdots \otimes \boldsymbol{u}_{\sigma(k)} \tag{2.2}$$

where $(-1)^{\sigma}$ denotes the sign of the permutation σ in the symmetric group of k elements, S_k. By requiring linearly, $\mathrm{Sym}^{\pm} : V^{\otimes} \to V^{\otimes}$ is defined on all tensors. A tensor A is called SYMMETRIC if $\mathrm{Sym}^{+}(A) = A$ and ANTISYMMETRIC if $\mathrm{Sym}^{-}(A) = A$.

Denote the image $\mathrm{Sym}^{-}(V^{\otimes})$ by S. The linear map $\mathrm{Sym}^{-} : V^{\otimes} \to S$ is *not* an algebra homomorphism with respect to the tensor product on S, since, e.g.,

$$\mathrm{Sym}^{-}(\boldsymbol{u} \otimes \boldsymbol{v}) = \frac{1}{2}(\boldsymbol{u} \otimes \boldsymbol{v} - \boldsymbol{v} \otimes \boldsymbol{u}) \neq \boldsymbol{u} \otimes \boldsymbol{v} = \mathrm{Sym}^{-}(\boldsymbol{u}) \otimes \mathrm{Sym}^{-}(\boldsymbol{v}).$$

However, Sym^{-} *is* a homomorphism if we instead equip $S \equiv (S, \wedge)$ with the antisymmetrising product $\wedge : S \times S \to S$ defined by

$$A \wedge B \coloneqq \mathrm{Sym}^{-}(A \otimes B). \tag{2.3}$$

With this algebra homomorphism, by theorem 1 we have

$$S \cong V^{\otimes} \big/ \ker \mathrm{Sym}^{-} . \tag{2.4}$$

Furthermore, note that the kernel of Sym^{-} consists of tensor products of linearly dependent vectors, and sums thereof,[18]

$$\ker \mathrm{Sym}^{-} = \operatorname{span}\{\boldsymbol{u}_1 \otimes \cdots \otimes \boldsymbol{u}_k \mid k \in \mathbb{N}, \{\boldsymbol{u}_i\} \text{ linearly dependent}\},$$

which is exactly the ideal $\{\!\{\boldsymbol{u} \otimes \boldsymbol{u}\}\!\}$. Therefore, the right-hand side of eq. (2.4) is identically the exterior algebra of definition 9. Hence, we have an algebra isomorphism $\mathrm{Sym}^{-}(V^{\otimes}) \cong \bigwedge V$, where the left-hand side is equipped with the product (2.3). This gives an alternative construction of the exterior algebra as the subalgebra of antisymmetric tensors.

18

Proof. If $A = \boldsymbol{u}_1 \otimes \cdots \otimes \boldsymbol{u}_k$ where two vectors $\boldsymbol{u}_i = \boldsymbol{u}_j$ are equal, then $\mathrm{Sym}^{-}(A) = 0$ since each term in the sum in eq. (2.2) is paired with an equal and opposite term with $i \leftrightarrow j$ swapped. If $\{\boldsymbol{u}_i\}$ is linearly dependent, any one vector is a sum of the others, so A is a sum of blades with at least two vectors repeated. □

Note on conventions

The factor of $\frac{1}{k!}$ present in eq. (2.2) is not necessary to derive the isomorphism $\mathrm{Sym}^-(V^\otimes) \cong \wedge V$. Indeed, some authors omit the normalisation factor, which has the effect of changing eq. (2.3) to[19]

[19] Written here with Sym^- *including* the factor $\frac{1}{k!}$, as in (2.2).

$$A \wedge B = \frac{(p+q)!}{p!q!}\mathrm{Sym}^-(A \otimes B)$$

for A and B of respective grades p and q. These different normalisations of $\wedge$ lead to distinct identifications of multivectors in $\wedge V$ with tensors in $S \subset V^\otimes$. Both conventions are present in literature, and their differences are clarified in table 2.1.

Kobayashi–Nomizu [15]	*Spivak* [16]
$A \wedge B \coloneqq \mathrm{Sym}^-(A \otimes B)$	$A \wedge B \coloneqq \frac{(p+q)!}{p!q!}\mathrm{Sym}^-(A \otimes B)$
$\boldsymbol{u} \wedge \boldsymbol{v} \equiv \frac{1}{2}(\boldsymbol{u} \otimes \boldsymbol{v} - \boldsymbol{v} \otimes \boldsymbol{u})$	$\boldsymbol{u} \wedge \boldsymbol{v} \equiv \boldsymbol{u} \otimes \boldsymbol{v} - \boldsymbol{v} \otimes \boldsymbol{u}$

Table 2.1.: Different embeddings of $\wedge V$ into $V^\otimes$. We employ the Kobayashi–Nomizu convention as this is coincides with the wedge product of geometric algebra. However, the Spivak convention is dominant for exterior differential forms in physics.

2.2.2. Exterior forms

The exterior algebra is most frequently encountered by physicists as an operation on *exterior (differential) forms*, which are alternating[20] multilinear maps. We *could* use the exterior algebra $\wedge V^*$ over the dual space of linear maps $V \to \mathbb{R}$ as a model for exterior forms, though we will not choose to do this, instead defining them separately.

[20] An ALTERNATING linear map is one which changes sign upon transposition of any pair of arguments.

To see why, consider $\wedge V^*$ as a model for exterior forms. Any element $f \in \wedge^k V^*$ has the form $f = f_{i_1 \cdots i_k} \boldsymbol{e}^{i_1} \wedge \cdots \wedge \boldsymbol{e}^{i_k}$ for a basis $\{\boldsymbol{e}^i\} \subset V^*$, and each component acts on $\boldsymbol{u}_1 \otimes \cdots \otimes \boldsymbol{u}_k \in V^{\otimes k}$ as

$$(\boldsymbol{e}^{i_1} \wedge \cdots \wedge \boldsymbol{e}^{i_k})(\boldsymbol{u}_1 \otimes \cdots \otimes \boldsymbol{u}_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (-1)^\sigma \boldsymbol{e}^{i_{\sigma(1)}}(\boldsymbol{u}_1) \cdots \boldsymbol{e}^{i_{\sigma(k)}}(\boldsymbol{u}_k)$$
$$= \frac{1}{k!} \det\left[\boldsymbol{e}^{i_m}(\boldsymbol{u}_n)\right]_{mn}. \tag{2.5}$$

However, this differs from the standard definition of exterior forms (as in [16, 17]) in two important ways:

1. In eq. (2.5), the dual vectors $\boldsymbol{e}^i \in V^*$ are permuted while the order of the arguments $\boldsymbol{u}_i$ are preserved; but for standard exterior forms, the opposite is true. This forbids the proper extension of $\bigwedge V^*$ to *non-Abelian vector-valued forms*, where the values $\boldsymbol{e}^i(\boldsymbol{u}_j)$ may not commute.

2. More trivially, we insist on the Kobayashi—Nomizu convention of normalisation factor for $\bigwedge V^*$; but the Spivak convention for exterior forms is much more common in physics.

For these reasons, we will not define exterior forms in terms of the exterior algebra $\bigwedge V^*$, but separately, to better agree with convention.

Definition 11. *For a vector space V over $\mathbb{F}$, a k-FORM $\varphi \in \Omega^k(V)$ is an alternating multilinear map $\varphi : V^{\otimes k} \to \mathbb{F}$.*

For another vector space A, an A-VALUED k-FORM $\varphi \in \Omega^k(V, A)$ is such a map with codomain A (instead of $\mathbb{F}$).

The evaluation of a form is denoted $\varphi(\boldsymbol{u}_1 \otimes \cdots \otimes \boldsymbol{u}_k)$ or $\varphi(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$, and the wedge product of a p-form φ and q-form ϕ is defined (in the Spivak convention) as

$$\varphi \wedge \phi = \frac{(p+q)!}{p!q!} (\varphi \otimes \phi) \circ \mathrm{Sym}^-. \tag{2.6}$$

Explicitly, eq. (2.6) acts to antisymmetrise *arguments*. To see this, choose a basis $\{\theta^\mu\}$ of $\Omega(V)$, and compare to eq. (2.5),

$$(\theta^{\mu_1} \wedge \cdots \wedge \theta^{\mu_k})(\boldsymbol{u}_1 \otimes \cdots \otimes \boldsymbol{u}_k) = \sum_{\sigma \in S_k} (-1)^\sigma \theta^{\mu_1}(\boldsymbol{u}_{\sigma(1)}) \cdots \theta^{\mu_k}(\boldsymbol{u}_{\sigma(k)}).$$

If $\varphi, \phi \in \Omega(V, A)$ are A-valued forms, where A is equipped with a bilinear product $\circledast : A \times A \to A$, then scalar multiplication may be replaced by $\circledast$ so that

$$(\varphi \wedge \phi)(\boldsymbol{u}_1 \otimes \cdots \otimes \boldsymbol{u}_k) = \sum_{\sigma \in S_k} (-1)^\sigma \varphi(\boldsymbol{u}_1 \otimes \cdots \otimes \boldsymbol{u}_p) \circledast \phi(\boldsymbol{u}_1 \otimes \cdots \otimes \boldsymbol{u}_q).$$

The product $\circledast$ need not be commutative nor associative. In particular, we may have Lie algebra–valued forms. For example, if $\varphi, \phi \in \Omega^1(V, \mathfrak{g})$, then we may take the Lie bracket $[\;,\;] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ to be the bilinear product on $\mathfrak{g}$, obtaining

$$(\varphi \wedge \phi)(\boldsymbol{u}, \boldsymbol{v}) = [\varphi(\boldsymbol{u}), \phi(\boldsymbol{v})] - [\varphi(\boldsymbol{v}), \phi(\boldsymbol{u})].$$

Note that this implies that $\varphi \wedge \varphi$ does not necessarily vanish for non-Abelian forms.[21]

[21] E.g., in the case above, $(\varphi \wedge \varphi)(\boldsymbol{u}, \boldsymbol{v}) = 2[\varphi(\boldsymbol{u}), \varphi(\boldsymbol{v})]$.

The two algebras $\bigwedge^V$ and $\Omega(V)$ are isomorphic, but differ in that elements of $\Omega(V)$ are identified as alternating maps (using the Spivak convention). In particular, this means that non-homogeneous exterior forms are often disallowed, since they cannot be identified with multilinear maps — while non-homogeneous elements of $\bigwedge^V$ are certainly allowed.[22]

[22] Non-homogeneous elements play important roles in the geometric algebra, in which $\bigwedge^V$ is embedded.

2.3. The Metric: Length and Angle

The tensor and exterior algebras considered so far are built from a vector space V alone. Notions of length and angle are central to geometry, but are not intrinsic to a vector space — this additional structure may be provided by a *metric.*

Definition 12. *A* METRIC[23] *is a function $\eta : V \times V \to \mathbb{F}$, often written $\langle \boldsymbol{u}, \boldsymbol{v} \rangle \equiv \eta(\boldsymbol{u}, \boldsymbol{v})$, which satisfies*

- *symmetry, $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$; and*
- *linearity, $\langle \alpha \boldsymbol{u} + \beta \boldsymbol{v}, \boldsymbol{w} \rangle = \alpha \langle \boldsymbol{u}, \boldsymbol{w} \rangle + \beta \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ for $\alpha, \beta \in \mathbb{F}$.*

[23] a.k.a. an *inner product*, or *symmetric bilinear form*

Linearity in either argument implies linearity in the other by symmetry, so η is bilinear. A metric is NON-DEGENERATE if $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$ for all $\boldsymbol{u}$ implies that $\boldsymbol{v}$ is zero. With respect to a basis $\{\boldsymbol{e}_i\}$ of V, the metric components $\eta_{ij} = \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle$ are defined. Non-degeneracy means that $\det \eta \neq 0$ when viewing $\eta = [\eta_{ij}]$ as a matrix, and in this case the matrix inverse η^{ij} is also defined and satisfies $\eta^{ik}\eta_{kj} = \delta^i_j$. Throughout, we will not have need to consider degenerate metrics,[24] so we assume non-degeneracy.

[24] Degenerate signatures do find use in computer graphics, especially via *projective geometric algebra* [18, 19].

A vector space V together with a metric η is called an INNER PRODUCT SPACE (V, η). Alternatively, instead of a metric, an inner product space may be constructed with a quadratic form:

Definition 13. *A* QUADRATIC FORM *is a function* $q : V \to \mathbb{F}$ *satisfying*

- $q(\lambda \boldsymbol{v}) = \lambda^2 q(\boldsymbol{v})$ *for all* $\lambda \in \mathbb{F}$*; and*
- *the requirement that the* POLARIZATION OF q,

$$(\boldsymbol{u}, \boldsymbol{v}) \mapsto q(\boldsymbol{u} + \boldsymbol{v}) - q(\boldsymbol{u}) - q(\boldsymbol{v}),$$

  is bilinear.

To any quadratic form q there is a unique associated bilinear form, which is *compatible* in the sense that $q(\boldsymbol{u}) = \langle \boldsymbol{u}, \boldsymbol{u} \rangle$. It is recovered[25] by the *polarization identity*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \frac{1}{2}\big(q(\boldsymbol{u} + \boldsymbol{v}) - q(\boldsymbol{u}) - q(\boldsymbol{v})\big).$$

[25] Except, of course, if the characteristic of $\mathbb{F}$ is two. We only consider fields of characteristic zero.

The prescription of either η or q is therefore equivalent — but the notion of a metric is more common in physics, whereas the mathematical viewpoint often starts with a quadratic form.

Covectors and dual bases

The dual space $V^* \coloneqq \{f : V \to \mathbb{F} \mid f \text{ linear}\}$ of a vector space consists of DUAL VECTORS or COVECTORS, which are linear maps from V into its underlying field. Summation convention dictates that components of vectors be written superscript, $\boldsymbol{u} = u^i \boldsymbol{e}_i \in V$, and covectors subscript, $\varphi = \varphi_i \boldsymbol{e}^i \in V^*$, for bases $\{\boldsymbol{e}_i\} \subset V$ and $\{\boldsymbol{e}^i\} \subset V^*$.

A metric η on V defines an isomorphism between V and its dual space. Collectively known as the MUSICAL ISOMORPHISMS, the maps $\flat : V \to V^*$ and its inverse $\sharp : V^* \to V$ are defined by

$$\boldsymbol{u}^\flat(\boldsymbol{v}) = \langle \boldsymbol{u}, \boldsymbol{v} \rangle \quad \text{and} \quad \left\langle \varphi^\sharp, \boldsymbol{u} \right\rangle = \varphi(\boldsymbol{u})$$

for $\boldsymbol{u}, \boldsymbol{v} \in V$ and $\varphi \in V^*$. The names become justified when working with a basis: the relations

$$(\boldsymbol{u}^\flat)_i = \eta_{ij} \boldsymbol{u}^j \quad \text{and} \quad (\varphi^\sharp)^i = \eta^{ij} \varphi_j$$

show that $\flat$ *lowers* indices, while $\sharp$ *raises* them.

Given a metric, a choice of basis $\{\boldsymbol{e}_i\} \subset V$ also defines a DUAL BASIS $\{\boldsymbol{e}^i\} \subset V^*$ OF V via $\boldsymbol{e}^i \coloneqq \eta^{ij} \boldsymbol{e}_j^\flat$. Note that basis vectors and covectors defined in this way do not exist in the same vector space, but are related by their evaluation on one another by $\boldsymbol{e}^i(\boldsymbol{e}_j) = \delta^i_j$. In some contexts, we will define a dual basis $\{\boldsymbol{e}^i\}$ in V (not in V^*), a.k.a. a RECIPROCAL BASIS, and by this we mean $\boldsymbol{e}^i \coloneqq \eta^{ij} \boldsymbol{e}_j$. Then, dual and non-dual basis vectors are related via $\left\langle \boldsymbol{e}^i, \boldsymbol{e}_j \right\rangle = \delta^i_j$.

We use both senses of the term "dual basis". In particular, V^* is never needed in the geometric algebra; its role is filled by reciprocal bases. Often, the distinction can be safely ignored (since, after all, $V \cong V^*$).

2.3.1. Metrical exterior algebra

In an exterior algebra $\bigwedge V$ with a metric defined on V, there is an induced metric on k-vectors defined by

$$\begin{aligned} \langle \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_k, \boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{v}_k \rangle &= \sum_{\sigma \in S_k} (-1)^\sigma \left\langle \boldsymbol{u}_1, \boldsymbol{v}_{\sigma(1)} \right\rangle \cdots \left\langle \boldsymbol{u}_k, \boldsymbol{v}_{\sigma(k)} \right\rangle \\ &= \det[\langle \boldsymbol{u}_m, \boldsymbol{u}_n \rangle]_{mn}. \end{aligned}$$

In particular, a metric on $\bigwedge V$ defines a magnitude for pseudoscalars.

Definition 14. *Let V be an n-dimensional vector space with a metric. There are two volume elements $\mathbb{I} \in \bigwedge^n V$ of the metrical exterior algebra $\bigwedge V$ is a pseudoscalar satisfying $\langle \mathbb{I}, \mathbb{I} \rangle = 1$, each differing by sign.*

A choice of ORIENTATION *is a choice of volume element $\mathbb{I}$.*

Given an ordered orthonormal basis $\{\boldsymbol{e}_i\}$ with $\langle \boldsymbol{e}_i, \boldsymbol{e}_i \rangle = \pm 1$, the basis is called right-handed if $\boldsymbol{e}_1 \wedge \cdots \wedge \boldsymbol{e}_n = \mathbb{I}$ is the chosen volume element, and left-handed otherwise.

Hodge–dual multivectors

Hodge duality: [17, 20], [21, §16].

A useful duality operation can be defined in an exterior algebra $\bigwedge V$ with a metric and orientation, which relates the k- and $(n-k)$-grade subspaces.

Definition 15. *Let $\bigwedge V$ be a metrical exterior algebra with base dimension n and volume element $\mathbb{I}$. The* HODGE DUAL $\star$ *is the unique linear operator satisfying*

$$A \wedge \star B = \langle A, B \rangle \, \mathbb{I} \tag{2.7}$$

for any k-vectors $A, B \in \bigwedge^k V$.

The Hodge dual $\star : \bigwedge^k V \to \bigwedge^{n-k} V$ defines an isomorphism between pairs of fixed-grade subspaces of the same dimension; in particular, scalars with pseudoscalars via $\star 1 = \mathbb{I}$.

Lemma 4. *The hodge dual of a p-vector $A = A^{i_1 \cdots i_p} \boldsymbol{e}_{i_1} \wedge \cdots \wedge \boldsymbol{e}_{i_p}$ in n dimensions has components*

$$(\star A)^{j_1 \cdots j_q} = \frac{1}{p!} A_{i_1 \cdots i_p} \varepsilon^{i_1 \cdots i_p j_1 \cdots j_q}$$

where $q = n - p$ and $A_{i_1 \cdots i_p} = A^{k_1 \cdots k_p} \eta_{i_1 k_1} \cdots \eta_{i_p k_p}$.

The LEVI-CIVITA SYMBOL *$\varepsilon^{i_1 \cdots i_k}$ is the sign of the permutation $(1, \ldots, k) \mapsto (i_1, \ldots, i_k)$ if it exists, and zero otherwise.*

Proof. Begin by writing the left-hand side of eq. (2.7) in component form,

$$\begin{aligned} A \wedge \star B &= A^{i_1 \cdots i_p} \, (\star B)^{j_1 \cdots j_p} \, \boldsymbol{e}_{i_1} \wedge \cdots \wedge \boldsymbol{e}_{i_p} \wedge \boldsymbol{e}_{j_1} \wedge \cdots \wedge \boldsymbol{e}_{j_p} \\ &= A^{i_1 \cdots i_p} \, (\star B)^{j_1 \cdots j_p} \, \varepsilon_{i_1 \cdots i_p j_1 \cdots j_p} \mathbb{I}. \end{aligned} \tag{2.8}$$

If the multi-indices i_k and $j_k = i_{\sigma(i)}$ are permutations of each other, then

$$\det\left[\eta_{i_m j_n}\right]_{mn} = (-1)^\sigma \eta_{i_1 j_{\sigma(1)}} \cdots \eta_{i_p j_{\sigma(p)}}.$$

Using this, the right-hand side of eq. (2.7) is

$$\begin{aligned}\langle A, B\rangle\, \mathbb{I} &= A^{i_1\cdots i_p} B^{j_1\cdots j_p} \det\left[\eta_{i_m j_n}\right]_{mn} \mathbb{I} \\ &= A^{i_1\cdots i_p} B^{j_1\cdots j_p} (-1)^\sigma \eta_{i_1 j_{\sigma(1)}} \cdots \eta_{i_p j_{\sigma(p)}} \mathbb{I} \\ &= A^{i_1\cdots i_p} B_{i_{\sigma^{-1}(1)}\cdots i_{\sigma^{-1}(p)}} (-1)^\sigma \mathbb{I} = A^{i_1\cdots i_p} B_{i_1\cdots i_p} \mathbb{I}\end{aligned} \tag{2.9}$$

where we absorb the factor $(-1)^\sigma$ by permuting indices of B. Equating coefficients of the arbitrary $A^{i_1\cdots i_p}$ components, eqs. (2.8) and (2.9) gives

$$(\star B)^{j_1\cdots j_p} \varepsilon_{i_1\cdots i_p j_1\cdots j_p} = B_{i_1\cdots i_p}$$

Contracting with $\varepsilon_{i_1\cdots i_p k_1\cdots k_p}$ gives

$$(\star B)^{k_1\cdots k_p} = \frac{1}{p!} B_{i_1\cdots i_p} \varepsilon^{i_1\cdots i_p k_1\cdots k_p}$$

since $\varepsilon_{i_1\cdots i_p j_1\cdots j_p} \varepsilon^{i_1\cdots i_p k_1\cdots k_p} = (-1)^\sigma p!$ where σ is the permutation sending $j_i \mapsto k_i$. The factor of $(-1)^\sigma$ is absorbed since $(\star B)^{j_1\cdots j_p} = (-1)^\sigma (\star B)^{k_1\cdots k_p}$. Finally, replacing $A \leftrightarrow B$ is the result. □

Chapter 3.

The Geometric Algebra

In chapter 2, we defined the metric-independent exterior algebra of multivectors over a vector space V. While metrical operations can be achieved by introducing the Hodge dual (of section 2.3.1), tacking it onto $\bigwedge V$, the geometric algebra is a generalisation of $\bigwedge V$ which has the metric (and concomitant notions of orientation and duality) built-in.

Geometric algebras are also known as real *Clifford algebras* $Cl(V, q)$ after their first inventor [4]. Especially in mathematics, Clifford algebras are defined in terms of a quadratic form q, and the vector space V is usually complex. However, in physics, where V is taken to be real and a metric η is usually supplied instead of q, the name "geometric algebra" is preferred.[26]

[26] The newer name was coined by David Hestenes in the 1970s, who popularised Clifford algebra for physics [22, 23].

3.1. Construction and Overview

Informally put, the geometric algebra is obtained by enforcing the single rule

$$\boldsymbol{u}^2 = \langle \boldsymbol{u}, \boldsymbol{u} \rangle \tag{3.1}$$

for any vector $\boldsymbol{u}$, along with the associative algebra axioms of definition 1. The rich algebraic structure which follows from this is remarkable. Formally, we may give the geometric algebra as a quotient, just like our presentation of $\bigwedge V$.

Definition 16. *Let V be a finite-dimensional real vector space with metric. The* GEOMETRIC ALGEBRA *over V is*

$$\mathscr{G}(V,\eta) := V^{\otimes} \big/ \{\!\{ \boldsymbol{u} \otimes \boldsymbol{u} - \langle \boldsymbol{u}, \boldsymbol{u} \rangle \}\!\} .$$

The ideal defines the congruence generated by $\boldsymbol{u} \otimes \boldsymbol{u} \sim \langle \boldsymbol{u}, \boldsymbol{u} \rangle$, encoding eq. (3.1). This uniquely defines the associative (but not generally commutative) *geometric product* which we denote by juxtaposition.

As 2^n-dimensional vector spaces, $\mathscr{G}(V,\eta)$ and $\bigwedge V$ are isomorphic, each with a $\binom{n}{k}$-dimensional subspace for each grade k. Denoting the k-grade subspace $\mathscr{G}_k(V,\eta)$, we have the vector space decomposition

$$\mathscr{G}(V,\eta) = \bigoplus_{k=0}^{\infty} \mathscr{G}_k(V,\eta).$$

Note that this is not a $\mathbb{Z}$ grading of the geometric algebra: the quotient is by *inhomogeneous* elements $\boldsymbol{u} \otimes \boldsymbol{u} - \langle \boldsymbol{u}, \boldsymbol{u} \rangle \in V^{\otimes 2} \oplus V^{\otimes 0}$, and therefore the geometric product of a p-vector and a q-vector is not generally a $(p+q)$-vector. However, the congruence *is* homogeneous with respect to the $\mathbb{Z}_2$-grading, so $\mathscr{G}(V,\eta)$ is $\mathbb{Z}_2$-graded. This shows that the algebra separates into 'even' and 'odd' subspaces

$$\mathscr{G}(V,\eta) = \mathscr{G}_+(V,\eta) \oplus \mathscr{G}_-(V,\eta) \quad \text{where} \quad \begin{cases} \mathscr{G}_+(V,\eta) = \bigoplus_{k=0}^{\infty} \mathscr{G}_{2k}(V,\eta) \\ \mathscr{G}_+(V,\eta) = \bigoplus_{k=0}^{\infty} \mathscr{G}_{2k+1}(V,\eta) \end{cases}$$

where $\mathscr{G}_+(V,\eta)$ is closed under the geometric product, forming the EVEN SUBALGEBRA.

The geometric product of vectors

By expanding $(\boldsymbol{u}+\boldsymbol{v})^2 = \langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v} \rangle$, it directly follows that

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \frac{1}{2}(\boldsymbol{u}\boldsymbol{v} + \boldsymbol{v}\boldsymbol{u}).$$

We recognise this as the symmetrised product of two vectors. The remaining antisymmetric part coincides with the *alternating* or *wedge* product familiar from exterior algebra

$$\boldsymbol{u} \wedge \boldsymbol{v} = \frac{1}{2}(\boldsymbol{uv} - \boldsymbol{vu}).$$

This is a 2-vector, or bivector, in $\mathscr{G}_2(V, \eta)$. Thus, the geometric product on vectors is

$$\boldsymbol{uv} = \langle \boldsymbol{u}, \boldsymbol{v} \rangle + \boldsymbol{u} \wedge \boldsymbol{v},$$

and some important features are immediate:

- *Parallel vectors commute, and vice versa:* If $\boldsymbol{u} = \lambda \boldsymbol{v}$, then $\boldsymbol{u} \wedge \boldsymbol{v} = 0$ and $\boldsymbol{uv} = \langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle = \boldsymbol{vu}$.

- *Orthogonal vectors anti-commute, and vice versa:* If $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$, then $\boldsymbol{uv} = \boldsymbol{u} \wedge \boldsymbol{v} = -\boldsymbol{v} \wedge \boldsymbol{u} = -\boldsymbol{vu}$.

In particular, if $\{\boldsymbol{e}_i\} \subset V$ is an orthonormal basis, then we have $\boldsymbol{e}_i^2 = \langle \boldsymbol{e}_i, \boldsymbol{e}_i \rangle$ and $\boldsymbol{e}_i \boldsymbol{e}_j = -\boldsymbol{e}_j \boldsymbol{e}_i$, which can be summarised by the anticommutation relation $\boldsymbol{e}_i \boldsymbol{e}_j + \boldsymbol{e}_j \boldsymbol{e}_i = 2\eta_{ij}$.

- *Vectors are invertible under the geometric product:* If $\boldsymbol{u}$ is a vector for which the scalar $\boldsymbol{u}^2$ is non-zero, then $\boldsymbol{u}^{-1} = \boldsymbol{u}/\boldsymbol{u}^2$.

- *Geometric multiplication produces objects of mixed grade:* The product $\boldsymbol{uv}$ has a scalar part $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$ and a bivector part $\boldsymbol{u} \wedge \boldsymbol{v}$.

Higher-grade elements

As with two vectors, the geometric product of two homogeneous multivectors is generally inhomogeneous. We can gain insight by separating geometric products into grades and studying each part.

Definition 17. *The* GRADE k PROJECTION *of a multivector* $A \in \mathscr{G}(V, \eta)$ *is*

$$\langle A \rangle_k = \begin{cases} A & \text{if } A \in \mathscr{G}_k(V, \eta) \\ 0 & \text{otherwise.} \end{cases}$$

We can generalise the definition of the wedge product of vectors $\boldsymbol{u} \wedge \boldsymbol{v} = \langle \boldsymbol{uv} \rangle_2$ to arbitrary homogeneous multivectors by taking the highest-grade part of their product,

$$A \wedge B = \langle AB \rangle_{p+q},$$

where $A \in \mathscr{G}_p(V, \eta)$ and $B \in \mathscr{G}_q(V, \eta)$. Dually, we can define an inner product on homogeneous multivectors by taking the lowest-grade part, $|p - q|$. These can be extended by linearity to inhomogeneous elements.

Definition 18. *Let $A, B \in \mathscr{G}(V, \eta)$ be possibly inhomogeneous multivectors. The* WEDGE PRODUCT *is*

$$A \wedge B := \sum_{p,q} \Big\langle \langle A \rangle_p \langle B \rangle_q \Big\rangle_{p+q},$$

and the GENERALISED INNER PRODUCT*, or "fat dot" product,*[27] *is*

$$A \bullet B := \sum_{p,q} \Big\langle \langle A \rangle_p \langle B \rangle_q \Big\rangle_{|p-q|}.$$

[27] Various 'inner products' have been proposed for geometric algebra, but the definition here is arguably the simplest and best behaved; see [24] for closer discussion.

With the wedge product defined on all of $\mathscr{G}(V, \eta)$, we use language of multivectors as we did with the exterior algebra, so that $\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_k \in \mathscr{G}_k(V, \eta)$ is a k-blade, and a sum of k-blades is a k-multivector, etcetea. The products in definition 18 work together nicely, and extend the notion of a dual vector basis to a dual basis of blades.

Lemma 5. *If $\{\boldsymbol{e}_i\} \subset V$ is a basis with dual $\boldsymbol{e}^i \bullet \boldsymbol{e}_j = \delta^i_j$, then*

Note the reverse order of the j indices.

$$(\boldsymbol{e}^{i_1} \wedge \cdots \wedge \boldsymbol{e}^{i_k}) \bullet (\boldsymbol{e}_{j_k} \wedge \cdots \wedge \boldsymbol{e}_{j_1}) = \varepsilon^{i}_{j}$$

where $\varepsilon^i_j = (-1)^\sigma$ is the sign of the permutation sending $\sigma(i_p) = j_p$ for $1 \leq p \leq k$, or zero if there is no such permutation or if i or j contain repeated indices.

Proof. If i or j contain repeated indices, then the left-hand side vanishes by antisymmetry of the wedge product, and the right-hand side by definition. If i contains no repeated indices, and the j indices are some permutation $j_p = \sigma(i_p)$, then $\boldsymbol{e}^{i_1} \wedge \cdots \wedge \boldsymbol{e}^{i_k} = \boldsymbol{e}^{i_1} \cdots \boldsymbol{e}^{i_k}$ by orthogonality. Rewriting the left-hand side,

$$\left\langle \boldsymbol{e}^{i_1} \cdots \boldsymbol{e}^{i_k} \boldsymbol{e}_{j_k} \cdots \boldsymbol{e}_{j_1} \right\rangle = (-1)^\sigma \left\langle \boldsymbol{e}^{i_1} \cdots \underbrace{\boldsymbol{e}^{i_k} \boldsymbol{e}_{i_k}}_{1} \cdots \boldsymbol{e}_{i_1} \right\rangle = (-1)^\sigma.$$

Finally, if i contains no repeated indices, but j is not a permutation of i, then there is at least one pair of indices in the symmetric difference of $\{i_p\}$ and $\{j_p\}$, say i_r and j_s. Commuting this pair $\boldsymbol{e}^{i_r}$ and $\boldsymbol{e}_{j_s}$ together shows that the left-hand side vanishes, since $\boldsymbol{e}^{i_r} \boldsymbol{e}_{j_s} = 0$. □

3.2. Relations to Other Algebras

An efficient way to become familiar with the geometric algebra is to exemplify its relationships and isomorphisms with other algebras and with itself.

3.2.1. Fundamental algebra automorphisms

Operations such complex conjugation $\overline{AB} = \overline{A}\,\overline{B}$ or matrix transposition $(AB)^\top = B^\top A^\top$ are useful because they preserve or reverse multiplication. Linear functions with this property are called algebra automorphisms or antiautomorphisms, respectively. The geometric algebra possesses this (anti)automorphism operations.

Isometries of (V, η) are linear functions $f : V \to V$ which preserve the metric, so that $\langle f(\boldsymbol{u}), f(\boldsymbol{v}) \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle$ for any $\boldsymbol{u}, \boldsymbol{v} \in V$. Vector spaces always possess the involution isometry $\boldsymbol{u} \mapsto -\boldsymbol{u}$, as well as the trivial isometry. An isometry extends uniquely to an algebra (anti)automorphism by defining $f(AB) = f(A)f(B)$ or $f(AB) = f(B)f(A)$. Thus, by extending the two fundamental isometries of (V, η) in the two possible ways, we obtain four fundamental (anti)automorphisms of $\mathscr{G}(V, \eta)$.

Definition 19. *Let $\boldsymbol{u} \in \mathscr{G}_1(V, \eta)$ be a vector and $A, B \in \mathscr{G}(V, \eta)$ possubly inhomogeneous multivectors in a geometric algebra.*

- *Reversion* $\dagger$ *is the identity map on vectors* $\boldsymbol{u}^\dagger = \boldsymbol{u}$ *extended to general multivectors by the rule* $(AB)^\dagger = B^\dagger A^\dagger$.

- *Grade involution* $\star$ *is the extension of the involution* $\boldsymbol{u}^\star = -\boldsymbol{u}$ *to general multivectors by the rule* $(AB)^\star = A^\star B^\star$.

Note that if $A \in \mathscr{G}_k(V, \eta)$ is a k-vector, then $\iota(A) = (-1)^k A$ and $A^\dagger = s_k A$ where

$$s_k = (-1)^{\frac{(k-1)k}{2}} \tag{3.2}$$

is the sign of the reverse permutation on k symbols.

Reversion and grade involution together generate the four fundamental automorphisms

$\star \circ \dagger$ is sometimes referred to as the Clifford conjugate

id	$\star$	automorphisms
$\dagger$	$\star \circ \dagger$	anti-automorphisms

which form a group isomorphic to $\mathbb{Z}_2^2$ under composition.

These operations are very useful in practice. In particular, the following result follows easily from reasoning about grades.

Lemma 6. *If* $A \in \mathscr{G}_k(V, \eta)$ *is a* k*-vector, then* A^2 *is a* $4\mathbb{N}$*-multivector, i.e., a sum of blades of grade* $\{0, 4, 8, \dots\}$ *only.*

Proof. The multivector A^2 is its own reverse, since $(A^2)^\dagger = (A^\dagger)^2 = (\pm A)^2 = A^2$, and hence has parts of grade $\{4n, 4n+1 \mid n \in \mathbb{N}\}$. Similarly, A^2 is self-involutive, since $(A^2)^\star = (A^\star)^2 = (\pm A)^2 = A^2$. It is thus of even grade, leaving the possible grades $\{0, 4, 8, \dots\}$. $\square$

3.2.2. Even subalgebra isomorphisms

As noted above, multivectors of even grade are closed under the geometric product, and form the even subalgebra $\mathscr{G}_+(p, q)$. There is an isomorphism $\mathscr{G}_+(p, q) \cong \mathscr{G}_+(q, p)$ given by $\bar{\boldsymbol{e}}_i \coloneqq \boldsymbol{e}_i$ with opposite signature

$\bar{\boldsymbol{e}}_i{}^2 := -\boldsymbol{e}_i^2$, since the factor of -1 occurs only an even number of times for even elements.

The even subalgebras are also isomorphic to full geometric algebras of one dimension less:

Lemma 7. *There are isomorphisms*

$$\mathscr{G}_+(p,q) \cong \mathscr{G}(p,q-1) \quad \textit{and} \quad \mathscr{G}_+(p,q) \cong \mathscr{G}(q,p-1)$$

when $q \geq 1$ and $p \geq 1$, respectively.

Proof. Select a unit vector $\boldsymbol{u} \in \mathscr{G}(p,q)$ with $\boldsymbol{u}^2 = -1$, and define a linear map $\Psi_{\boldsymbol{u}} : \mathscr{G}(p,q-1) \to \mathscr{G}_+(p,q)$ by

$$\Psi_{\boldsymbol{u}}(A) = \begin{cases} A & \text{if } A \text{ is even} \\ A \wedge \boldsymbol{u} & \text{if } A \text{ is odd} \end{cases}.$$

Note we are taking $\mathscr{G}(p,q-1) \subset \mathscr{G}(p,q)$ to be the subalgebra obtained by removing $\boldsymbol{u}$ (i.e., restricting V to $\boldsymbol{u}^\perp$) so there is a canonical inclusion from the domain of $\Psi_{\boldsymbol{u}}$ to the codomain. Let $A \in \mathscr{G}(p,q-1)$ be a k-vector. Note that $A \wedge \boldsymbol{u} = A\boldsymbol{u}$ since $\boldsymbol{u} \perp \mathscr{G}(p,q-1)$, and that A commutes with $\boldsymbol{u}$ if k is even and anticommutes if k is odd.

To so $\Psi_{\boldsymbol{u}}$ is a homomorphism, suppose $A, B \in \mathscr{G}(p,q-1)$ are both even; then $\Psi_{\boldsymbol{u}}(AB) = AB = \Psi_{\boldsymbol{u}}(A)\Psi_{\boldsymbol{u}}(B)$. If both are odd, then AB is even and $\Psi_{\boldsymbol{u}}(AB) = AB = -AB\boldsymbol{u}^2 = A\boldsymbol{u}B\boldsymbol{u} = \Psi_{\boldsymbol{u}}(A)\Psi_{\boldsymbol{u}}(B)$. If A is odd and B even, then $\Psi_{\boldsymbol{u}}(AB) = AB\boldsymbol{u} = A\boldsymbol{u}B = \Psi_{\boldsymbol{u}}(A)\Psi_{\boldsymbol{u}}(B)$ and similarly for A even and B odd. Injectivity and surjectivity are clear, so $\Psi_{\boldsymbol{u}}$ is an algebra isomorphism. □

The special case $\mathscr{G}_+(1,3) \cong \mathscr{G}(3)$ is of great relevance to special relativity, and is discussed in detail in section 4.1. Here the isomorphism $\Psi_{\boldsymbol{u}}$ is called a *space/time split* with respect to an observer of velocity $\boldsymbol{u}$. This provides an impressively efficient algebraic method for transforming relativistic quantities between inertial frames.

3.2.3. Relation to exterior forms

The geometric algebra is a generalisation of the exterior algeba. Indeed, in the special case of a completely degenerate metric (i.e., $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_0 = 0$ for all vectors), then there is an algebra isomorphism $\mathscr{G}(V, 0) \cong \bigwedge V$.

Exterior forms as multivectors

Exterior forms can be mimicked in the geometric algebra by making use of a dual basis V, as in the following lemma. Note that the dual space V^* does not make an appearance — all elements belong to $\mathscr{G}(V, \eta)$.

Lemma 8. *If $A \in \mathscr{G}_k(V, \eta)$ is a k-vector and $\varphi \in \Omega^k(V)$ is a k-form whose components coincide (i.e., $A_{i_1 \cdots i_k} = \varphi_{i_1 \cdots i_k}$ given a common basis of V) then*

$$A \bullet (\boldsymbol{u}_k \wedge \cdots \wedge \boldsymbol{u}_1) = k!\, \varphi(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k).$$

Note the reversed order of the wedge products on the left-hand side. The factor of $k!$ is due to the Spivak convention for exterior forms (replace $k! \mapsto 1$ for the Kobayashi–Nomizu convention).

Proof. Fix an orthonormal basis $\{\boldsymbol{e}_i\} \subset V$ and a dual basis $\boldsymbol{e}^i \bullet \boldsymbol{e}_j = \delta^i_j$. Expanding the right-hand side with respect to his basis,

$$A \bullet (\boldsymbol{u}_k \wedge \cdots \wedge \boldsymbol{u}_1) = A_{i_1 \cdots i_k} (\boldsymbol{e}^{i_1} \wedge \cdots \wedge \boldsymbol{e}^{i_k}) \bullet (\boldsymbol{e}_{j_k} \wedge \cdots \wedge \boldsymbol{e}_{j_1}) u_k^{j_k} \cdots u_1^{j_1}.$$

By lemma 5, the dot product of k-blades is $(-1)^\sigma$ is the sign of the permutation $\sigma(i_p) = j_p$, and zero for all non-permutation terms in the sum. Thus, for each (non-zero) term in the sum we have

$$u_1^{j_1} \cdots u_k^{j_k} = u_1^{\sigma^{-1}(j_1)} \cdots u_k^{\sigma^{-1}(j_k)} = u_{\sigma(1)}^{i_1} \cdots u_{\sigma(k)}^{i_k},$$

where the last equality is obtained by permuting the scalar components $u_{\sigma(p)}^{i_p}$ by σ. Putting this together,

$$A \bullet (\boldsymbol{u}_k \wedge \cdots \wedge \boldsymbol{u}_1) = \sum_{\sigma \in S_k} (-1)^\sigma A_{i_1 \cdots i_k} u_{\sigma(1)}^{i_1} \cdots u_{\sigma(k)}^{i_k},$$

which by $A_{i_1 \cdots i_k} = \varphi_{i_1 \cdots i_k}$ is equal to

$$\cdots = \sum_{\sigma \in S_k} (-1)^\sigma \varphi(\boldsymbol{u}_{\sigma(1)}, \dots, \boldsymbol{u}_{\sigma(k)}) = k!\, \varphi(\boldsymbol{u}_{\sigma(1)}, \dots, \boldsymbol{u}_{\sigma(k)})$$

where all $k!$ terms are equal due to the alternating property of φ. □

Pseudoscalars and Hodge duality

Since the metric is built into the geometric algebra, so are the features of metrical exterior algebra of section 2.3.1, including the Hodge dual. In geometric algebra, Hodge duality is achieved simply by multiplication by the volume element.

Consider two k-vectors A and B. The object $B\mathbb{I}$ is then a $(n-k)$-vector, and its wedge product with A is a pseudoscalar. From associativity of the geometric product, we immediately have

$$A \wedge (B\mathbb{I}) = \langle A(B\mathbb{I}) \rangle_n = \langle (AB)\mathbb{I} \rangle_n = \langle AB \rangle \mathbb{I}.$$

This generalises the defining relation of the Hodge dual, eq. (2.7). The duality given by $A \mapsto A\mathbb{I}$ is trivial to manipulate algebraically, compared to the implicitly-defined exterior algebraic equivalent (definition 15).

Using the Hodge dual, the geometric product (of vectors) may be defined entirely within the exterior algebra as

$$\boldsymbol{uv} := (-1)^s \star (\boldsymbol{u} \wedge \star \boldsymbol{v}) + \boldsymbol{u} \wedge \boldsymbol{v} \tag{3.3}$$

where s is the signature of the metric (so that $(-1)^s = \det \eta$). Indeed, eq. (3.3) reduces to the familiar formula

$$\boldsymbol{uv} = (-1)^s \langle \boldsymbol{u}, \boldsymbol{v} \rangle \star \mathbb{I} + \boldsymbol{u} \wedge \boldsymbol{v} = \langle \boldsymbol{u}, \boldsymbol{v} \rangle + \boldsymbol{u} \wedge \boldsymbol{v}$$

by eq. (2.7) and $\star\mathbb{I} = (-1)^s$. However, eq. (3.3) does not apply to general multivectors, and the equivalent formulae for higher-grade objects are more complex and tend to obscure the underlying simplicity of the geometric product.

3.2.4. Common algebra isomorphisms

While $\mathscr{G}(V,\eta)$ encompasses the exterior algebra $\bigwedge V$, one important difference between the two algebras is that while inhomogeneous multivectors find little use in exterior algebra,[28] they have a significant role in describing reflections and rotations in $\mathscr{G}(V,\eta)$. Many familiar algebraic structures concerning rotations in relativistic or quantum physics are in fact special cases of geometric algebra.

[28] In fact, some authors [20] leave sums of terms of $\bigwedge V$ of differing grade undefined.

- **Complex numbers:** $\mathscr{G}_+(2) \cong \mathbb{C}$

  The complex plane is contained within $\mathscr{G}(2)$ as the even subalgebra, with the isomorphism

  $$\mathbb{C} \ni x + iy \leftrightarrow x + y\boldsymbol{e}_1\boldsymbol{e}_2 \in \mathscr{G}_+(2)$$

  Complex conjugation in $\mathbb{C}$ coincides with reversion in $\mathscr{G}(2)$.

- **Quaternions:** $\mathscr{G}_+(3) \cong \mathbb{H}$

  Similarly, the quaternions are the even subalgebra $\mathscr{G}_+(3)$, with the isomorphism[29]

  $$q_0 + q_1\hat{\boldsymbol{\imath}} + q_2\hat{\boldsymbol{\jmath}} + q_3\hat{\boldsymbol{k}} \longleftrightarrow q_0 + q_1\boldsymbol{e}_2\boldsymbol{e}_3 - q_2\boldsymbol{e}_3\boldsymbol{e}_1 + q_3\boldsymbol{e}_1\boldsymbol{e}_2.$$

  Again, quaternion conjugation corresponds to reversion in $\mathscr{G}(3)$.

[29] Note the minus sign. Viewed as rotations through their respective normal planes, $(\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ form a *left*-handed basis. This is because Hamilton chose $\hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}} = -1$, not $+1$.

- **Complexified quaternions**: $\mathscr{G}_+(1,3) \cong \mathbb{C} \otimes \mathbb{H}$

  The complexified quaternion algebra, which has been applied to special relativity [7, 9, 10], is isomorphic to the subalgebra $\mathscr{G}_+(1,3)$. The isomorphism

  $$\begin{aligned}\mathbb{C} \otimes \mathbb{H} \ni (x + yi) \otimes (q_0 + q_1\hat{\boldsymbol{\imath}} + q_2\hat{\boldsymbol{\jmath}} + q_3\hat{\boldsymbol{k}}) &\longleftrightarrow \\ (x + y\boldsymbol{e}_{0123})(q_0 + q_1\boldsymbol{e}_{23} - q_2\boldsymbol{e}_{31} + q_3\boldsymbol{e}_{12}) &\in \mathscr{G}_+(1,3)\end{aligned}$$

  associates quaternion units with bivectors, and the complex plane with the scalar–pseudoscalar plane. Reversion in $\mathscr{G}(1,3)$ corresponds to quaternion conjugation (preserving the complex i).

- **The Pauli algebra**: $\mathscr{G}(3) \cong \{\sigma_i\}_{i=1}^3$

The *algebra of physical space*, $\mathscr{G}(3)$, admits a complex representation $\boldsymbol{e}_i \longleftrightarrow \sigma_i$ via the Pauli spin matrices

$$\sigma_1 = \begin{pmatrix} 0 & +1 \\ +1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ +i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Reversion in $\mathscr{G}(3)$ corresponds to the adjoint (Hermitian conjugate), and the volume element $\mathbb{I} \coloneqq \boldsymbol{e}_{123} \longleftrightarrow \sigma_1\sigma_2\sigma_3 = i$ corresponds to the unit imaginary.

- **The Dirac algebra**: $\mathscr{G}(1,3) \cong \left\{\gamma_\mu\right\}_{\mu=0}^{3}$

  The relativistic analogue to the Pauli algebra is the Dirac algebra, generated by the 4×4 complex Dirac matrices

$$\gamma_0 = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \gamma_1 = \begin{pmatrix} 0 & +\sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}, \qquad \gamma_2 = \begin{pmatrix} 0 & -i\sigma_2 \\ +i\sigma_2 & 0 \end{pmatrix}, \qquad \gamma_3 = \begin{pmatrix} 0 & +\sigma_3 \\ -\sigma_3 & 0 \end{pmatrix}.$$

  These form a complex representation of the *algebra of spacetime*, $\mathscr{G}(1,3)$, via $\boldsymbol{e}_\mu \longleftrightarrow \gamma_\mu$. Again, reversion corresponds to the adjoint, and $\mathbb{I} \coloneqq \boldsymbol{e}_0\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3 \longleftrightarrow \gamma_0\gamma_1\gamma_2\gamma_3 = -i\gamma_5$.

3.3. Rotors and the Associated Lie Groups

There is a consistent pattern in the algebra isomorphisms listed in section 3.2.4. Note how the complex numbers $\mathbb{C}$ are fit for describing SO(2) rotations in the plane, and the quaternions $\mathbb{H}$ describe SO(3) rotations in $\mathbb{R}^3$. Common to both their respective isomorphisms with $\mathscr{G}_+(2)$ and $\mathscr{G}_+(3)$ is the identification of each "imaginary unit" in $\mathbb{C}$ or $\mathbb{H}$ with a *unit bivector* in $\mathscr{G}(n)$.

- In 2d, there is one linearly independent bivector, $\boldsymbol{e}_1\boldsymbol{e}_2$, and one imaginary unit, i.
- In 3d, there are $\dim \mathscr{G}_2(3) = \binom{3}{2} = 3$ such bivectors, and so three imaginary units $\left\{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\right\}$ are needed.

- In (1+3)d, we have $\dim \mathscr{G}_2(1,3) = \binom{4}{2} = 6$, corresponding to three 'spacelike' $\{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\}$ and three 'timelike' $\{i\hat{\boldsymbol{\imath}}, i\hat{\boldsymbol{\jmath}}, i\hat{\boldsymbol{k}}\}$ units of $\mathbb{C} \otimes \mathbb{H}$.

The interpretation of a bivector is clear: it takes the role of an 'imaginary unit', generating a rotation through the oriented plane which it spans.

To see how bivectors act as rotations, observe that rotations in the $\mathbb{C}$-plane may be described as mappings $z \mapsto e^{\theta i} z$, while $\mathbb{R}^3$ rotations are described in $\mathbb{H}$ using a double-sided transformation law, $u \mapsto e^{\theta \hat{\boldsymbol{n}}/2} u e^{-\theta \hat{\boldsymbol{n}}/2}$, where $\hat{\boldsymbol{n}} \in \operatorname{span}\{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\}$ is a unit quaternion defining the plane of rotation. Due to the commutativity of $\mathbb{C}$, the double-sided transformation law is actually general to both $\mathbb{C}$ and $\mathbb{H}$.

Similarly, rotations in a geometric algebra are described as

$$u \mapsto e^{\theta \hat{b}/2} u e^{-\theta \hat{b}/2},$$

where $\hat{b} \in \mathscr{G}_2(V, \eta)$ is a unit bivector. Multivectors of the form $R = e^{\sigma}$ for $\sigma \in \mathscr{G}_2(V, \eta)$ are called *rotors*. Immediate advantages to the rotor formalism are clear:

- *It is general to n dimensions, and to any metric signature.*

  Rotors describe generalised rotations,[30] depending on the metric and algebraic properties of the exponentiated unit bivector σ. If $\sigma^2 < 0$, then e^{σ} describes a Euclidean rotation; if $\sigma^2 > 0$, then e^{σ} is a hyperbolic rotation or *Lorentz boost*.

- *Vectors are distinguished from bivectors.*

  One of the subtler points about quaternions is their transformation properties under reflection. A quaternion 'vector' $v = x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}}$ reflects through the origin as $v \mapsto -v$, but a quaternion 'rotor' of the same value is invariant — vectors and pseudovectors are confused with the same kind of object. Not so in the geometric algebra: vectors are 1-vectors, and $\mathbb{R}^3$ pseudovectors are bivectors.

  The price of introducing more algebraic objects is hardly a cost but a benefit: the generalisation to arbitrary dimensions is immediate and elegant, and the geometric role of objects becomes clear.[31]

[30] a.k.a., proper orthogonal transformations

[31] See [5, 22, 25] for similarly impassioned testaments to the elegance of geometric algebra.

3.3.1. The rotor groups

Rotors, spin groups: [12, §4.2] [22, 26]

We will now see more rigorously how the rotor formalism arises. An orthogonal transformation in n dimensions may be achieved by the composition of at most n reflections.[32] A reflection may be described in the geometric algebra by conjugation with an invertible vector. For instance, the linear map

$$A \mapsto -\boldsymbol{v}A\boldsymbol{v}^{-1} \tag{3.4}$$

[32] This is the Cartan–Dieudonné theorem [27].

reflects the multivector A along the vector $\boldsymbol{v}$ — that is, across the hyperplane with normal $\boldsymbol{v}$. By composing reflections of this form, any orthogonal transformation may be built, acting on multivectors as

$$A \mapsto \pm RAR^{-1} \tag{3.5}$$

for some $R = \boldsymbol{v}_1\boldsymbol{v}_2 \cdots \boldsymbol{v}_3$, where the sign is positive for an even number of reflections, and negative for odd.

Scaling the axis of reflection $\boldsymbol{v}$ by a non-zero scalar λ does not affect the reflection map (3.4), since $\boldsymbol{v} \mapsto \lambda\boldsymbol{v}$ is cancelled out by $\boldsymbol{v}^{-1} \mapsto \lambda^{-1}\boldsymbol{v}^{-1}$. Therefore, a more direct correspondence exists between reflections and normalised vectors $\hat{\boldsymbol{v}}^2 = \pm 1$ (although there still remains an overall ambiguity in sign). For an orthogonal transformation built using normalised vectors,

$$R^{-1} = \hat{\boldsymbol{v}}_3^{-1} \cdots \hat{\boldsymbol{v}}_2^{-1}\hat{\boldsymbol{v}}_1^{-1} = \pm R^\dagger$$

since $\hat{\boldsymbol{v}}^{-1} = \pm\hat{\boldsymbol{v}}$, and hence eq. (3.5) may be written in terms of reversion instead of inversion:

$$A \mapsto \pm RAR^\dagger \tag{3.6}$$

All such elements $R^{-1} = \pm R^\dagger$ taken together form a group under the geometric product. This is called the *pin* group:

$$\mathrm{Pin}(p, q) := \left\{R \in \mathscr{G}(p, q) \mid RR^\dagger = \pm 1\right\}$$

There are two "pinors" for each orthogonal transformation, since $+R$ and $-R$ give the same map (3.6). Thus, the pin group forms a double cover of the orthogonal group $\mathrm{O}(p, q)$.

Furthermore, the even-grade elements of $\mathrm{Pin}(p, q)$ form a subgroup, called the *spin* group:

$$\mathrm{Spin}(p, q) := \left\{ R \in \mathcal{G}_+(p, q) \mid RR^\dagger = \pm 1 \right\}$$

This forms a double cover of $\mathrm{SO}(p, q)$.

Finally, the additional requirement that $RR^\dagger = 1$ defines the restricted spinor group, or the *rotor* group:

$$\begin{tikzcd}
\mathrm{Spin}^+ \arrow[r, phantom, "\subseteq"] \arrow[d, two heads] & \mathrm{Spin} \arrow[r, phantom, "\subset"] \arrow[d, two heads] & \mathrm{Pin} \arrow[d, two heads] \\
\mathrm{SO}^+ \arrow[r, phantom, "\subseteq"] & \mathrm{SO} \arrow[r, phantom, "\subset"] & \mathrm{O}
\end{tikzcd}$$

Figure 3.1.: Relationships between Lie groups associated with a geometric algebra. An arrow $A \twoheadrightarrow B$ signifies that A is a double-cover of B.

$$\mathrm{Spin}^+(p, q) := \left\{ R \in \mathcal{G}_+(p, q) \mid RR^\dagger = 1 \right\}$$

The rotor group is a double cover of the restricted special orthogonal group $\mathrm{SO}^+(p, q)$. Except for the degenerate case of $\mathrm{Spin}^+(1, 1)$, the rotor group is simply connected to the identity.

3.3.2. The bivector subalgebra

The multivector commutator product

$$A \times B := \frac{1}{2}(AB - BA) \tag{3.7}$$

enjoys several useful properties, particularly when acting on bivectors.

Lemma 9. *Commutation by a multivector A is a derivation,*

$$A \times (BC) = (A \times B)C + B(A \times C).$$

Proof. By expanding both sides,

$$\frac{1}{2}(ABC - BCA) = \frac{1}{2}(ABC - CAB + BAC - ACB).$$

□

Lemma 10. *Commutation by a bivector σ is a grade-preserving operation; i.e., $\sigma \times \langle A \rangle_k = \langle \sigma \times A \rangle_k$.*

Proof. If $A = \langle A \rangle_k$ then $A\sigma$ and σA are $\{k-2, k, k+2\}$-multivectors. The $k \pm 2$ parts are

$$\langle A \times \sigma \rangle_{k\pm 2} = \frac{1}{2}\big(\langle A\sigma \rangle_{k\pm 2} - \langle \sigma A \rangle_{k\pm 2}\big).$$

However, $\langle \sigma A \rangle_{k\pm 2} = s_{k\pm 2}\big\langle A^\dagger \sigma^\dagger \big\rangle_{k\pm 2} = -s_{k\pm 2} s_k \langle A\sigma \rangle_{k\pm 2}$ and the reversion signs[33] satisfy $s_{k\pm 2} s_k = -1$ for any k. Hence, $\langle A \times \sigma \rangle_{k\pm 2} = 0$, leaving only the grade k part, $A \times \sigma = \langle A \times \sigma \rangle_k$. □

[33] Recall from eq. (3.2) that $A^\dagger = s_k A$ for a k-vector where $s_k = (-1)^{\frac{(k-1)k}{2}}$.

A corollary of lemma 10 is that the commutator is closed on the space of bivectors, $\mathscr{G}_2$. Clearly eq. (3.7) is bilinear and satisfies the Jacobi identity, so $\mathscr{G}_2$ in fact forms a Lie algebra with the bivector commutator $\times$ as the Lie bracket.

Because the even subalgebra $\mathscr{G}_+ \supset \mathscr{G}_2$ is closed under the geometric product, the exponential $e^\sigma = 1 + \sigma + \frac{1}{2}\sigma^2 + \cdots$ of a bivector is an even multivector. Furthermore, note that the reverse $(e^\sigma)^\dagger = e^{(\sigma^\dagger)} = e^{-\sigma}$ is the inverse, and also that e^σ is continuously connected to the identity by the path $e^{\lambda\sigma}$ for $\lambda \in [0, 1]$. Therefore, $e^\sigma \in \mathrm{Spin}^+$ is a rotor, and we have a Lie algebra–Lie group correspondence shown in fig. 3.2. Thus, both the rotor groups and their Lie algebras are directly represented within the mother algebra $\mathscr{G}(p, q)$.

$$\begin{tikzcd}
\mathrm{Spin}^+(p,q) \arrow[r, two heads] & \mathrm{SO}^+(p,q) \\
\mathscr{G}_2(p,q) \arrow[u, "\exp"] \arrow[r, phantom, "\cong"] & \mathfrak{so}(p,q) \arrow[u, "\exp"]
\end{tikzcd}$$

Figure 3.2.: The Lie algebras $\mathfrak{so}(p, q)$ and $\mathscr{G}_2(p, q)$ under $\times$ are isomorphic, and are associated respectively to $\mathrm{SO}^+(p, q)$ and its universal double cover $\mathrm{Spin}^+(p, q)$.

3.4. Higher Notions of Orthogonality

As discussed at the start of this chapter, the lack of a $\mathbb{Z}$-grading means that a geometric product of blades is generally an inhomogeneous multivector. Geometrically, the grade k part of product of blades reveals the degree to which the two blades are 'orthogonal' or 'parallel', in a certain k-dimensional sense.

To see this, first consider the special case where the product of blades a and b is a homogeneous k-blade. This occurs when there exists a common orthonormal basis $\{\boldsymbol{e}_i\}$ such that

$$a = \alpha \boldsymbol{e}_{i_1} \cdots \boldsymbol{e}_{i_p} \qquad \text{and} \qquad b = \beta \boldsymbol{e}_{j_1} \cdots \boldsymbol{e}_{j_q}$$

simultaneously, for scalars α, β. Then, the product is

$$ab = \pm\alpha\beta \boldsymbol{e}_{h_1} \cdots \boldsymbol{e}_{h_k}.$$

Each pair of parallel basis vectors in a and b contributes an overall factor of $\boldsymbol{e}_i^2 = \pm 1$, and each transposition required to bring each pair together flips the overall sign.

The resulting grade k is the number of basis vectors $\boldsymbol{e}_{h_i}$ which are not common to both a and b; i.e., $\{h_i\}$ is the symmetric difference of i and j. Thus, the possible values of k are separated by steps of two, with the maximum $k = p + q$ attained when no basis vectors are common to a and b. In terms of the spans of the blades, we have

$$\begin{aligned} k &= \underbrace{\dim \operatorname{span}\{a\}}_{p} + \underbrace{\dim \operatorname{span}\{b\}}_{q} - \underbrace{2\dim(\operatorname{span}\{a\} \cap \operatorname{span}\{b\})}_{2m} \\ &\in \{|p-q|, |p-q|+2, \dots, p+q-2, p+q\}. \end{aligned} \tag{3.8}$$

Solving for the dimension of the intersection, we have

$$m = \frac{1}{2}(p + q - k).$$

Thus, the higher the grade k of the product ab, the lower the dimension m of the intersection of their spans.

We are used to the geometric meaning of two vectors being parallel or orthogonal. In terms of vector spans, they imply that the intersection is one or zero dimensional, respectively. Similarly, blades of higher grade can be 'parallel' or 'orthogonal' to varying degrees, depending on the dimension of their intersection, m.

For example, the intersection of two 2-blades may be of dimension two, one or (in four or more dimensions) zero. The notion of parallel (i.e., being a scalar multiple) remains clear ($m = 2$), but there are now two different types of orthogonality for 2-blades ($m = 1$ and $m = 0$). An example of the first type can be pictured as two planes meeting at right-angles along a line; the second type requires at least four dimensions.

Definition 20. *A p-blade a and q-blade b satisfying $ab = \langle ab \rangle_k$ are called Δ-ORTHOGONAL where $\Delta = \frac{1}{2}(k - |p - q|)$.*

Informally, Δ-orthogonality of a and b means that ab is of the Δth grade above the minimum possible grade $|p-q|$. The higher Δ, the fewer linearly independent directions are shared by (the spans of) a and b. Different cases are exemplified in table 3.1.

p	q	k	$\langle ab \rangle_k$	Δ	m	*commutativity*	*geometric interpretation of $ab = \langle ab \rangle_k$*
1	1	0	$a \cdot b$	0	1	commuting	vectors are parallel; $a \parallel b \iff a = \lambda b$
1	1	2	$a \wedge b$	1	0	anticommuting	vectors are orthogonal $a \perp b$
2	2	0	$a \cdot b$	0	2	commuting	bivectors are parallel $a = \lambda b$
2	2	2	$a \times b$	1	1	anticommuting	bivectors are at right-angles to each other
2	2	4	$a \wedge b$	2	0	commuting	bivectors are 2-*orthogonal*
1	2	1	$a \cdot b$	0	1	anticommuting	vector a lies in plane of bivector b
1	2	3	$a \wedge b$	1	0	commuting	vector a is normal to plane of bivector b
2	3	1	$a \cdot b$	0	2	commuting	bivector a lies in span of trivector b
2	3	3	$\langle ab \rangle_3$	1	1	anticommuting	a and b are 1-*orthogonal*
2	3	5	$a \wedge b$	2	0	commuting	a and b are 2-*orthogonal*

Table 3.1.: Geometric interpretation of the k-blade $ab = \langle ab \rangle_k$ where a and b are of grades p and q respectively, and where $m = \dim(\operatorname{span}\{a\} \cap \operatorname{span}\{b\})$.

Familiarity with some special cases may aid intuition when considering general products of blades. For instance, if the product of two bivectors is $\sigma_1\sigma_2 = \sigma_1 \cdot \sigma_2 + \sigma_1 \times \sigma_2$, then it is understood that σ_1 has a component parallel to σ_2, and a component which meets σ_2 at right-angles along a line of intersection. In other words, σ_1 and σ_2 are planes that intersect along a line with some angle between them (see fig. 3.3). On the other hand, if $\sigma_1\sigma_2 = \sigma_1 \wedge \sigma_2$, then the bivectors exist in orthogonal planes — a scenario requiring at least four dimensions.

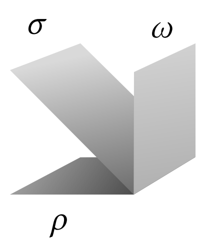


Figure 3.3.: $\{\rho, \omega\}$ are 1-orthogonal ($\rho\omega = \rho \times \omega$) and $\{\sigma, \rho\}$ have both 0- and 1-orthogonal components ($\sigma\rho = \sigma \cdot \rho + \sigma \times \rho$).

3.5. More Graded Products

All operations in the geometric algebra can be expressed in terms of the fundamental geometric product along with grade projection operators $\langle\ \rangle_k$. For example, we have seen that the wedge and inner products ($\wedge$ and $\bullet$ of definition 18) are merely combinations of multiplication and projection.

There are other similar constructions which are useful enough warrant their own definitions, inclusing left and right *contractions*.

Definition 21.

LEFT CONTRACTION $$A \rfloor B = \sum_{p,q} \Big\langle \langle A\rangle_p \langle B\rangle_q \Big\rangle_{q-p}$$

RIGHT CONTRACTION $$A \lfloor B = \sum_{p,q} \Big\langle \langle A\rangle_p \langle B\rangle_q \Big\rangle_{p-q}$$

Observe that $(A \rfloor B)^\dagger = A^\dagger \lfloor B^\dagger$, so these are in essentially the same operation — only one is viewed in a mirror.[34]

[34] I.e., every statement involving $\rfloor$ produces, under reversion, an equivalent statement involving $\lfloor$.

The fat dot product reduces to a contraction on homogeneous multivectors, depending on which multivector has the higher grade. If A is a p-vector and B a q-vector, then

$$A \bullet B = \begin{cases} A \rfloor B & p \leq q \\ A \lfloor B & q \geq p \end{cases},$$

with $A \bullet B = A \rfloor B = A \lfloor B = \langle AB\rangle$ when $p = q$. While in some expressions the grades of multivectors are obvious so that it is clear how the fat dot product acts, the contractions are arguably better behaved algebraically: the conditional comparison of grades is reincorporated into the products themselves, allowing for more useful identities to be written with fewer grade-based exceptions [24].[35]

[35] E.g., $\boldsymbol{u}A = \boldsymbol{u} \bullet A + \boldsymbol{u} \wedge A$ holds if A has zero scalar part, but $\boldsymbol{u}A = \boldsymbol{u} \rfloor A + \boldsymbol{u} \wedge A$ holds for *any* A.

Lemma 11. *For any vector $\boldsymbol{u}$ and multivector A,*

$$\boldsymbol{u} \rfloor A = \frac{1}{2}(\boldsymbol{u}A - A^\star \boldsymbol{u}), \qquad \boldsymbol{u} \wedge A = \frac{1}{2}(\boldsymbol{u}A + A^\star \boldsymbol{u}).$$

Proof. Begin by assuming A is of grade k. The geometric product contains two grades,

$$\boldsymbol{u}A = \langle \boldsymbol{u}A \rangle_{k-1} + \langle \boldsymbol{u}A \rangle_{k+1} \equiv \boldsymbol{u} \rfloor A + \boldsymbol{u} \wedge A.$$

Now consider the reversed product, and rearrange terms using the fact that $a^\dagger = s_p a$ if a is a p-vector.

$$\begin{aligned} A\boldsymbol{u} &= A \lfloor \boldsymbol{u} + A \wedge \boldsymbol{u} \\ &= s_{k-1}\, \boldsymbol{u}^\dagger \rfloor A^\dagger + s_{k+1}\, \boldsymbol{u}^\dagger \wedge A^\dagger \\ &= s_{k-1}s_k\, \boldsymbol{u} \rfloor A + s_{k+1}s_k\, \boldsymbol{u} \wedge A \end{aligned}$$

With reference to eq. (3.2), notice that $s_{k\pm1}s_k = \pm(-1)^k$. Thus,

$$A^\star \boldsymbol{u} = (-1)^k A\boldsymbol{u} = -\boldsymbol{u} \rfloor A + \boldsymbol{u} \wedge A.$$

Taking the sum and difference of $\boldsymbol{u}A$ and $A^\star\boldsymbol{u}$ as above yields the two results, respectively — at least for homogeneous A. Since the expressions are linear in A, and are written without reference to k, they extend by linearity to general multivectors. □

Lemma 12. *For a bivector σ and multivector A,*

$$\sigma A = \sigma \rfloor A + \sigma \times A + \sigma \wedge A,$$

where $a \times b = \frac{1}{2}(ab - ba)$ is the commutator product.

Proof. Suppose A is a k-vector. The geometric product with a bivector then contains three grades,

$$\sigma A = \langle \sigma A \rangle_{k-2} + \langle \sigma A \rangle_k + \langle \sigma A \rangle_{k+2} \qquad \equiv \sigma \rfloor A + \langle \sigma A \rangle_k + \sigma \wedge A.$$

Consider the reverse product,

$$A\sigma = A \lfloor \sigma + \langle A\sigma \rangle_k + A \wedge \sigma$$

reverse each term, noting that $\sigma^\dagger = -\sigma$ and $A^\dagger = s_k A$,

$$= -s_k\big(s_{k-2}\, \sigma \rfloor A + s_k\, \langle \sigma A \rangle_k + s_{k+2}\, \sigma \wedge A\big)$$

and simplify with $s_k s_{k\pm2} = -1$.

$$= \sigma \rfloor A - \langle \sigma A \rangle_k + \sigma \wedge A$$

Thus, $\langle \sigma A \rangle_k = \frac{1}{2}(\sigma A - A\sigma) \equiv \sigma \times A$, and so the result holds for homogeneous multivectors, and by linearity for general multivectors. □

Lemma 13. *For $i, j, k \geq 0$, the following conditions are equivalent.*

$$|i-j| \leq k \leq i+j, \qquad |k-i| \leq j \leq k+i, \qquad |j-k| \leq i \leq j+k.$$

Proof. There exists a triangle in the Euclidean plane with side lengths i, j, k if and only if $|i-j| \leq k \leq i+j$. By relabelling its sides, it follows that the other relations are equivalent. □

Lemma 14. *The three terms*

$$\left\langle \langle A\rangle_p \langle B\rangle_q \right\rangle_k, \qquad \left\langle \langle A\rangle_k \langle B\rangle_p \right\rangle_q, \qquad \left\langle \langle A\rangle_q \langle B\rangle_k \right\rangle_p$$

all vanish unless $|p-q| \leq k \leq p+q$.

Proof. From eq. (3.8) it follows that $\left\langle \langle A\rangle_p \langle B\rangle_q \right\rangle_k \neq 0$ implies $|p-q| \leq k \leq p+q$. By lemma 13, it also holds under permutations of the grade projections. □

Lemma 15. *For any multivectors A, B, C,*

$$(A \lfloor B) \lfloor C = A \lfloor (B \wedge C), \qquad A \rfloor (B \rfloor C) = (A \wedge B) \rfloor C.$$

Proof. It suffices to derive the identities for homogeneous multivectors; they extend by linearity to general multivectors. Thus, let (A, B, C) be multivectors of grade (a, b, c), respectively.

Consider $\langle\langle AB\rangle_k C\rangle_{a-b-c}$ and assume it to be non-zero. By lemma 14, this is zero unless $k \leq c + (a-b-c) = a-b$. However, $\langle AB\rangle_k$ is zero unless $|a-b| \leq k$, hence $k = a-b$. Therefore,

$$\langle (AB)C\rangle_{a-b-c} = \langle\langle AB\rangle_{a-b} C\rangle_{a-b-c},$$

since the only non-zero contribution from the product AB is the part of grade $a-b$.

Similarly, assume that $\langle A\langle BC\rangle_k\rangle_{a-b-c}$ is non-zero. Again by lemma 14 we have $|a-(a-b-c)| \leq k$ implying $b+c \leq k$. Since $\langle BC\rangle_k$ is zero unless $k \leq b+c$, we have $k = b+c$ exactly and

$$\langle A(BC)\rangle_{a-b-c} = \langle A\langle BC\rangle_{b+c}\rangle_{a-b-c}.$$

By associativity of the geometric product, we have shown

$$\langle\langle AB\rangle_{a-b}C\rangle_{(a-b)-c} = \langle A\langle BC\rangle_{b+c}\rangle_{a-(b+c)},$$

which is definitionally equivalent to

$$(A \lfloor B) \lfloor C = A \lfloor (B \wedge C).$$

Reversion yields the corresponding identity for left contraction. □

To summarise these results, for *any* multivectors (A, B, C), we have

$$\begin{aligned} (A \lfloor B) \lfloor C &= A \lfloor (B \wedge C), & A \rfloor (B \rfloor C) &= (A \wedge B) \rfloor C, \\ (A \rfloor B) \lfloor C &= A \rfloor (B \lfloor C), & \boldsymbol{u} \bullet (B \bullet \boldsymbol{v}) &= (\boldsymbol{u} \bullet B) \bullet \boldsymbol{v}. \end{aligned}$$

The last equation is a specialisation of the upper right, which in particular means that parentheses are unnecessary when defining the components of a bivector $F = F^{ij}\boldsymbol{e}_i \wedge \boldsymbol{e}_j$ with the expression $F_{ij} = \boldsymbol{e}_i \bullet F \bullet \boldsymbol{e}_j$.

Chapter 4.

The Algebra of Spacetime

Special relativity is geometry with a Lorentzian signature. The SPACETIME ALGEBRA (STA) is the name given to the geometric algebra of a Minkowski vector space, $\mathscr{G}(1,3) \equiv \mathscr{G}(\mathbb{R}^4, \eta)$, where $\eta = \pm\mathrm{diag}(-{+}{+}{+})$. Other introductory material on the STA can be found in [28–30].

We denote the standard vector basis by $\{\boldsymbol{\gamma}_\mu\}$, where Greek indices run over $\{0,1,2,3\}$. This is a deliberate allusion to the Dirac γ-matrices, whose algebra is isomorphic to the STA – however, the $\boldsymbol{\gamma}_\mu \in \mathbb{R}^{1+3}$ of STA are real, genuine spacetime vectors. A basis for the entire 2^4-dimensional STA is then

$$\underset{+}{\overset{\text{1 scalar}}{\{1\}}} \cup \overset{\text{4 vectors}}{\{\underset{+}{\boldsymbol{\gamma}_0},\ \underset{-}{\boldsymbol{\gamma}_i}\}} \cup \overset{\text{6 bivectors}}{\{\underset{+}{\boldsymbol{\gamma}_0\boldsymbol{\gamma}_i},\ \underset{-}{\boldsymbol{\gamma}_j\boldsymbol{\gamma}_k}\}} \cup \overset{\text{4 trivectors}}{\{\underset{-}{\boldsymbol{\gamma}_0\boldsymbol{\gamma}_j\boldsymbol{\gamma}_k},\ \underset{+}{\boldsymbol{\gamma}_1\boldsymbol{\gamma}_2\boldsymbol{\gamma}_3}\}} \cup \overset{\text{1 pseudoscalar}}{\{\underset{-}{\mathbb{I} \coloneqq \boldsymbol{\gamma}_0\boldsymbol{\gamma}_1\boldsymbol{\gamma}_2\boldsymbol{\gamma}_3}\}}$$

where lowercase Latin indices range over spacelike components, $\{1,2,3\}$. Blades shown on the left-hand side of $\{\quad,\quad\}$ are called TIMELIKE, and those in on right-hand side SPACELIKE. The sign below each basis blade shows its signature (the sign of its scalar square). Multivectors of any kind which square to zero are called NULL.

The pseudoscalar and duality

The right-handed unit pseudoscalar $\mathbb{I}$ represents an oriented unit 4-volume. It anticommutes with odd elements of the STA (vectors and trivectors) and commutes with even elements (bivectors and (pseudo)scalars).

Since $\mathbb{I}^2 = -1$, the scalar–pseudoscalar plane $\mathscr{G}_{0,4}(1,3) = \text{span}_{\mathbb{R}}\{1, \mathbb{I}\}$ is isomorphic to the complex plane $\mathbb{C}$. Thus, for the sake of computation, operations on $\{0,4\}$-multivectors may be regarded as operations on complex numbers. In particular, we define the principal root $\sqrt{a}$ of a $\{0,4\}$-multivector $a \in \mathscr{G}_{0,4}(1,3)$ in the same way as it is defined in $\mathbb{C}$ with a branch cut at $\theta = \pi$. It is worth emphasising that there are many square roots of -1 in the spacetime algebra, each with distinct geometrical meanings.[36] We single to single out $\sqrt{-1} = \mathbb{I}$ as 'the' principal root as this proves to be useful notationally.[37]

[36] E.g., the spacelike bivector $(\gamma_i\gamma_j)^2 = -1$ represents a directed spacelike plane.

[37] In electromagnetism, the imaginary unit i often represents the volume element $\mathbb{I}$. E.g., in the Riemann–Silberstein vector [31], both i and $\mathbb{I}$ play roles similar to the Hodge dual [30].

As in section 3.2.3, Hodge duality is accomplished by (right) multiplication by the volume element. In particulat, this establishes a duality between vectors and trivectors, and between spacelike and timelike bivectors.

4.1. The space/time Split

While we actually live in $\mathbb{R}^{1,3}$ spacetime, to any particular observer it appears that space is $\mathbb{R}^3$ with a separate scalar time parameter. This is reflected in the fact that $\mathscr{G}_+(1,3)$ and $\mathscr{G}(3)$ are isomorphic, from lemma 7. In fact, there is a separate isomorphism associated to each timelike direction, corresponding to each inertial observer's personal spacetime split. Such a SPACE/TIME SPLIT identifies *even* multivectors in the spacetime algebra $\mathscr{G}_+(1,3)$ with $\mathscr{G}(3)$ multivectors, providing an efficient, purely algebraic method for switching between inertial frames [28].

Let K be an inertial observer and for simplicity choose the standard basis $\{\gamma_\mu\}$ so that γ_0 is the instantaneous velocity of the K frame. The three RELATIVE VECTORS $\vec{\sigma}_i := \gamma_i\gamma_0$ form a vector basis for $\mathscr{G}(3)$, since the $\gamma_i\gamma_0$ indeed satisfy $\vec{\sigma}_i^2 = -\gamma_i^2\gamma_0^2 = 1$ and $\vec{\sigma}_i\vec{\sigma}_j = -\vec{\sigma}_j\vec{\sigma}_i$ for $i \neq j$. Because of the dependence on the frame's velocity vector $\boldsymbol{\gamma}_0$, the relative vectors $\vec{\sigma}_i$ are particular to the K frame. With respect to the K frame, we may view $\mathscr{G}(3) \subset \mathscr{G}(1,3)$ as embedded in the STA, allowing us to consider multivectors as belonging to both spaces. Note that the same volume element $\mathbb{I} = \vec{\sigma}_1\vec{\sigma}_2\vec{\sigma}_3 = \boldsymbol{\gamma}_0\boldsymbol{\gamma}_1\boldsymbol{\gamma}_2\boldsymbol{\gamma}_3$ is shared by the algebras.

For example a spacetime bivector $F = F^{\mu\nu}\boldsymbol{\gamma}_\mu\boldsymbol{\gamma}_\nu$ may be separated into timelike F^{i0} and spacelike F^{ij} components with respect to the K frame and viewed as a $\{1,2\}$-multivector in $\mathscr{G}(3)$,

$$F = F^{i0}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_0 + F^{ij}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_j = E^i\vec{\sigma}_i + B^i\mathbb{I}\vec{\sigma}_i = \vec{E} + \mathbb{I}\vec{B}, \tag{4.1}$$

where we use $\boldsymbol{\gamma}_i\boldsymbol{\gamma}_j = (\boldsymbol{\gamma}_i\boldsymbol{\gamma}_0)(\boldsymbol{\gamma}_j\boldsymbol{\gamma}_0) = -\vec{\sigma}_i\vec{\sigma}_j = -\varepsilon_{ijk}\mathbb{I}\vec{\sigma}_k$. Note that the relativistic representation F is *equal* to the frame-dependent representation — they are the same spacetime object. Equation (4.1) performs the frame-dependent decomposition of a spacetime bivector (or "2-form") into two $\mathbb{R}^3$ vectors familiar from electromagnetic theory.

Of particular interest are space/time splits on bivector generators associated to rotors. A proper orthochronous Lorentz transformation $\Lambda \in \mathrm{SO}^+(1,3)$ acts as a 'sandwich' product $\Lambda(A) = e^\sigma A e^{-\sigma}$, where the rotor $e^\sigma \in \mathrm{Spin}^+(1,3)$ is generated by a spacetime bivector $\sigma \in \mathscr{G}_2(1,3)$. This bivector σ can be represented in the K frame as

$$\sigma = \frac{1}{2}(\xi^i\boldsymbol{\gamma}_i + \theta^i\mathbb{I}\boldsymbol{\gamma}_i)\boldsymbol{\gamma}_0 = \frac{1}{2}(\boldsymbol{\xi} + \mathbb{I}\boldsymbol{\theta}) \tag{4.2}$$

where $\boldsymbol{\xi} = \xi^i\vec{\sigma}_i \in \mathscr{G}_1(3)$ is a rapidity vector and $\mathbb{I}\boldsymbol{\theta} \in \mathscr{G}_2(3)$ is a rotation bivector.

4.1.1. On the choice of metric signature

Both metric signatures $\eta = \mathrm{diag}(-+++)$ and $\eta = \mathrm{diag}(+---)$ are appropriate for relativistic physics, and both are used in the literature. While the overall physics is agnostic to this choice, expressions written in the STA are generally not independent of the overall sign. It is a useful reference to note what changes and what is constant under both choices.

One of the most important properties of the space/time split is the union of the $\mathscr{G}(3)$ and $\mathscr{G}_+(1,3)$ volume elements, $\mathbb{I} = \vec{\sigma}_1\vec{\sigma}_2\vec{\sigma}_3 = \boldsymbol{\gamma}_0\boldsymbol{\gamma}_1\boldsymbol{\gamma}_2\boldsymbol{\gamma}_3$.

If this equality is to hold, then switching the metric signature is concomitant with a switch in sign of the relative vectors, $\vec{\sigma}_i \mapsto -\vec{\sigma}_i$. Another noticable difference is in the space/time split of a position vector $\boldsymbol{X} \in \mathscr{G}_1(1,3)$ into components $\boldsymbol{X}^0 = ct$ and $(\boldsymbol{X}^i) = \vec{x}$, achieved by multiplication with the frame velocity $\boldsymbol{\gamma}_0$. For example, the equations

$$\boldsymbol{X}\boldsymbol{\gamma}_0 = ct + \vec{x}, \qquad \boldsymbol{\gamma}_0\boldsymbol{X} = ct - \vec{x}$$

hold in the $(+---)$ signature, but both change by an overall sign in the $(-+++)$ signature.[38] Both these points are summarised in table 4.1.

[38] In all cases, reversion $\boldsymbol{X}\boldsymbol{\gamma}_0 \mapsto (\boldsymbol{X}\boldsymbol{\gamma}_0)^\dagger = \boldsymbol{\gamma}_0\boldsymbol{X}$ simply negates the spacetime bivector part, $\vec{x} \to -\vec{x}$.

signature	*preferred* $\vec{\sigma}_i$	$\boldsymbol{\gamma}_0\boldsymbol{X}$	$\boldsymbol{X}\boldsymbol{\gamma}_0$
$(+---)$	$\vec{\sigma}_i := \boldsymbol{\gamma}_i\boldsymbol{\gamma}_0$	$ct - \vec{x}$	$ct + \vec{x}$
$(-+++)$	$\vec{\sigma}_i := \boldsymbol{\gamma}_0\boldsymbol{\gamma}_i$	$-ct + \vec{x}$	$-ct - \vec{x}$

Table 4.1.: Comparison of space/time split in each metric signature. The spacetime vector X has contravariant components $X^0 = ct$ and $(X^i) = \vec{x}$ in the $\boldsymbol{\gamma}_0$-frame. Relative vectors are defined so that the spacetime volume element and volume element under a space/time split are equal.

A choice of metric sign may be avoided by using sign-agnostic expressions. An invariant definition of relative vectors and their duals in the $\boldsymbol{\gamma}_0$-frame is

$$\vec{\sigma}_i := \boldsymbol{\gamma}_i\boldsymbol{\gamma}^0, \qquad \vec{\sigma}^i = \boldsymbol{\gamma}_0\boldsymbol{\gamma}^i.$$

These satisfy $\mathbb{I} = \vec{\sigma}_1\vec{\sigma}_2\vec{\sigma}_3 = \boldsymbol{\gamma}_0\boldsymbol{\gamma}_1\boldsymbol{\gamma}_2\boldsymbol{\gamma}_3$ and $\mathbb{I}^{-1} = \vec{\sigma}^1\vec{\sigma}^2\vec{\sigma}^3 = \boldsymbol{\gamma}^0\boldsymbol{\gamma}^1\boldsymbol{\gamma}^2\boldsymbol{\gamma}^3$ in either signature. In particular, the following expressions hold in either signature, and are useful when performing space/time splits.

$$\begin{aligned} \boldsymbol{\gamma}^0\boldsymbol{X} &= ct - \vec{x} & \boldsymbol{X}\boldsymbol{\gamma}^0 &= ct + \vec{x} \\ \boldsymbol{\gamma}_0\,\boldsymbol{\partial} &= \frac{1}{c}\frac{\partial}{\partial t} + \vec{\nabla} & \boldsymbol{\partial}\,\boldsymbol{\gamma}_0 &= \frac{1}{c}\frac{\partial}{\partial t} - \vec{\nabla} \end{aligned}$$

Here, the spacetime vector derivative $\boldsymbol{\partial} = \boldsymbol{\gamma}^\mu\partial_\mu$ decomposes into a scalar time derivative $\partial_0 = c^{-1}\partial_t$ and the spatial derivative $\vec{\nabla} = \vec{\sigma}^i\partial_i$.

4.2. The Invariant Bivector Decomposition

There is a clear analogy between the space/time split of a bivector (4.1), into spacelike and timelike components, and the Cartesian form of a complex number, $x + iy$, into real and imaginary parts. This similarity can be made more precise: just as we may express complex numbers in polar form $re^{i\phi} = x+iy$, we may use the invariant bivector decomposition to write $\rho e^{\mathbb{I}\sigma} = E + \mathbb{I}B$, since $\mathbb{I}^2 = i^2 = -1$. This is distinct from the space/time split in that it is frame *independent*, and the bivector E is not necessarily timelike, and so need not correspond to any relative vector $\vec{E} \in \mathscr{G}_1(3)$.

Non-null spacetime bivectors $\sigma \in \mathscr{G}_2(1,3)$ may be *normalised*, in the sense that there always exists some $N_\sigma \in \mathscr{G}_{0,4}(1,3)$ such that

$$\sigma = N_\sigma \hat{\sigma} = \hat{\sigma} N_\sigma \quad \text{where} \quad \hat{\sigma}^2 = 1.$$

In the null case $\sigma^2 = 0$, we let $\hat{\sigma}^2 = 0$ instead. This is possible because the square of a bivector is a $\{0,4\}$-multivector (lemma 6), which always has a principle square root (since $\mathscr{G}_{0,4}(1,3) \cong \mathbb{C}$). Explicitly, let $\sigma^2 = \alpha + \mathbb{I}\beta = \rho^2 e^{2\mathbb{I}\phi}$ for scalars $\alpha, \beta, \rho, \phi$, so that

$$N_\sigma \coloneqq \sqrt{\sigma^2} = \rho e^{\mathbb{I}\phi},$$

assuming without loss of generality that $\rho > 0$ and $\phi \in (-\pi/2, \pi/2]$. Thus, the INVARIANT BIVECTOR DECOMPOSITION

$$\sigma = \rho e^{\mathbb{I}\phi}\hat{\sigma} = \underbrace{(\rho\cos\phi)\hat{\sigma}}_{\sigma_+} + \underbrace{(\rho\sin\phi)\mathbb{I}\hat{\sigma}}_{\sigma_-}$$

separates σ into commuting parts, $[\sigma_+, \sigma_-] = 0$, each of which satisfy $\pm\sigma_\pm^2 > 0$. This makes it a useful device for algebraic manipulations. Furthermore, the decomposition is unique, and does not depend on any particular space/time split.

The decomposition can be used to show the non-injectivity of the exponential map in the STA. Take some bivector written in decomposed form, $\sigma = \lambda_+\hat{\sigma} + \lambda_-\mathbb{I}\hat{\sigma}$. For $n \in \mathbb{Z}$, each bivector in the family

$$\sigma_n = \lambda_+\hat{\sigma} + (\lambda_- + n\pi)\mathbb{I}\hat{\sigma}$$

exponentiates to the same rotor, up to an overall sign:

$$e^{\sigma_n} = e^{\sigma_0}e^{n\pi\mathbb{I}\hat{\sigma}} = (-1)^n e^{\sigma_0} \tag{4.3}$$

Note that $e^{\hat{\sigma}+\mathbb{I}\hat{\sigma}} = e^{\hat{\sigma}}e^{\mathbb{I}\hat{\sigma}}$ since $[\hat{\sigma}, \mathbb{I}\hat{\sigma}] = 0$. All the rotors in eq. (4.3) correspond to the same $\mathrm{SO}^+(1,3)$ Lorentz transformation. Equation (4.3) shows that every Lorentz rotor $\pm e^{\sigma_0}$ is equal to a pure bivector exponential e^{σ_n} with a shifted rotational part $\lambda_- \mapsto \lambda_- + n\pi$.

4.3. Lorentz Conjugacy Classes

As shown above, every proper Lorentz transformation $\Lambda \in \mathrm{SO}^+(1,3)$ is generated by a bivector exponential $\Lambda(\boldsymbol{u}) = e^{\sigma}\boldsymbol{u}e^{-\sigma}$. This rotor formulation of the Lorentz group makes some of its more subtle properties clear, including its decomposition into conjugacy classes.

Definition 22. *The* CONJUGACY CLASS *of a group element $g \in G$ is the set*

$$[g] := \left\{hgh^{-1} \,\middle|\, h \in G\right\} = \{g' \in G \mid g' \sim g\}$$

of elements conjugate[39] *to g.*

[39] Group elements $g \sim g'$ are conjugate iff there extists $h \in G$ such that $g = hg'h^{-1}$.

Since conjugacy is an equivalence relation, the conjugacy classes partition the group G.

In the case of the proper Lorentz group, the set of conjugacy classes further partitions into five categories, or 'kinds', according to basis-invariant properties of the constituent Lorentz transformations. Using the STA, the 'kind' of a Lorentz transformation (or of its associated rotors) is given by basic properties of its generating bivector.[40]

[40] One rotor has many generating bivectors, but any one will do.

Definition 23. *Let $\sigma \in \mathscr{G}_2(1,3)$ be a bivector. If σ^2 is a scalar, then σ is called*

- TRIVIAL *if $\sigma = 0$; and if $\sigma \neq 0$,*
- ELLIPTIC *if $\sigma^2 < 0$;*
- PARABOLIC *if $\sigma^2 = 0$;*
- HYPERBOLIC *if $\sigma^2 > 0$; and*
- LOXODROMIC *if $\sigma^2 = \alpha + \mathbb{I}\beta$ is not a scalar but a $\{0,4\}$-multivector.*

Lemma 16. *The square of a bivector is constant within each conjugacy class.*

Proof. Let $\Lambda : \boldsymbol{u} \mapsto e^{\sigma}\boldsymbol{u}e^{-\sigma}$ be a proper Lorentz transformation, and consider its conjugation with some other transformation Γ,

$$\Gamma\Lambda\Gamma^{-1} : \boldsymbol{u} \mapsto e^{\rho}e^{\sigma}e^{-\rho}\boldsymbol{u}e^{-\rho}e^{-\sigma}e^{\rho}.$$

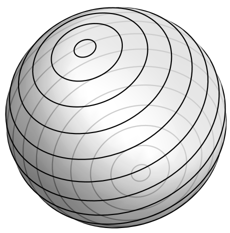
Note that $e^{\rho}e^{\sigma}e^{-\rho} = e^{e^{\rho}\sigma e^{-\rho}}$ by the automorphism property of rotor application. Therefore, conjugacy of $\Lambda \sim \Gamma\Lambda\Gamma^{-1}$ translates to bivectors as

$$\sigma \sim \sigma' \coloneqq e^{\rho}\sigma e^{-\rho}$$

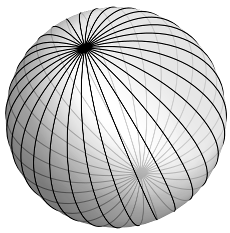
for some ρ. Hence, the conjugate bivectors have common square,

$$\sigma'^2 = (e^{\rho}\sigma e^{-\rho})^2 = e^{\rho}\sigma^2 e^{-\rho} = \sigma^2$$

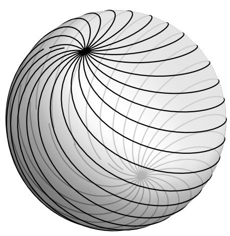
since $e^{\pm\rho}$ commutes with the $\{0,4\}$-multivector σ^2. □



(a) Elliptical



(b) Hyperbolic



(c) Loxodromic

Figure 4.1.: Lorentz transformations on the celestial sphere, taking curves to themselves.

Corollary 1. *Conjugacy classes of* $\mathrm{SO}^+(1,3)$ *fall into the five categories in definition 23 by considering the generating bivector of any representative Lorenz rotor.*

Elliptical Lorentz transformations are *rotations*, whose rotors are generated by spacelike 2-blades; hyperbolic transformations are *boosts*, with timelike 2-blades generators. Parabolic transformations are sometimes called *null rotations*, and fall in between the previous two, with null 2-blades as generators.

The final class of loxodromic transformations are a combination of a rotation and a boost where the axis of rotation is parallel with the boost direction (in a particular frame). A loxodromic generator is *not* a 2-blade, but a bivector comprising mutually 2-orthogonal[41] 2-blades, one timelike and one spacelike.

[41] in the sense of definition 20, section 3.4

These can be helpfully visualised by making use of the isomorphism $\mathrm{SO}^+(1,3) \cong \mathrm{Aut}(\mathbb{C} \cup \{\infty\})$ of the Lorentz group with the Möbius group of conformal transformations on the sphere. An observer undergoing a change of frame will see the celestial sphere transform conformally, as in fig. 4.1.

Chapter 5.

Composition of Rotors in terms of their Generators

In studying proper orthogonal transformations, it is often easier to represent them in terms of their generators $\sigma_i \in \mathscr{G}(p,q)$ which belong to the Lie algebra $\mathfrak{so}(p,q)$. A fundamental question is how such transformations compose in terms of these generators: "given σ_1 and σ_2, what is σ_3 such that $e^{\sigma_1}e^{\sigma_2} = e^{\sigma_3}$?" This is of theoretical interest and is useful practically when representing transformations in terms of their generators is cheaper. One may use the Baker–Campbell–Hausdorff–Dynkin[42] (BCHD) formula $\sigma_1 \odot \sigma_2 \coloneqq \log(e^{\sigma_1}e^{\sigma_2})$ which is well studied in general Lie theory [32]. However, the general BCHD formula

[42] Often simply Baker–Campbell–Hausdorff and permutations thereof.

$$a \odot b = a + b + \frac{1}{2}[a,b] + \frac{1}{12}[a,[a,b]] + \frac{1}{12}[[a,b],b] + \cdots \tag{5.1}$$

involves an infinite series of nested commutators and may not obviously admit a useful closed form.

In the case of Lorentz transformations $\mathrm{SO}^+(1,3)$, some closed-form expressions for eq. (5.1) have been found using a 2-form representation of $\mathfrak{so}(1,3)$ [33, 34], but the expressions are complicated and do not clearly reduce to well-known formulae in, for example, the special cases of pure rotations or pure boosts. The rotor formalism of geometric algebra leads to an elegant closed form of eq. (5.1) which, in the case of Lorentzian spacetime, is inexpensive to compute.

5.1. A Geometric BCHD Formula

Suppose $\sigma \in \mathscr{G}_2(p,q)$ is a bivector in a geometric algebra of dimension $p+q \leq 4$. By their definitions as formal power series, we have $e^\sigma = \cosh\sigma + \sinh\sigma$, where 'cosh' involves even powers of σ and 'sinh' odd powers. For convenience, define the linear projections onto SELF-REVERSE and ANTI-SELF-REVERSE parts respectively as

$$\{A\} \coloneqq \frac{1}{2}\big(A + A^\dagger\big) \qquad \text{and} \qquad [\![A]\!] \coloneqq \frac{1}{2}\big(A - A^\dagger\big). \tag{5.2}$$

Since any bivector obeys $\sigma^\dagger = -\sigma$, it follows that $(e^\sigma)^\dagger = e^{-\sigma} = \cosh\sigma - \sinh\sigma$. Using the notation (5.2), the self-reverse and anti-self-reverse projections of e^σ are $\{e^\sigma\} = \cosh\sigma$ and $[\![e^\sigma]\!] = \sinh\sigma$, respectively. Furthermore, these two projections commute, and so

$$[\![e^\sigma]\!]\{e^\sigma\}^{-1} = \{e^\sigma\}^{-1}[\![e^\sigma]\!] = \frac{[\![e^\sigma]\!]}{\{e^\sigma\}} = \tanh\sigma$$

which leads to an expression for the logarithm of any rotor $\mathscr{R} = \pm e^\sigma$.

$$\sigma = \log(\mathscr{R}) = \operatorname{arctanh}\left(\frac{[\![\mathscr{R}]\!]}{\{\mathscr{R}\}}\right) \tag{5.3}$$

Note that the overall sign of the rotor is not recovered, and $\log(+\mathscr{R}) = \log(-\mathscr{R})$ according to eq. (5.3). However, this does not affect the Lorentz transformation $\mathrm{R} \in \mathrm{SO}^+(p,q)$, since it is defined by $R(\boldsymbol{u}) = \mathscr{R}\boldsymbol{u}\mathscr{R}^\dagger$. The exact sign can be recovered by considering the relative signs of $[\![\mathscr{R}]\!]$ and $\{\mathscr{R}\}$, as in [35, §5.3].

From eq. (5.3) we may derive a BCHD formula by substituting $\mathscr{R} = e^{\sigma_1}e^{\sigma_2}$ for any two bivectors $\sigma_i \in \mathscr{G}_2(p,q)$. Using the shorthand $\mathrm{C}_i \coloneqq \cosh\sigma_i$ and $\mathrm{S}_i \coloneqq \sinh\sigma_i$, the composite rotor is

$$\mathscr{R} = e^{\sigma_1}e^{\sigma_2} = (\mathrm{C}_1 + \mathrm{S}_1)(\mathrm{C}_2 + \mathrm{S}_2) = \mathrm{C}_1\mathrm{C}_2 + \mathrm{S}_1\mathrm{C}_2 + \mathrm{C}_1\mathrm{S}_1 + \mathrm{S}_1\mathrm{S}_2.$$

For $p+q < 4$, any even function of a bivector (such as C_i) is a scalar, and for $p+q = 4$, is a $\{0,4\}$-multivector $\alpha + \beta\mathbb{I}$. In either case, the C_i commute with even multivectors, so $[\mathrm{C}_i, \mathrm{C}_j] = [\mathrm{C}_i, \mathrm{S}_j] = 0$. Therefore, the self-reverse and anti-self-reverse parts are

$$\{\mathscr{R}\} = \mathrm{C}_1\mathrm{C}_2 + \frac{1}{2}\{\mathrm{S}_1, \mathrm{S}_2\} \quad \text{and} \quad [\![\mathscr{R}]\!] = \mathrm{S}_1\mathrm{C}_2 + \mathrm{C}_1\mathrm{S}_2 + \frac{1}{2}[\mathrm{S}_1, \mathrm{S}_2]. \tag{5.4}$$

Hence, from eq. (5.3) we obtain an explicit BCHD formula.

Theorem 3 (rotor BCHD formula). *If $\sigma_1, \sigma_2 \in \mathscr{G}_2(p,q)$ are bivectors in $p+q \leq 4$ dimensions, then $e^{\sigma_1}e^{\sigma_2} = \pm e^{\sigma_1 \odot \sigma_2}$ where*

$$\sigma_1 \odot \sigma_2 = \operatorname{arctanh}\left(\frac{\mathrm{T}_1 + \mathrm{T}_2 + \frac{1}{2}[\mathrm{T}_1, \mathrm{T}_2]}{1 + \frac{1}{2}\{\mathrm{T}_1, \mathrm{T}_2\}}\right) \tag{5.5}$$

where we abbreviate $\mathrm{T}_i := \tanh \sigma_i$. Note that this satisfies the rotor equation with an overall ambiguity in sign.

We may wish to express eq. (5.5) in terms of geometrically significant products instead of (anti)commutators. A bivector product is generally a $\{0,2,4\}$-multivector

$$\begin{aligned} ab &= \langle ab \rangle_0 + \langle ab \rangle_2 + \langle ab \rangle_4 \\ &= a \cdot b + a \times b + a \wedge b. \end{aligned} \tag{5.6}$$

where $a \times b = \langle ab \rangle_2 = \frac{1}{2}[a,b]$ is the commutator product. We may then write eq. (5.5) so that the grade of each term is explicit:

$$\sigma_1 \odot \sigma_2 = \operatorname{arctanh}\left(\frac{\mathrm{T}_1 + \mathrm{T}_2 + \mathrm{T}_1 \times \mathrm{T}_2}{1 + \mathrm{T}_1 \cdot \mathrm{T}_2 + \mathrm{T}_1 \wedge \mathrm{T}_2}\right) \tag{5.7}$$

The numerator is a bivector, while the denominator contains scalar ($\mathrm{T}_1 \cdot \mathrm{T}_2$) and 4-vector ($\mathrm{T}_1 \wedge \mathrm{T}_2$) terms.

5.1.1. Zassenhaus-type formulae

It is interesting to generalise the BCHD formula (5.1) to three rotors $e^{\sigma_1}e^{\sigma_2}e^{\sigma_3} = e^{\sigma}$ in an algebra $\mathscr{G}(p,q)$ with $p+q \leq 4$. A solution to this rotor equation is

$$\sigma = \log(\pm e^{\sigma}) = \operatorname{arctanh}\left(\frac{[\![e^{\sigma_1}e^{\sigma_2}e^{\sigma_3}]\!]}{\{e^{\sigma_1}e^{\sigma_2}e^{\sigma_3}\}}\right),$$

by eq. (5.3).

We will find it convenient to define the ANTICOMMUTATOR PRODUCT $A \curlywedge B := \frac{1}{2}\{A,B\}$ to complement the commutator product $A \times B$. The symbol "$\curlywedge$" is motivated by the fact that, for bivectors, we have $\sigma \curlywedge \rho = \sigma \cdot \rho + \sigma \wedge \rho$ and thus

$$\sigma \curlywedge \rho \coloneqq \frac{1}{2}(\sigma\rho + \rho\sigma) = \{\sigma\rho\}, \qquad \sigma \times \rho \coloneqq \frac{1}{2}(\sigma\rho - \rho\sigma) = [\![\sigma\rho]\!]. \tag{5.8}$$

Because $e^{\sigma_1}e^{\sigma_2}e^{\sigma_3} \in \mathscr{G}_+(p,q)$ is an even multivector, the anti-self-reverse projection is exactly the bivector part, $[\![e^{\sigma_1}e^{\sigma_2}e^{\sigma_3}]\!] = \langle e^{\sigma_1}e^{\sigma_2}e^{\sigma_3}\rangle_2$, and the self-reverse projection is the $\{0,4\}$-multivector part.[43] Decomposing $e^{\sigma_i} = \mathrm{C}_i + \mathrm{S}_i$, we find 2^3 terms which separate into

[43] Recall $A^\dagger = s_k A$ for a k-vector A where $(s_1 \cdots s_4) = (++--)$.

$$\begin{aligned}
[\![e^{\sigma_1}e^{\sigma_2}e^{\sigma_3}]\!] &= \mathrm{S}_1\mathrm{C}_2\mathrm{C}_3 + \mathrm{C}_1\mathrm{S}_2\mathrm{C}_3 + \mathrm{C}_1\mathrm{C}_2\mathrm{S}_3 + (\mathrm{C}_1\mathrm{S}_2 + \mathrm{S}_1\mathrm{C}_2) \times \mathrm{S}_3 + (\mathrm{S}_1 \times \mathrm{S}_2)\mathrm{C}_3 + [\![\mathrm{S}_1\mathrm{S}_2\mathrm{S}_3]\!], \\
\{e^{\sigma_1}e^{\sigma_2}e^{\sigma_3}\} &= \mathrm{C}_1\mathrm{C}_2\mathrm{C}_3 + (\mathrm{C}_1\mathrm{S}_2 + \mathrm{S}_1\mathrm{C}_2) \curlywedge \mathrm{S}_3 + (\mathrm{S}_1 \curlywedge \mathrm{S}_2)\mathrm{C}_3 + \{\mathrm{S}_1\mathrm{S}_2\mathrm{S}_3\}.
\end{aligned}$$

The $\{0,4\}$-multivectors C_i commute with the bivectors S_i, and products of C_i and S_j are themselves bivectors. Therefore, terms containing one S_i factor are bivectors, and terms containing two S_i factors, such as $\mathrm{S}_1\mathrm{S}_2\mathrm{C}_3$, are products of bivectors, or $\{0,2,4\}$-multivectors. These terms are split into bivectors $(\mathrm{S}_1 \times \mathrm{S}_2)\mathrm{C}_3$ and $\{0,4\}$-multivectors $(\mathrm{S}_1 \curlywedge \mathrm{S}_2)\mathrm{C}_3$.

Cancelling factors of $\mathrm{C}_1\mathrm{C}_2\mathrm{C}_3$, we then have

$$\frac{[\![e^{\sigma_1}e^{\sigma_2}e^{\sigma_3}]\!]}{\{e^{\sigma_1}e^{\sigma_2}e^{\sigma_3}\}} = \frac{\mathrm{T}_1 + \mathrm{T}_2 + \mathrm{T}_3 + (\mathrm{T}_1 + \mathrm{T}_2) \times \mathrm{T}_3 + \mathrm{T}_1 \times \mathrm{T}_2 + [\![\mathrm{T}_1\mathrm{T}_2\mathrm{T}_3]\!]}{1 + (\mathrm{T}_1 + \mathrm{T}_2) \curlywedge \mathrm{T}_3 + \mathrm{T}_1 \curlywedge \mathrm{T}_2 + \{\mathrm{T}_1\mathrm{T}_2\mathrm{T}_3\}} \tag{5.9}$$

where $\mathrm{T}_i \coloneqq \tanh\sigma_i$. The next lemma is used to rewrite the rightmost terms with (anti)commutator products (5.8).

Lemma 17. *For any bivectors $\sigma, \rho, \omega \in \mathscr{G}_2(p,q)$ where $p + q \leq 4$,*

$$[\![\sigma\rho\omega]\!] = (\sigma \curlywedge \rho) \curlywedge \omega + (\sigma \times \rho) \times \omega, \qquad \{\sigma\rho\omega\} = (\sigma \times \rho) \curlywedge \omega.$$

Proof. Observe that $[\![\sigma\rho\omega]\!] = \langle\sigma\rho\omega\rangle_2$ since $\sigma\rho\omega$ is a $\{0,2,4\}$-multivector, of which only the bivector part is anti-self-reverse. Using associativity and linearity,

$$\langle\sigma\rho\omega\rangle_2 = \langle(\sigma \curlywedge \rho)\omega\rangle_2 + \langle(\sigma \times \rho)\omega\rangle_2 = (\sigma \curlywedge \rho)\omega + (\sigma \times \rho) \times \omega.$$

The product $(\sigma \wedge \rho)\omega = (\sigma \wedge \rho) \wedge \omega$ is between a $\{0,4\}$-multivector and a bivector, which may only contain bivector components. The product $(\sigma \times \rho)\omega$ is between two bivectors, having bivector part $(\sigma \times \rho) \times \omega$.

Similarly, note that

$$\{\sigma\rho\omega\} = \langle(\sigma \wedge \rho)\omega\rangle_{0,4} + \{(\sigma \times \rho)\omega\} = (\sigma \times \rho) \wedge \omega,$$

where the first term vanishes since $(\sigma \wedge \rho)\omega$ is a bivector. □

This allows us to collect the terms in eq. (5.9) as

$$\frac{[\![e^{\sigma_1}e^{\sigma_2}e^{\sigma_3}]\!]}{\{e^{\sigma_1}e^{\sigma_2}e^{\sigma_3}\}} = \frac{\mathrm{T}_{12} + \mathrm{T}_3 + \mathrm{T}_{12} \times \mathrm{T}_3 + (\mathrm{T}_1 \wedge \mathrm{T}_2) \wedge \mathrm{T}_3}{1 + \mathrm{T}_{12} \wedge \mathrm{T}_3 + \mathrm{T}_1 \wedge \mathrm{T}_2}$$

where $\mathrm{T}_{12} \coloneqq \mathrm{T}_1 + \mathrm{T}_2 + \mathrm{T}_1 \times \mathrm{T}_2$. This leads us to the following result.

Lemma 18. *For bivectors $\sigma_i \in \mathscr{G}_2(p,q)$ with $p+q \leq 4$,*

$$e^{\sigma_1+\sigma_2} = e^{\sigma_1}e^{\sigma_2}e^{\rho}$$

where

$$\begin{aligned}
\rho &= \operatorname{arctanh}\left(\frac{F - R - R \times F + S \wedge F}{1 - R \wedge F + S}\right), \\
F &= \tanh(\sigma_1 + \sigma_2), \\
R &= \tanh(\sigma_1) \times \tanh(\sigma_2) + \tanh(\sigma_1) + \tanh(\sigma_2), \\
S &= \tanh(\sigma_1) \wedge \tanh(\sigma_2).
\end{aligned}$$

{TO DO: First order corrections? Don't know where to go with this.}

5.1.2. In low dimensions: Rodrigues' rotation formula

It is illustrative to see how the BCHD formula (5.5) reduces in low-dimensional special cases. Indeed, in two dimensions, all bivectors are scalar multiples of $\mathbb{I} = \boldsymbol{e}_1\boldsymbol{e}_2$, and we recover the trivial case $e^a e^b = e^{a+b}$. Specifically, in the Euclidean $\mathscr{G}(2)$ plane (or anti-Euclidean $\mathscr{G}(0,2)$ plane) we have $\mathbb{I}^2 = -1$, and eq. (5.5) simplifies by way of the tangent angle addition identity

$$\arctan\left(\frac{\tan\theta_1 + \tan\theta_1}{1 - \tan\theta_1 \tan\theta_2}\right) = \theta_1 + \theta_2.$$

This identity encodes how angles add when given as the gradients of lines; $m = \tan\theta$.

Similarly, in the hyperbolic plane $\mathscr{G}(1,1)$ with basis $\{\boldsymbol{e}_+, \boldsymbol{e}_-\}$, $\boldsymbol{e}_\pm^2 = \pm 1$, the pseudoscalar $\mathbb{I} = \boldsymbol{e}_+\boldsymbol{e}_-$ generates *hyperbolic* rotations $e^{\mathbb{I}\xi} = \cosh\xi + \mathbb{I}\sinh\xi$ owing to the fact that $\mathbb{I}^2 = -\boldsymbol{e}_+^2\boldsymbol{e}_-^2 = +1$. Then, eq. (5.5) simplifies by the hyperbolic angle addition identity

$$\operatorname{arctanh}\left(\frac{\tanh\xi_1 + \tanh\xi_1}{1 + \tanh\xi_1 \tanh\xi_2}\right) = \xi_1 + \xi_2$$

which encodes how collinear rapidities add when given as relativistic velocities; $\beta = \tanh\xi$.

Less trivially, a rotation in $\mathbb{R}^3$ by θ may be represented by its Rodrigues vector[44] $\boldsymbol{r} = \hat{\boldsymbol{r}}\tan\frac{\theta}{2}$ pointing along the axis of rotation. The composition of two rotations is then succinctly encoded in Rodrigues' composition formula

[44] Olinde Rodrigues first originated the formula in 1840 [36, pp. 406].

$$\boldsymbol{r}_{12} = \frac{\boldsymbol{r}_1 + \boldsymbol{r}_2 - \boldsymbol{r}_1 \times \boldsymbol{r}_2}{1 - \boldsymbol{r}_1 \cdot \boldsymbol{r}_2} \tag{5.10}$$

involving the standard vector dot and cross products.

We can easily derive eq. (5.10) as a special case of eq. (5.7) as follows: Let $\sigma_1, \sigma_2 \in \mathscr{G}_2(3)$ be two bivectors defining the rotors e^{σ_1} and e^{σ_2} in three dimensions. In $\mathscr{G}(3)$, the only 4-vector is trivial, so $\sigma_1 \wedge \sigma_2 = 0$ and for the composite rotor $e^{\sigma_3} \coloneqq e^{\sigma_1}e^{\sigma_2}$ we have

$$\sigma_3 = \sigma_1 \odot \sigma_2 = \operatorname{arctanh}\left(\frac{\tanh\sigma_1 + \tanh\sigma_2 + \tanh\sigma_1 \times \tanh\sigma_2}{1 + \tanh\sigma_1 \cdot \tanh\sigma_2}\right)$$

where $a \times b$ is the commutator product of bivectors as in eq. (5.6), not the vector cross product. Observe that Euclidean bivectors $\sigma_i \in \mathscr{G}_2(3)$ have negative square (e.g., $(\boldsymbol{e}_1\boldsymbol{e}_2)^2 = -\boldsymbol{e}_1^2\boldsymbol{e}_2^2 = -1$) and relate to their dual normal vectors by $\boldsymbol{u}_i$ by $\sigma_i = \boldsymbol{u}_i\mathbb{I}$. Therefore, by rewriting $\tanh\sigma_i =$

$\tanh(\boldsymbol{u}_i\mathbb{I}) = (\tan \boldsymbol{u}_i)\mathbb{I}$, we obtain the formula in terms of plain vectors and the vector cross product.

$$\boldsymbol{u}_{12} = (\boldsymbol{u}_1\mathbb{I} \odot \boldsymbol{u}_2\mathbb{I})\mathbb{I}^{-1} = \arctan\left(\frac{\tan \boldsymbol{u}_1 + \tan \boldsymbol{u}_2 - \tan \boldsymbol{u}_1 \times \tan \boldsymbol{u}_2}{1 - \tan \boldsymbol{u}_1 \cdot \tan \boldsymbol{u}_2}\right)$$

Indeed, a bivector $\sigma_i = \boldsymbol{u}_i\mathbb{I}$ generates an $\mathbb{R}^3$ rotation through an angle $\theta = 2\|\boldsymbol{u}_i\|$ via the double-sided transformation law $a \mapsto e^{\boldsymbol{u}\mathbb{I}} a e^{-\boldsymbol{u}\mathbb{I}}$. Hence, $\tan \boldsymbol{u}_i = \hat{\boldsymbol{v}}_i \tan \frac{\theta}{2} \equiv \boldsymbol{r}_i$ are exactly the half-angle Rodrigues vectors, and we recover eq. (5.10).

The necessity of the half-angle in the Rodrigues vectors reflects the fact that they actually generate *rotors*, not direct rotations, and hence belong to the underlying spin representation of $\mathrm{SO}^+(3)$ — a fact made clearer in the context of geometric algebra.

5.1.3. In higher dimensions

In fewer than four dimensions, the 4-vector $\mathrm{T}_1 \wedge \mathrm{T}_2 = 0$ appearing in the geometric BCHD formula is trivial, and so eq. (5.5) involves only bivector addition and scalar multiplication. In four dimensions, there is one linearly independent 4-vector — the pseudoscalar — which necessarily commutes with all even multivectors. However, in more than four dimensions, 4-vectors do *not* necessarily commute with bivectors, and the assumptions underlying eq. (5.4) and hence the main result (5.5) fail.

On the face of it, the BCHD formula (5.5) in the four-dimensional case appears deceptively simple — it hides complexity in the calculation of the trigonometric functions of arbitrary bivectors,

$$\tanh \sigma_i = \sigma - \frac{1}{3}\sigma^3 + \frac{2}{15}\sigma^5 + \cdots \qquad \text{and} \qquad \operatorname{arctanh} \sigma_i = \sigma + \frac{1}{3}\sigma^3 + \frac{1}{5}\sigma^5 + \cdots. \tag{5.11}$$

In fewer dimensions, σ^2 is a scalar, and so these power series are as easy to compute as their real equivalents.[45] But in four dimensions, σ^2 is in general a $\{0, 4\}$-multivector (by lemma 6) and the power series (5.11) are more complicated. However, if $\sigma^2 \neq 0$ has a square root $N_\sigma = \alpha + \beta\mathbb{I}$ in the scalar–pseudoscalar plane, then one has $\sigma = N_\sigma\hat{\sigma} = \hat{\sigma}N_\sigma$ where

[45] If $\sigma^2 = N_\sigma^2 \in \mathbb{R}$, then we have simply $\tanh \sigma = (\tanh N_\sigma)N_\sigma^{-1}\sigma$.

$\hat{\sigma} \coloneqq \sigma/N_\sigma$ so that $\hat{\sigma}^2 = 1$. With a bivector $\sigma = N_\sigma\hat{\sigma}$ expressed in this form, the valuation of a formal power series $f(z) = \sum_{n=1}^{\infty} f_n z^n$ simplifies to

$$
\begin{aligned}
(f \text{ even}) \quad f(\sigma) &= \sum_{n=1}^{\infty} f_{2n}\sigma^{2n} = \sum_{n=1}^{\infty} f_{2n}N_\sigma^{2n} = f(N_\sigma), \\
(f \text{ odd}) \quad f(\sigma) &= \sum_{n=1}^{\infty} f_{2n+1}\sigma^{2n+1} = \sum_{n=1}^{\infty} f_{2n}N_\sigma^{2n+1}\hat{\sigma} = f(N_\sigma)\hat{\sigma}.
\end{aligned}
$$

This is especially useful in the case of Minkowski spacetime $\mathscr{G}(1,3)$ because the scalar–pseudoscalar plane is isomorphic to $\mathbb{C}$ and square roots always exist (see section 4.2). From now on, we focus on the special case of Minkowski spacetime, and consider practical and theoretical applications.

5.2. BCHD Composition in Spacetime

Because the geometric BCHD formula is constructed from sums and products of bivectors, it involves only even spacetime multivectors. Therefore, in numerical applications, it is not necessary to represent the full STA, but only the even subalgebra $\mathscr{G}_+(1,3) \cong \mathscr{G}(3)$.

The algebra of physical space $\mathscr{G}(3)$ admits a faithful complex linear representation by the Pauli spin matrices [25, 28, 37]. The real dimension of both $\mathbb{C}^{2\times 2}$ and $\mathscr{G}(3)$ is eight, so there is no redundancy in the Pauli representation, so it is convenient for computer implementation.

An even $\mathscr{G}_+(1,3)$ multivector — or equivalently, a general $\mathscr{G}(3)$ multivector — may be parametrised by four complex scalars $q^\mu = \mathfrak{R}(q^\mu) + i\mathfrak{I}(q^\mu) \in \mathbb{C}$ as

$$
A = \mathfrak{R}(q^0) + \mathfrak{R}(q^i)\vec{\sigma}_i + \mathfrak{I}(q^i)\mathbb{I}\vec{\sigma}_i + \mathfrak{I}(q^0)\mathbb{I},
$$

where the $\vec{\sigma}_i$ may be read both as spacetime bivectors $\vec{\sigma}_i \equiv \boldsymbol{\gamma}_0\boldsymbol{\gamma}_i \in \mathscr{G}_+(1,3)$ or as basis vectors of $\mathscr{G}(3)$ under a space/time split. The Pauli matrices $\sigma_i \in \mathbb{C}^{2\times 2}$ form a linear representation of $\mathscr{G}(3)$ by the association $\vec{\sigma}_i \equiv \sigma_i$.

Explicitly, identifying

$$\vec{\sigma}_1 \equiv \begin{bmatrix} 0 & +1 \\ +1 & 0 \end{bmatrix} \qquad \vec{\sigma}_2 \equiv \begin{bmatrix} 0 & -i \\ +i & 0 \end{bmatrix} \qquad \vec{\sigma}_3 \equiv \begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix}$$

along with $1 \equiv I$ and $\mathbb{I} \equiv iI$ where I is the identity matrix, we obtain a representation of the multivector A by a 2×2 complex matrix:

$$\mathrm{A} \equiv \begin{bmatrix} q^0 + q^3 & q^1 - iq^2 \\ q^1 + iq^2 & q^0 - q^3 \end{bmatrix}. \tag{5.12}$$

A proper Lorentz transformation $\Lambda \in \mathrm{SO}^+(1,3)$ is determined in the K frame by a vector rapidity $\boldsymbol{\xi} \in \mathbb{R}^3$ and axis-angle vector $\boldsymbol{\theta} \in \mathbb{R}^3$. The standard 4×4 matrix representation of Λ is obtained as the exponential of the generator

$$\begin{bmatrix} 0 & \boldsymbol{\xi}^T \\ \boldsymbol{\xi} & \varepsilon_{ijk}\theta^k \end{bmatrix} = \begin{bmatrix} 0 & \xi^1 & \xi^2 & \xi^3 \\ \xi^1 & 0 & +\theta^3 & -\theta^2 \\ \xi^2 & -\theta^3 & 0 & +\theta^1 \\ \xi^3 & +\theta^2 & -\theta^1 & 0 \end{bmatrix} \in \mathfrak{so}(1,3). \tag{5.13}$$

In the spin representation, the transformation Λ corresponds to a rotor $\mathscr{L} = e^{\sigma}$, and the generating bivector (4.2) may be expressed via eq. (5.12) as the traceless complex matrix

$$\Sigma = q^k \sigma_k = \begin{bmatrix} +q^3 & q^1 - iq^2 \\ q^1 + iq^2 & -q^3 \end{bmatrix}, \tag{5.14}$$

where $q^k \coloneqq \frac{1}{2}(\xi^k + i\theta^k) \in \mathbb{C}$. Note that, since the square of a spacetime bivector is a $\{0,4\}$-multivector, its representative matrix Σ squares to a complex scalar multiple of the identity.

Given two generators σ_i with matrix representations Σ_i, the geometric BCHD formula (5.5) reads

$$\Sigma_3 \coloneqq \Sigma_1 \odot \Sigma_2 = \tanh^{-1}\left(\frac{\mathrm{T}_1 + \mathrm{T}_2 + \mathrm{A}}{\mathrm{I} + \mathrm{S}}\right), \tag{5.15}$$

where $\mathrm{T}_i \coloneqq \tanh \Sigma_i$. To efficiently compute T_i, make use of the fact that $\Sigma_i^2 = \lambda_i^2 \mathrm{I}$ is a complex multiple of the identity matrix and evaluate

$\mathrm{T}_i = (\tanh\lambda_i)\lambda_i^{-1}\Sigma_i$. In the null case $\Sigma_i^2 = \lambda = 0$, we have trivially $\tanh\Sigma_i = \Sigma_i = \tanh^{-1}\Sigma_i$.

The commutator $\mathrm{A} \coloneqq \frac{1}{2}[\mathrm{T}_1, \mathrm{T}_2]$ and anti-commutator $\mathrm{S} \coloneqq \frac{1}{2}\{\mathrm{T}_1, \mathrm{T}_2\}$ terms may be efficiently computed by separating the single matrix product $\Pi \coloneqq \mathrm{T}_1\mathrm{T}_2 = \mathrm{A} + \mathrm{S}$ into off-diagonal and diagonal components, respectively; i.e.,

$$\mathrm{A}_{ij} = (1 - \delta_{ij})\Pi_{ij} \quad \text{and} \quad \mathrm{S}_{ij} = \delta_{ij}\Pi_{ij}.$$

The numerator of eq. (5.15) is therefore a matrix with zeros on the diagonal, and the denominator is a complex scalar multiple of the identity, so the argument of $\tanh^{-1}$, call it M, is in the form (5.14). Computing $\tanh^{-1}\mathrm{M}$ again simply amounts to $\Sigma_3 = \tanh^{-1}\mathrm{M} = (\tanh^{-1}\lambda)\lambda^{-1}\mathrm{M}$ where $\mathrm{M}^2 = \lambda^2\mathrm{I}$.

The Lorentz generator in the standard vector representation (5.13) can then be recovered from Σ_3 with the relations $\xi^k = 2\mathfrak{R}(q^k)$ and $\theta^k = 2\mathfrak{I}(q^k)$, and the final $\mathrm{SO}^+(1,3)$ vector transformation is its 4×4 matrix exponential.

5.2.1. Relativistic 3-velocities and the Wigner angle

As an example of its theoretical utility, we shall use the geometric BCHD formula (5.5) to derive the composition law for arbitrary relativistic 3-velocities.

The innocuous problem of composing relativistic velocities has been called “paradoxical” [38–40], owing in part to the fact that *irrotational* boosts are not closed under composition, and that explicit matrix analysis becomes cumbersome. Of course, in reality there is no paradox, and the full description of the composition of boosts is pedagogically valuable as it highlights aspects of special relativity which differ from spatial intuition.

We may speak of a rotation or boost as being PURE relative to the K frame. Technically, σ generates a pure rotation (or pure boost) if, under the space/time split relative to the K frame, $\sigma = \langle\sigma\rangle_2$ is a pure bivector (or a pure vector) in $\mathscr{G}(3)$. A pure rotation or pure boost relative to K is *not* pure in all other frames.

The restriction of the BCHD formula to pure boosts is not as simple as the restriction to rotations (5.10), because pure boosts do not form a closed subgroup of $\mathrm{SO}^+(1,3)$ as pure rotations do. Instead, the composition of two pure boosts $\mathscr{B}_i$ is a pure boost composed with a pure rotation (or vice versa),

$$\mathscr{B}_1\mathscr{B}_2 = \mathscr{B}\mathscr{R}. \tag{5.16}$$

The direction of the boost $\mathscr{B}$ lies within the plane defined by the boost directions of $\mathscr{B}_1$ and $\mathscr{B}_2$, and $\mathscr{R}$ is a rotation through this plane by the *Wigner angle* [40]. Applying eq. (5.5) to this case immediately yields formulae for the resulting boost and rotation.[46]

[46] These results are equivalent to those in [9] which are formulated using complexified quaternions.

For ease of algebra, we conduct the following analysis under a space/time split with respect to the K frame. Under this split, a pure boost $\mathscr{B}$ is generated by an $\mathbb{R}^3$ vector $\frac{\boldsymbol{\xi}}{2}$, and a pure rotation $\mathscr{R}$ is generated by an $\mathbb{R}^3$ bivector $\frac{\theta}{2}\hat{r}$. Here, $\boldsymbol{\xi} \in \mathscr{G}_1(3)$ is the *vector rapidity*, related to the velocity by $\boldsymbol{v}/c = \boldsymbol{\beta} = \tanh \boldsymbol{\xi}$, and the rotation is through an angle θ in the plane spanned by the bivector $\hat{r} \in \mathscr{G}_2(3)$. Equation (5.5) with two pure boosts $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ is

$$\tanh\left(\frac{\boldsymbol{\xi}_1}{2} \odot \frac{\boldsymbol{\xi}_2}{2}\right) = \frac{\boldsymbol{w}_1 + \boldsymbol{w}_2 + \boldsymbol{w}_1 \wedge \boldsymbol{w}_2}{1 + \boldsymbol{w}_1 \cdot \boldsymbol{w}_2} \tag{5.17}$$

where $\boldsymbol{w}_i \coloneqq \tanh \frac{\boldsymbol{\xi}_i}{2}$ are the *relativistic half-velocities*, also defined in [9, 10]. The generator (5.17) has vector and bivector (namely $\boldsymbol{w}_1 \wedge \boldsymbol{w}_2$) parts, indicating that the Lorentz transformation it describes is indeed some combination of a boost and a rotation.

Similarly, for an arbitrary pure boost and pure rotation,

$$\tanh\left(\frac{\boldsymbol{\xi}}{2} \odot \frac{\theta}{2}\hat{r}\right) = \frac{\boldsymbol{w} + \rho + \frac{1}{2}[\boldsymbol{w}, \rho]}{1 + \boldsymbol{w} \wedge \rho} \tag{5.18}$$

where $\rho \coloneqq \tanh \frac{\theta\hat{r}}{2} = \hat{r} \tan \frac{\theta}{2}$ is a bivector. In general, eq. (5.18) has vector, bivector *and* pseudoscalar parts (the commutator $\frac{1}{2}[\boldsymbol{w}, \rho] = \langle \boldsymbol{w}\rho \rangle_1 + \boldsymbol{w} \wedge \rho$ and the denominator both have grade-three part $\boldsymbol{w} \wedge \rho$). However, eqs. (5.17) and (5.18) are equal by supposition of eq. (5.16). By comparing parts of equal grade, we deduce the pseudoscalar part of eq. (5.18) is zero. This requires $\boldsymbol{w} \wedge \rho = 0$ or, equivalently, that $\boldsymbol{w}$ lies in the plane defined by ρ — meaning the resulting boost is coplanar with the Wigner rotation as expected. Hence, for a coplanar boost and rotation, eq. (5.18) is simply

$$\tanh\left(\frac{\xi}{2} \odot \frac{\theta}{2}\hat{r}\right) = \boldsymbol{w} + \rho + \boldsymbol{w}\rho. \tag{5.19}$$

The term $\boldsymbol{w}\rho = \langle \boldsymbol{w}\rho \rangle_1 = -\rho\boldsymbol{w}$ is a vector orthogonal to $\boldsymbol{w}$ in the plane defined by ρ.

Equating the bivector parts of eqs. (5.17) and (5.19) determines the rotation

$$\rho = \frac{\boldsymbol{w}_1 \wedge \boldsymbol{w}_2}{1 + \boldsymbol{w}_1 \cdot \boldsymbol{w}_2}, \quad \text{implying} \quad \theta = 2\tan^{-1}\left(\frac{w_1 w_2 \sin\phi}{1 + w_1 w_2 \cos\phi}\right)$$

where ϕ is the angle between the two initial boosts (in the K frame). The angle θ is precisely the Wigner angle. Equating the vector parts determines the boost

$$\boldsymbol{w} = \frac{\boldsymbol{w}_1 + \boldsymbol{w}_2}{1 + \boldsymbol{w}_1 \cdot \boldsymbol{w}_2}(1 + \rho)^{-1},$$

noting that $\boldsymbol{w}_i$ and ρ do not commute. Substituting ρ leads to the remarkably succinct composition law $\boldsymbol{w} = (\boldsymbol{w}_1 + \boldsymbol{w}_2)(1 + \boldsymbol{w}_1\boldsymbol{w}_2)^{-1}$ exhibited in [9], with the final relativistic velocity being $\boldsymbol{\beta} = \tanh\boldsymbol{\xi} = \tanh\left(2\tanh^{-1}\boldsymbol{w}\right)$.

Chapter 6.

Calculus in Flat Geometries

So far, we have been concerned with special relativity at a single point in spacetime. We move now toward the description of *fields* — quantities extending across regions of spacetime. The first step in this direction is the calculus of *flat spacetime*. In a flat geometry, we may assume that

- points in spacetime are elements of a vector space, with differences of points being physically meaningful; and that
- fields are parametric functions of a single vector argument representing a location in spacetime.

We reserve the word FIELD to mean a map with a fixed vector space codomain. For instance, the electromagnetic bivector *field* in flat space $F : \mathbb{R}^4 \to \wedge^2\mathbb{R}^4$ is a function between fixed vector spaces. In particular, the values of such a vector field at different points in spacetime belong to the same space, and for example, their sum $F(x) + F(y) \in A$ is well-defined.

These assumptions are acceptable in special relativity, but in arbitrary regions of spacetime and in the presence of gravity, curvature prevents spacetime from admitting a meaningful vector space structure. It is then *un-physical* to compare field values at different points in spacetime. (Consideration of curvature leads to differential geometry and comprises part II.)

This chapter introduces the *exterior* and *vector derivatives* as instances of the 'algebraic derivative' in the exterior and geometric algebras, respectively. These devices combine derivative information with the geometrical structure inherent in the respective algebras. To demonstrate their utility, a case study of Maxwell's equations of electromagnetism are given in terms of these algebraic derivatives.

6.1. Differentiation

The directional derivative of a vector field $F : V \to A$ in the direction $\boldsymbol{u} \in V$ at $\boldsymbol{x} \in V$ may be defined in the usual way,

$$\partial_{\boldsymbol{u}} F(\boldsymbol{x}) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} F(\boldsymbol{x} + \varepsilon \boldsymbol{u})\bigg|_{\varepsilon=0} = \lim_{\varepsilon \to 0} \frac{F(\boldsymbol{x} + \varepsilon\boldsymbol{u}) - F(\boldsymbol{x})}{\varepsilon}.$$

The directional derivative is linear in both its argument and direction.[47] We define the notation $\partial_a := \partial_{\boldsymbol{e}_a}$ for brevity, so long as it is understood that this is not a partial derivative with respect to a scalar coordinate, $\frac{\partial}{\partial x^a}$. Of course, it may be viewed as such by setting $f(x^1, \dots, x^n) = f(x^i \boldsymbol{e}_i)$ so that

$$\partial_{\boldsymbol{e}_a} f(x^i \boldsymbol{e}_i) = \frac{\partial}{\partial x^a} f(x^1, \dots, x^n),$$

though this is a basis-dependent definition.

[47] By a change of variables, $\partial_{u^a \boldsymbol{e}_a} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} F(x + \varepsilon u^a \boldsymbol{e}_a)\big|_{\varepsilon=0} = u^a \frac{\mathrm{d}}{\mathrm{d}\bar\varepsilon} F(x + \bar\varepsilon \boldsymbol{e}_a)\big|_{\bar\varepsilon=0} = u^a \partial_{\boldsymbol{e}_a}$ (summation on a).

Suppose $F : V \to A$ is some algebra–valued field. It is useful to define a kind of "total" derivative $\mathrm{D}\,F$ which does not depend on a direction $\boldsymbol{u}$, but instead encompasses, in a sense, all directional derivatives in a single object $\mathrm{D}\,F : V \to A$. The motivation for this is that the soon-to-be-defined exterior derivative (of exterior algebra) and vector derivative (of geometric algebra) are realised as special cases of such a construction. The derivative D will be defined whenever there is a canonical inclusion $\iota : V^* \to A$ of dual vectors into the algebra, which is automatic if A is a quotient of $(V^*)^{\otimes}$.

Definition 24. *Let $F : V \to A$ be a field with values in an algebra A with product $\circledast$, equipped with an inclusion $\iota : V^* \to A$. The* ALGEBRAIC

DERIVATIVE of F is

$$\mathrm{D}\,F \coloneqq \iota(\boldsymbol{e}^a) \circledast \partial_{\boldsymbol{e}_a} F \tag{6.1}$$

(summation on a) where $\{\boldsymbol{e}_a\} \subset V$ and $\{\boldsymbol{e}^a\} \subset V^*$ are dual bases.

To understand this definition, consider the simple case of the free tensor algebra $F : V \to (V^*)^{\otimes}$. We leave the canonical inclusion $\iota : V^* \to (V^*)^{\otimes}$ implicit. Given a basis $\{\boldsymbol{e}^a\} \subset V^*$, the algebraic derivative is $\mathrm{D}\,F = \boldsymbol{e}^a \otimes \partial_a F$, which simply encodes the partial derivatives of a k-vector F in a $(k+1)$-grade object. In component language, $(\mathrm{D}\,F)_{aa_1\cdots a_k} = \partial_a F_{a_1\cdots a_k}$.

6.1.1. The exterior derivative

Consider a vector field $F : V \to \wedge V^*$ with values in the (dual) exterior algebra. In this case eq. (6.1) is the EXTERIOR DERIVATIVE

$$\mathrm{d}F = \boldsymbol{e}^a \wedge \partial_a F,$$

where $\{\boldsymbol{e}^a\} \subset V^*$ also form a basis of $\wedge V^*$ (so $\iota : V^* \to \wedge V^*$ may be omitted). If $F : V \to \wedge^k V^*$ is a k-vector field, then $\mathrm{d}F = \partial_a F_{a_1\cdots a_k} \boldsymbol{e}^a \wedge \boldsymbol{e}^{a_1} \wedge \cdots \wedge \boldsymbol{e}^{a_k}$ is a $(k+1)$-vector.

Using the equivalence of $\wedge V^*$ with the subspace of antisymmetric tensors (see section 2.2.1), the exterior derivative is seen to be the totally anti-symmetrised partial derivative. In components, $(\mathrm{d}F)_{a_1\cdots a_k} = \partial_{[a_1} F_{a_2\cdots a_k]}$.

The treatment of exterior forms is identical. An exterior form field $\varphi : V \to \Omega^k(V, U)$ is called a U-valued EXTERIOR DIFFERENTIAL k-FORM, with exterior derivative formally defined by its action on vectors,

$$\begin{aligned}(\mathrm{d}\varphi)(\boldsymbol{u}_0, \boldsymbol{u}_1, \dots, \boldsymbol{u}_k) &= (\boldsymbol{e}^a \wedge \partial_a \varphi)(\boldsymbol{u}_0, \boldsymbol{u}_1, \dots, \boldsymbol{u}_k) \\ &= \frac{1}{k!} \sum_{\sigma \in S_{k+1}} (-1)^\sigma \boldsymbol{e}^a(\boldsymbol{u}_{\sigma(0)})\, \partial_a \varphi(\boldsymbol{u}_{\sigma(1)} \cdots \boldsymbol{u}_{\sigma(k)}) \\ &= \sum_{i=0}^{k} (-1)^i\, \partial_{\boldsymbol{u}_i} \varphi(\boldsymbol{u}_0, \dots, \widehat{\boldsymbol{u}}_i, \dots, \boldsymbol{u}_k),\end{aligned}$$

under the Spivak convention (see section 2.2.2). Note that the directional derivative acts on the position dependence of φ only — the vectors $\boldsymbol{u}_i \in V$ are *fixed* input vectors to the field $\mathrm{d}\varphi$. This changes when generalising to forms defined on a *manifold*, where correction terms are needed to account for partial derivatives of input vectors (discussed in part II).

6.1.2. The vector derivative

The algebraic derivative in the tensor and exterior algebras are somewhat uninteresting because they are easily expressible in component form (e.g., $\partial_a F_{a_1\cdots a_k}$ or $\partial_{[a}F_{a_1\cdots a_k]}$). This is not possible in the geometric algebra, however, because $\mathscr{G}(V,\eta)$ is not $\mathbb{Z}$-graded, and we would face the problem of notating inhomogeneous objects with a variable number of indices. The algebraic derivative is, however, still geometrically significant and extremely useful in geometric algebra.

In $\mathscr{G}(V,\eta)$, the algebraic derivative is called the VECTOR DERIVATIVE, denoted $\boldsymbol{\partial}$. Explicitly, if $F : V \to \mathscr{G}(V,\eta)$ is a multivector field, then from eq. (6.1) we take $\circledast$ to be with the geometric product and take the inclusion to be[48] $V^* \ni \boldsymbol{u} \mapsto \iota(\boldsymbol{u}^\sharp) \in \mathscr{G}(V,\eta)$. Here, we use the canonical inclusion $\iota : V \equiv \mathscr{G}_1(V,\eta) \to \mathscr{G}(V,\eta)$ and the metric relating $V^* \to V$. The vector derivative is then

$$\boldsymbol{\partial}F = \boldsymbol{e}^a\,\partial_{\boldsymbol{e}_a}F$$

(summation on a) where $\{\boldsymbol{e}_a\} \subset V$ and $\{\boldsymbol{e}^a\} \subset V^*$ are dual bases, and juxtaposition denotes the geometric product. If F is a homogeneous k-vector, then we may write its components as $F = F_{a_1\cdots a_k}\boldsymbol{e}^{a_1} \wedge \cdots \wedge \boldsymbol{e}^{a_k}$ and hence

$$\boldsymbol{\partial}F = \partial_{\boldsymbol{e}^a}F_{a_1\cdots a_k}\,\boldsymbol{e}^a(\boldsymbol{e}^{a_1} \wedge \cdots \wedge \boldsymbol{e}^{a_k}).$$

Note that these terms are not $(k+1)$-blades, but geometric products of vectors $\boldsymbol{e}^a$ with k-blades — in general, $(k \pm 1)$-multivectors.

We may regard the vector derivative itself as an operator-valued vector,

$$\boldsymbol{\partial} = \boldsymbol{e}^a\partial_a,$$

[48] We could just as well consider fields $V \to \mathscr{G}(V^*,\eta)$, avoiding the need for the isomorphism $\sharp : V^* \to V$. But the metric is already defined, and we prefer to think about multivectors over ‘dual multivectors’.

reflecting the fact that $\boldsymbol{\partial}$ behaves algebraically like a vector. For instance, the derivative of a vector $\boldsymbol{u}$ has scalar and bivector parts, $\boldsymbol{\partial u} = \boldsymbol{\partial} \cdot \boldsymbol{u} + \boldsymbol{\partial} \wedge \boldsymbol{u}$, just like the geometric product of two vectors, $\boldsymbol{uv} = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \wedge \boldsymbol{v}$. For a general multivector F, then, we have

$$\boldsymbol{\partial} F = \boldsymbol{\partial} \rfloor F + \boldsymbol{\partial} \wedge F.$$

The $(k+1)$-grade part $\boldsymbol{\partial} \wedge F$ is the *curl* of F, and coincides with the exterior derivative $\mathrm{d}F$. The $(k-1)$-grade part involves the metric, and can be related to the 'interior' derivative $\star\mathrm{d}\star A$ via Hodge duality.[49]

[49] Observe that $\partial \rfloor A = \langle \partial \mathbb{I}^{-1} \mathbb{I} A \rangle_{k-1} = \pm \mathbb{I} \langle \partial (\mathbb{I} A) \rangle_{1+n-k} = \pm (\mathbb{I}\partial) \wedge (\mathbb{I} A)$.

6.2. Case Study: Maxwell's Equations

Expressed in the standard vector calculus of $\mathbb{R}^3$, Maxwell's equations for the electric $\boldsymbol{E}$ and magnetic $\boldsymbol{B}$ fields in the presence of a source are

$$\begin{aligned}
&\nabla \cdot \boldsymbol{E} = \frac{\rho}{\varepsilon_0} && \text{(Gauß' law)} \\
&\nabla \cdot \boldsymbol{B} = 0 && \text{(Absence of magnetic monopoles)} \\
&\nabla \times \boldsymbol{E} = -\partial_t \boldsymbol{B} && \text{(Faraday's law)} \\
&\nabla \times \boldsymbol{B} = \mu_0 (\boldsymbol{J} + \varepsilon_0 \partial_t \boldsymbol{E}) && \text{(Ampère's law)}
\end{aligned}$$

where ρ is the scalar charge density and $\boldsymbol{J}$ the current density. The constants ε_0 and μ_0 are the vacuum permittivity and permeability, respectively, related to the speed of light c by $\varepsilon_0 \mu_0 c^2 = 1$.

Non-relativistic

quantity	dimension
$\boldsymbol{E}$	$MQ^{-1}LT^{-2}$
$\boldsymbol{B}$	$MQ^{-1}T^{-1}$
ρ	QL^{-3}
$\boldsymbol{J}$	$QT^{-1}L^{-2}$
μ_0	$MQ^{-2}L$
ε_0	$M^{-1}Q^2L^{-3}T^2$
∇, ∂_t	L^{-1}, T^{-1}
c	LT^{-1}

Relativistic

quantity	dimension
F	$MQ^{-1}S^{-1}$
J	QS^{-3}
μ_0, ε_0^{-1}	$MQ^{-2}S$
∂	S^{-1}
c	1

Table 6.1.: Dimensions of physical quantities in Maxwell's equations. M is mass, Q is electric charge, T is duration and L is length. In the relativistic formulation, T and L are unified and replaced by *spacetime interval* S.

6.2.1. With tensor calculus

These can be expressed relativistically as eight scalar equations,

$$\boxed{\partial_\mu F^{\mu\nu} = \mu_0 J^\nu, \qquad \partial_\mu G^{\mu\nu} = 0} \tag{6.2}$$

where $F^{\mu\nu} = -F^{\nu\mu}$ is the Faraday tensor and $G^{\mu\nu}$ its Hodge dual, both encoding the electric and magnetic fields via

$$F^{i0} = \frac{E^i}{c}, \qquad F^{ij} = -\varepsilon^{ijk} B_k, \qquad G^{\mu\nu} = \frac{1}{2} \varepsilon^{\mu\nu}{}_{\rho\sigma} F^{\rho\sigma}, \tag{6.3}$$

and where J^μ encodes both the static charge density $J^0 = c\rho$ and current density $J^i = \boldsymbol{J}$. The left of eqs. (6.2) is the *source equation*, while the right is the *second Bianchi identity*. These equations assume the metric signature $(+---)$, where the equivalent equations under $(-+++)$ are obtained by a change of sign $F^{\mu\nu} \mapsto -F^{\mu\nu}$.

Proof. We show how the relativistic equations (6.2) reduce to the non-relativistic vector calculus equivalents. The 0-component of the source equation is $\partial_\mu F^{\mu 0} = \partial_i E^i/c = \mu_0 J^0 = \mu_0 c\rho$ implying $\nabla \cdot \boldsymbol{E} = \rho/\varepsilon_0$ (Gauß' law). The i-components are

$$\begin{aligned} \partial_0 F^{0i} + \partial_j F^{ji} &= \frac{1}{c}\partial_t\left(-\frac{E^i}{c}\right) - \partial_j \varepsilon^{jik} B_k = \mu_0 J^i \\ \text{or} \quad \partial_j \varepsilon^{ijk} B_k &= \mu_0 J^i + \mu_0 \varepsilon_0 \partial_t E^i, \end{aligned}$$

which is equivalent to Ampère's law. The 0-component of the Bianchi identity $\partial_\mu G^{\mu 0} = 0$ is

$$\frac{1}{2}\varepsilon^i{}_{jk}\partial_i F^{jk} = -\frac{1}{2}\varepsilon^i{}_{jk}\varepsilon^{jkl}\partial_i B_l = -\partial_i B^i = 0,$$

which using the identity $\varepsilon_{ijk}\varepsilon^{jkl} = 2\delta_i^l$ is $\nabla \cdot \boldsymbol{B} = 0$. Finally, the i-component gives

$$\begin{aligned} 0 = \partial_\mu G^{\mu i} = \frac{1}{2}\varepsilon^{\mu i}{}_{\rho\sigma}\partial_\mu F^{\rho\sigma} &= \frac{1}{2}\varepsilon^{0i}{}_{jk}\partial_0 F^{jk} + \varepsilon^{ji}{}_{k0}\partial_j F^{k0} \\ &= -\frac{1}{4}\varepsilon^i{}_{jk}\varepsilon^{jkl}\partial_0 B_l - \frac{1}{2c}\varepsilon^{ijk}\partial_j E_k = -\frac{1}{2c}\left(\partial_t B^i + \varepsilon^{ijk}\partial_j E_k\right) \end{aligned}$$

yielding Faraday's law $\nabla \times \boldsymbol{E} = -\partial_t \boldsymbol{B}$. □

6.2.2. With exterior calculus

It is easy to translate from the language of exterior calculus to tensor calculus, and hence vice versa, by identifying the former as the subalgebra of totally antisymmetric tensors (as in section 2.2.1). We will employ the Spivak convention, which in particular identifies 2-forms via

$$\boldsymbol{e}^\mu \wedge \boldsymbol{e}^\nu \equiv \boldsymbol{e}^\mu \otimes \boldsymbol{e}^\nu - \boldsymbol{e}^\nu \otimes \boldsymbol{e}^\mu$$

where $\boldsymbol{e}^\mu$ are spacetime basis vectors (having physical dimensions of spacetime interval, S). We then the electromagnetic bivector as $\mathscr{F} = \frac{1}{2}F_{\mu\nu}\boldsymbol{e}^\mu \wedge \boldsymbol{e}^\nu$ (omitting the $\frac{1}{2}$ in the Kobayashi–Nomizu convention).

Since the charge density $J \sim QS^{-3}$ has dimensions of charge per spacetime 3-volume, it is natural to interpret it as a *trivector*

$$\mathscr{J} = J^{\mu\nu\lambda}\boldsymbol{e}_\mu \wedge \boldsymbol{e}_\nu \wedge \boldsymbol{e}_\lambda := J^\mu \star \boldsymbol{e}_\mu = \frac{1}{3!}\varepsilon_{\mu\nu\lambda\alpha}J^\alpha \boldsymbol{e}^\mu \wedge \boldsymbol{e}^\nu \wedge \boldsymbol{e}^\lambda$$

so that the coefficients $J^{\mu\nu\lambda} \sim Q$ have dimensions of charge.[50]

[50] Note that dual vectors $\boldsymbol{e}_\mu$ have dimension S^{-1}.

The relativistic Maxwell equations are then

$$\boxed{\mathrm{d}\star\mathscr{F} = \mu_0\mathscr{J}, \qquad \mathrm{d}\mathscr{F} = 0.}$$

Proof. The first equation written in component form is

$$\frac{1}{4}\varepsilon_{\mu\nu\rho\sigma}\partial_\lambda F^{\rho\sigma} = \frac{1}{3!}\varepsilon_{\lambda\mu\nu\alpha}\mu_0 J^\alpha,$$

which, by contracting with $\varepsilon^{\mu\nu\lambda\beta}$ and using the identities $\varepsilon^{\mu\nu\lambda\beta}\varepsilon_{\mu\nu\rho\sigma} = 2(\delta^\lambda_\rho\delta^\beta_\sigma - \delta^\lambda_\sigma\delta^\beta_\rho)$ and $\varepsilon^{\mu\nu\lambda\beta}\varepsilon_{\lambda\mu\nu\alpha} = 3!\delta^\beta_\sigma$, reduces to

$$\frac{1}{2}(\partial_\lambda F^{\lambda\beta} - \partial_\lambda F^{\beta\lambda}) = \mu_0 J^\beta$$

or $\partial_\mu F^{\mu\nu} = \mu_0 J^\nu$, the source equation. The Bianchi identity can be rewritten as

$$\partial_\mu G^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu}{}_{\rho\sigma}\partial_\mu F^{\rho\sigma} = -\frac{1}{2}\varepsilon^{\nu[\mu\rho\sigma]}\partial_\mu F_{\rho\sigma} = -\frac{1}{2}\varepsilon^{\nu\mu\rho\sigma}\partial_{[\mu}F_{\rho\sigma]} = 0,$$

implying $\mathrm{d}\mathscr{F} = 0$. □

6.2.3. With geometric calculus

Using the spacetime algebra $\mathscr{G}(1,3)$ with vector basis $\{\boldsymbol{\gamma}_\mu\}$ as introduced in chapter 4, the electromagnetic bivector is[51]

$$F = F^{\mu\nu}\boldsymbol{\gamma}_\mu\boldsymbol{\gamma}_\nu \tag{6.4}$$

[51] This coincides with the electromagnetic bivector 2-form $\mathscr{F}$ in the Kobayashi–Nomizu convention, because the wedge product in geometric algebra is naturally normalised (see table 2.1).

and the current density is

$$\boldsymbol{J} = J^\mu \boldsymbol{\gamma}_\mu.$$

Maxwell's equations are equivalent to the *single* multivector equation

$$\boxed{\boldsymbol{\partial} F = \mu_0 \boldsymbol{J}.} \tag{6.5}$$

Proof. The multivector equation $\boldsymbol{\partial} F = \mu_0 \boldsymbol{J}$ separates into a vector part $\boldsymbol{\partial} \cdot F = \mu_0 \boldsymbol{J}$ and a trivector part $\boldsymbol{\partial} \wedge F = 0$. In terms of components, the vector part is

$$\boldsymbol{\partial} \cdot F = \partial_\lambda F^{\mu\nu} \boldsymbol{\gamma}^\lambda \cdot (\boldsymbol{\gamma}_\mu \boldsymbol{\gamma}_\nu) = \mu_0 J^\nu \boldsymbol{\gamma}_\nu.$$

The only non-zero components are those for which $\mu \neq \nu$. If λ, μ and ν are all distinct, then $\boldsymbol{\gamma}^\lambda \cdot (\boldsymbol{\gamma}_\mu \boldsymbol{\gamma}_\nu) = \left\langle \boldsymbol{\gamma}^\lambda \boldsymbol{\gamma}_\mu \boldsymbol{\gamma}_\nu \right\rangle_1 = 0$. There are then two cases, $\lambda = \mu$ and $\lambda = \nu$, which respectively simplify

$$\begin{aligned} \boldsymbol{\gamma}^\mu \cdot (\boldsymbol{\gamma}_\mu \boldsymbol{\gamma}_\nu) &= \left\langle \boldsymbol{\gamma}^\mu \boldsymbol{\gamma}_\mu \boldsymbol{\gamma}_\nu \right\rangle_1 = \boldsymbol{\gamma}_\nu, \\ \boldsymbol{\gamma}^\nu \cdot (\boldsymbol{\gamma}_\mu \boldsymbol{\gamma}_\nu) &= \left\langle \boldsymbol{\gamma}^\nu \boldsymbol{\gamma}_\mu \boldsymbol{\gamma}_\nu \right\rangle_1 = -\boldsymbol{\gamma}_\mu, \end{aligned}$$

so that

$$\boldsymbol{\partial} \cdot F = \left(\partial_\mu F^{\mu\nu} \boldsymbol{\gamma}_\nu - \partial_\nu F^{\mu\nu} \boldsymbol{\gamma}_\mu \right) = \partial_\mu F^{\mu\nu} \boldsymbol{\gamma}_\nu.$$

This recovers the source equation $2\partial_\mu F^{\mu\nu} = \mu_0 J^\nu$.

It is clear that the trivector part

$$\boldsymbol{\partial} \wedge F = \partial_\lambda F^{\mu\nu} \boldsymbol{\gamma}^\lambda \wedge (\boldsymbol{\gamma}_\mu \boldsymbol{\gamma}_\nu) = \partial_\lambda F_{\mu\nu} \boldsymbol{\gamma}^\lambda \wedge \boldsymbol{\gamma}^\mu \wedge \boldsymbol{\gamma}^\nu = 0$$

is equivalent to the exterior algebraic Bianchi identity $\mathrm{d}\mathscr{F} = 0$. □

In terms of electric and magnetic fields

It is worth showing how the relativistic Maxwell equation (6.5) splits into a frame-dependent description in the geometric algebra framework. As in section 4.1, we use the notation $\vec{u}$ to indicate relative vectors; i.e., time-like bivectors of the spacetime algebra $\mathscr{G}(1,3)$ which are simultaneously grade-1 vectors in the observer's algebra $\mathscr{G}(3)$.

From eqs. (6.3) and (6.4), the electromagnetic bivector is expressed in the $\boldsymbol{\gamma}_0$-frame as[52]

$$F = \frac{1}{c}\vec{E} + \mathbb{I}\vec{B}, \tag{6.6}$$

[52] We assume $(+{-}{-}{-})$ for concreteness; for $(-{+}{+}{+})$ replace $F \mapsto -F$.

where $\vec{E} = E^i\vec{\sigma}_i = E^i\boldsymbol{\gamma}_i\boldsymbol{\gamma}_0$ and

$$\mathbb{I}\vec{B} = B_i\mathbb{I}\vec{\sigma}^i = \frac{1}{2}B_i\varepsilon^{ijk}\vec{\sigma}_j\vec{\sigma}_k = \frac{1}{2}B_i\varepsilon^{ijk}\boldsymbol{\gamma}_j\boldsymbol{\gamma}_k.$$

Equation (6.6) should be compared with the Riemann-Silberstein vector [31] which has the form $\vec{F}_{\mathbb{C}} = \vec{E} + ic\vec{B}$.

Similarly, the current density spacetime vector J may be viewed under the space/time split by (left) multiplying by the frame velocity $\boldsymbol{\gamma}_0$,

$$\boldsymbol{\gamma}_0\boldsymbol{J} = c\rho - \vec{J},$$

where $J^0 = c\rho$ and $\vec{J} = J^i\vec{\sigma}_i$. Similarly for the vector derivative, we have

$$\boldsymbol{\gamma}_0\,\boldsymbol{\partial} = \frac{1}{c}\frac{\partial}{\partial t} + \vec{\nabla}$$

in either signature.

Putting these together, we split eq. (6.5) within the $\boldsymbol{\gamma}_0$-frame by left-multiplying by $\boldsymbol{\gamma}_0$;

$$\begin{aligned} \boldsymbol{\gamma}_0\,\boldsymbol{\partial}F &= \boldsymbol{\gamma}_0\mu_0\boldsymbol{J} \\ = \left(\frac{1}{c}\frac{\partial}{\partial t} + \vec{\nabla}\right)\left(\frac{1}{c}\vec{E} + \mathbb{I}\vec{B}\right) &= \mu_0\left(c\rho - \vec{J}\right). \end{aligned}$$

By expanding and equating grades, we obtain four equations,

$$\begin{aligned} \frac{1}{c}\vec{\nabla}\cdot\vec{E} &= \mu_0 c\rho && \text{(scalar)} \\ \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t} + \mathbb{I}(\vec{\nabla}\wedge\vec{B}) &= -\mu_0\vec{J} && \text{(vector)} \\ \frac{1}{c}\vec{\nabla}\wedge\vec{E} + \frac{\mathbb{I}}{c}\frac{\partial\vec{B}}{\partial t} &= 0 && \text{(bivector)} \\ \mathbb{I}(\vec{\nabla}\cdot\vec{B}) &= 0 && \text{(pseudoscalar)} \end{aligned}$$

Note that the cross product relates to the bivector curl in $\mathscr{G}(3)$ by

$$\boldsymbol{u} \wedge \boldsymbol{v} = \mathbb{I}(\boldsymbol{u} \times \boldsymbol{v}) \quad \text{so that} \quad \nabla \times \boldsymbol{X} = -\mathbb{I}(\vec{\nabla} \wedge \vec{X}).$$

Hence, by adjusting by factors of c and $\mathbb{I}$ (and using $\mu_0\varepsilon_0 c^2 = 1$), the above equations reduce immediately to Gauß's law, Ampère's law, Faraday's law and the magnetic monopole equation, respectively.

The calculations in this section were performed assuming $\eta = \operatorname{diag}(+---)$. In the $(-+++)$ signature, $\boldsymbol{\gamma}_0\boldsymbol{J} = -c\rho + \vec{J}$ differs by an overall sign, which is absorbed by $F \mapsto -F$.

Part II.

General Relativity and Manifold Geometry

Chapter 7.

Spacetime as a Manifold

The investigations of part I were restricted to *flat geometries*. Special relativity models spacetime as a homogeneous, isotropic Minkowski vector space. However, in the general theory of relativity, spacetime no longer has an intrinsic vector space structure, instead exhibiting curvature to incorporate gravity. The mathematical demands of curvature call for the *differential geometry of smooth manifolds*.

Here we give a pragmatic definition of a manifold as a space which locally looks like $\mathbb{R}^n$ upon which one can do calculus.[53]

[53] See [21, §1] for rigorous definitions in terms of charts and atlases.

Definition 25. *A* MANIFOLD $\mathscr{M}$ OF DIMENSION n *is a nice*[54] *topological space which is locally Euclidean, meaning for every $x \in \mathscr{M}$ there exist neighbourhoods $x \in \mathscr{U} \subseteq \mathscr{M}$ and subsets $U \subseteq \mathbb{R}^n$ with a homeomorphism (continuous bijection) $\mathscr{U} \hookrightarrow\!\!\!\!\rightarrow U$ between them.*

[54] Here, a 'nice' topological space is:

1. *Hausdorff*, meaning each distinct pair of points have mutually disjoint neighbourhoods (so it is "not too small"); and
2. *second-countable*, meaning there exists a countable base (so it is "not too large").

A SMOOTH MANIFOLD *is a manifold with the stricter requirement that $\mathscr{U} \hookrightarrow\!\!\!\!\rightarrow U$ be a diffeomorphism (differentiable bijection).*

Essentially, definition 25 is designed to guarantee that well-behaved local coordinates (corresponding to vectors in U) always exist.

Definition 26. *Let $\mathscr{M}$ be an n-dimensional manifold. A* (GLOBAL) COORDINATE CHART $\{x^i\} \equiv \{x^1, \dots, x^n\}$ *of $\mathscr{M}$ is a set of scalar fields $x^i : \mathscr{M} \to \mathbb{R}$ such that each point in $\mathscr{M}$ is specified uniquely by the coordinate values $(x^1, \dots, x^n) \in \mathbb{R}^n$.*

A LOCAL COORDINATE CHART *about a point $x \in \mathscr{M}$ is a coordinate chart of a neighbourhood of x.*

In other words, the coordinate functions $\{x^i\}$ specify a homeomorphism $\mathscr{M} \to \mathbb{R}^n$, or a diffeomorphism in the case of a smooth chart. We will often call a point $x \in \mathscr{M}$ by the same symbol as the local coordinates $x^i : \mathscr{M} \to \mathbb{R}$ without the index — but these objects are not interchangeable.

A structure-preserving map between manifolds is a continuous function; and between smooth manifolds, a differentiable function. For brevity, we assume the definitions that follow take place in the category of manifolds, and take *all maps between manifolds to be continuous.* Furthermore, if the qualifier "smooth" is present, we operate in the category of smooth manifolds and such maps are assumed differentiable. Thus, the coordinate scalars x^i of definition 26 are continuous functions, and are differentiable if the manifold is smooth, etcetera.

7.1. Derivatives of Smooth Maps

Manifolds themselves do not have inherent vector space structure. However, being locally Euclidean means there is a real vector space naturally associated to each point:

Definition 27. *The* TANGENT SPACE $\mathrm{T}_x \mathscr{M}$ *of a manifold at a point $x \in \mathscr{M}$ is the vector space of derivations on smooth functions at that point.*[55] *In any local coordinate chart $\{x^i\}_{i=1}^n$ of $\mathscr{M}$ containing x, this is*

$$\mathrm{T}_x \mathscr{M} \cong \operatorname{span}\left\{ \left.\frac{\partial}{\partial x^i}\right|_x \right\}_{i=1}^n .$$

The TANGENT BUNDLE $\mathrm{T}\mathscr{M}$ *is the disjoint union of all tangent spaces*

$$\mathrm{T}\mathscr{M} = \{(x, \boldsymbol{u}) \mid x \in \mathscr{M}, \boldsymbol{u} \in \mathrm{T}_x \mathscr{M}\}$$

equipped with an appropriate manifold topology.[56]

[55] More precisely, each vector $\boldsymbol{u} \in \mathrm{T}_x \mathscr{M}$ is an equivalence class of derivatives evaluated at the point x, where different derivations which agree at the point x are identified.

[56] Specifically, the topology of a fibre bundle (see section 7.2).

Given a smooth manifold, its tangent bundle comes for free: its construction is canonical and requires no additional data. Similarly, given a smooth function f between manifolds, its derivative $\mathrm{d}f$ (i.e., its 'tangent') also comes for free.

In the same way that the tangent bundle consists of 'directional derivatives of points' in the manifold (i.e., tangent vectors), the differential $\mathrm{d}f$ encodes the directional derivatives of f at all points in the domain.[57] Intuitively, if $\boldsymbol{u} \in \mathrm{T}_x \mathscr{M}$ is a vector at a point $x \in \mathscr{M}$, then the vector $\mathrm{d}f(\boldsymbol{u}) \in \mathrm{T}_{f(x)} \mathscr{N}$ is interpreted as the derivative of $f(x) \in \mathscr{N}$ in the direction $\boldsymbol{u}$.

[57] This parallel is precise: d and T form a functor in category of smooth manifolds, sending $f : \mathscr{M} \to \mathscr{N}$ to $\mathrm{d}f : \mathrm{T}\mathscr{M} \to \mathrm{T}\mathscr{N}$. Some authors use the symbol T for both.

Definition 28. *The* DIFFERENTIAL *or* PUSH FORWARD *of a map $f : \mathscr{M} \to \mathscr{N}$ between smooth manifolds is the map $\mathrm{d}f : \mathrm{T}\mathscr{M} \to \mathrm{T}\mathscr{N}$ defined by*

$$(\mathrm{d}f(\boldsymbol{u}))(\varphi)\big|_{f(x)} := \boldsymbol{u}(\varphi \circ f)\big|_x \tag{7.1}$$

for each point $x \in \mathscr{M}$, vector $\boldsymbol{u} \in \mathrm{T}_x \mathscr{M}$ and smooth function $\varphi : \mathscr{N} \to \mathbb{R}$.

In the definition above, vectors act on scalar functions as derivations; hence $\mathrm{d}f(\boldsymbol{u})$ is defined by its action on an arbitrary scalar field.

Note that $\mathrm{d}f(\boldsymbol{u})$ may not be defined everywhere on $\mathscr{N}$. If $\boldsymbol{u}|_x \in \mathrm{T}_x \mathscr{M}$ is now a family of vectors defined everywhere over $x \in \mathscr{M}$, then $\mathrm{d}f(\boldsymbol{u})|_{f(x)} = \mathrm{d}f(\boldsymbol{u}|_x)$ is defined only at each $f(x) \in \mathscr{N}$. This means that if f fails to be surjective, then $\mathrm{d}f(\boldsymbol{u})$ is not defined at those points lying outside the image $f(\mathscr{M}) \subset \mathscr{N}$. Likewise, if f fails to be injective at a point $y \in \mathscr{N}$, then $\mathrm{d}f(\boldsymbol{u})$ is *multivalued* at y. Only if f is bijective does $\mathrm{d}f(\boldsymbol{u})|_y$ have a single value everywhere.

The meaning of definition 28 may become clearer when expressed in coordinates. Suppose $\{x^i\}$ is a local chart of $\mathscr{M}$ containing a point $x \in \mathscr{M}$, and $\{y^j\}$ a chart of $\mathscr{N}$ containing $f(x)$. With associated coordinate bases $\mathrm{T}_x \mathscr{M} = \mathrm{span}\{\frac{\partial}{\partial x^i}\}$ and $\mathrm{T}_{f(x)} \mathscr{N} = \mathrm{span}\{\frac{\partial}{\partial y^j}\}$, eq. (7.1) takes the full form:

$$\left[\mathrm{d}f\left(u^i \frac{\partial}{\partial x^i}\right)\right]^j \frac{\partial \varphi}{\partial y^j}\bigg|_{f(x)} = u^i \frac{\partial \varphi \circ f}{\partial x^i}\bigg|_x = u^i \frac{\partial y^j \circ f}{\partial x^i}\bigg|_x \frac{\partial \varphi}{\partial y^j}\bigg|_{f(x)}$$

The first equality is the definition itself, and the second is an application of the chain rule. Since φ is an arbitrary smooth function, this holds as an equation of differential operators, and we may remove reference to any particular φ on which the operators act.

$$\left[\mathrm{d}f(u^i\partial_i)\right]^j \partial_j\Big|_{f(x)} = u^i \frac{\partial f^j}{\partial x^i}\Big|_x \partial_j\Big|_{f(x)} \tag{7.2}$$

We reduce typographical complexity with $\partial_i := \frac{\partial}{\partial x^i}$ and $\partial_j := \frac{\partial}{\partial y^j}$, being aware that these are basis vectors of *different* tangent spaces. We also abbreviate $f^j := y^j \circ f$ so that $f^j(x)$ is the jth coordinate of the point $f(x)$ in the y^j chart. Thus, the coordinate form of $\mathrm{d}f$ is precisely the Jacobian matrix,

$$[\mathrm{d}f(\partial_i)]^j = \frac{\partial f^j}{\partial x^i}.$$

Turning back to eq. (7.2), the partial derivatives $\partial/\partial x^i$ act on smooth functions $f^j : \mathscr{M} \to \mathbb{R}$ to produce smooth functions $\partial f^j/\partial x^i : \mathscr{M} \to \mathbb{R}$. However, since we have an intuitive picture of the directional derivative of the point $f(x)$ as x is displaced, it is useful to formally extend the notation $\partial/\partial x^i$ so that we may write the partial derivative of a mapping of *points* $f : \mathscr{M} \to \mathscr{N}$. That is, $\partial f/\partial x^i\big|_x \in \mathrm{T}_{f(x)}\mathscr{N}$ is the infinitesimal displacement vector of $f(x) \in \mathscr{N}$ caused by an infinitesimal variation in the ith coordinate of the source point x. This is the meaning of the last term in eq. (7.2), so the desired shorthand is

$$\frac{\partial f}{\partial x^i} := \frac{\partial f^j}{\partial x^i}\partial_j \quad \text{or, in full,} \quad \frac{\partial f}{\partial x^i}\Big|_x := \frac{\partial y^j \circ f}{\partial x^i}\Big|_x \frac{\partial}{\partial y^j}\Big|_{f(x)}.$$

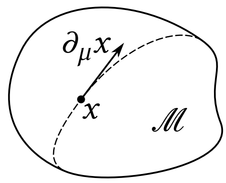


Figure 7.1.: The derivative of the point $x \in \mathscr{M}$ along the direction of increasing x^μ is a tangent vector $\partial_\mu x \in \mathrm{T}_x\mathscr{M}$. The vector is tangent to the dotted line, along which all coordinates but x^μ are constant.

With this, eq. (7.2) may be written as

$$\mathrm{d}f(\boldsymbol{u}) = u^i\frac{\partial f}{\partial x^i}. \tag{7.3}$$

This condensed notation is useful, despite being implicit: take for instance the coordinate functions $x^i : \mathscr{M} \to \mathbb{R}$ regarded as maps between manifolds. Then eq. (7.3) yields the defining property of the coordinate dual basis,

$$\mathrm{d}x^i(\partial_j) = \frac{\partial x^i}{\partial x^j} = \delta^i_j,$$

where we have identified the one-dimensional vector space $\mathrm{T}_{x^i}\mathbb{R}$ with $\mathbb{R}$ itself.

Lemma 19 (Chain rule). *If $f \circ g$ is a composition of maps between smooth manifolds, then* $\mathrm{d}(f \circ g) = \mathrm{d}f \circ \mathrm{d}g$.

Proof. Acting on a vector $\boldsymbol{u}$ and applying the forward-pushed vector to a scalar field φ, we obtain

$$\begin{aligned}(\mathrm{d}(f \circ g)(\boldsymbol{u}))(\varphi) &= \boldsymbol{u}(\varphi \circ f \circ g)\\ &= \boldsymbol{u}((\varphi \circ f) \circ g) = (\mathrm{d}g(\boldsymbol{u}))(\varphi \circ f) = \mathrm{d}f(\mathrm{d}g(\boldsymbol{u}))(\varphi)\end{aligned}$$

by three applications of definition 28. □

7.2. Fibre Bundles

In flat geometries, fields were modelled as functions into a fixed vector space. For example, in flat spacetime $\mathscr{M} = \mathbb{R}^{1+3}$, the electromagnetic bivector $F : \mathscr{M} \to \wedge^2\mathbb{R}^4$ makes no distinction between the vector space $\wedge^2\mathbb{R}^4$ evaluated at one point in spacetime over another. This would suggest that all values of a field are directly comparable, making expressions like $F(x) + F(y) \in \wedge^2\mathbb{R}^4$ geometrically meaningful for different points $x, y \in \mathscr{M}$. However, these kinds of expressions are ill-defined for general smooth manifolds $\mathscr{M}$, since they depend on the way tangent spaces are chosen. Instead, it is beneficial to distinguish between codomains *at each point in the domain*, and treat $F(x)$ and $F(y)$ as belonging to different spaces.

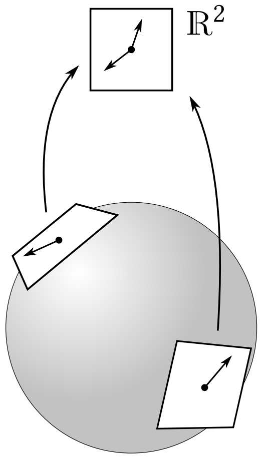


Figure 7.2.: Vectors in different tangent spaces, and their basis-dependent representation as an $\mathbb{R}^2$-valued field.

A motivating example is fluid flowing on a sphere: The instantaneous fluid velocity at a point is a vector lying in the sphere's tangent plane at that point. If the fluid flow is given as a map $f : \mathscr{S}^2 \to \mathbb{R}^2$, then any two velocity vectors exist in the "same" space, though *geometrically* they do not (fig. 7.2). This also means the fluid flow's representation as a field $f : \mathscr{S}^2 \to \mathbb{R}^2$ is *dependent on the choice of basis* — i.e., the way in which the single codomain $\mathbb{R}^2$ is identified with each tangent plane on the sphere. A basis-independent representation of f requires treating each tangent space as distinct.

Doing this, we are led to the tangent *bundle* $\mathrm{T}\,\mathscr{S}^2$, where all the tangent planes of $\mathscr{S}^2$ are collected in a disjoint union. The vector field on the sphere now becomes a *section* of $\mathrm{T}\,\mathcal{S}^2$, which is a map $f : \mathcal{S}^2 \to \mathrm{T}\,\mathcal{S}^2$ such that $f(x)$ belongs to the tangent space at x. No longer is the expression $f(x) + f(y)$ well-defined.

The tangent bundle is a special case of a *fibre bundle*, which is a manifold consisting of disjoint copies of a space (called the *fibre*) taken at every point in a base manifold.

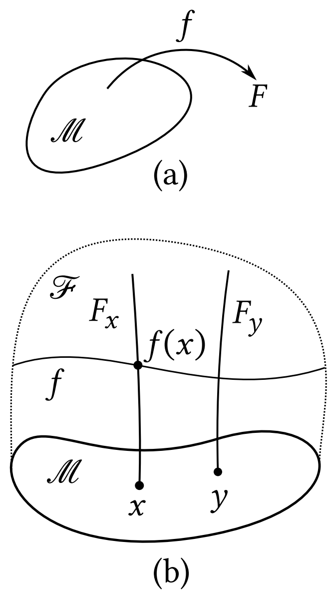


Figure 7.3.: (a) A field $f : \mathscr{M} \to F$, where values at any point can be compared. (b) A fibre bundle $F \hookrightarrow \mathscr{F} \twoheadrightarrow \mathscr{M}$ with a section $f \in \Gamma(\mathscr{F})$ whose individual fibres F are labelled by base point in $\mathscr{M}$.

Definition 29. *A* FIBRE BUNDLE $F \hookrightarrow \mathscr{F} \overset{\pi}{\twoheadrightarrow} \mathscr{M}$ *consists of*

- *a* BULK MANIFOLD $\mathscr{F}$*;*
- *a* BASE MANIFOLD $\mathscr{M}$*; and*
- *a surjection* $\pi : \mathscr{F} \to \mathscr{M}$*, the* PROJECTION*, such that*
- *the inverse image* $F_x \coloneqq \pi^{-1}(x)$ *of a base point* $x \in \mathscr{M}$ *is homeomorphic to the* FIBRE F.

Definition 29 takes place in the category of manifolds, so the projection $\pi : \mathscr{F} \to \mathscr{M}$ is continuous. In a SMOOTH FIBRE BUNDLE, the projection π is differentiable and F, $\mathscr{F}$ and $\mathscr{M}$ are all smooth manifolds.

Many different kinds of fibre bundle may be considered by giving F more structure. For example,

- a VECTOR BUNDLE is one where the fibre is a vector space;
- a PRINCIPLE BUNDLE is one where the fibre is a group (usually a Lie group); and
- an ALGEBRA BUNDLE is a vector bundle where each fibre is equipped with a (smoothly varying) algebraic product; and so on.

Trivialisations and coordinates

The bulk $\mathscr{F}$ of a fibre bundle $F \hookrightarrow \mathscr{F} \twoheadrightarrow \mathscr{M}$ is itself a manifold (of dimension $\dim \mathscr{F} = \dim \mathscr{M} + \dim F$) so we may always prescribe local coordinates on $\mathscr{F}$. If we already have coordinates $\{x^\mu\}$ on the base $\mathscr{M}$ and $\{x^a\}$ on a fibre F, then we often want to use the same coordinates $\{x^\mu, x^a\}$ to describe the bulk $\mathscr{F}$. This first requires a way of continuously splitting the bulk $\mathscr{F} \to \mathscr{M} \times F$ into its base and fibre "components", in a way which respects the fibred structure of the bundle. This splitting is known as a *trivialisation* of the bundle.

Definition 30. *A* TRIVIALISATION *of a fibre bundle* $F \hookrightarrow \mathscr{F} \overset{\pi}{\twoheadrightarrow} \mathscr{M}$ *is a homeomorphism* $\varphi : \mathscr{F} \to \mathscr{M} \times F$ *such that* $\mathrm{pr}_1 \circ \varphi = \pi$.

It is not always possible to find a global trivialisation of a fibre bundle, and if it is, the bundle is called a TRIVIAL FIBRE BUNDLE and there may be different possible trivialisations.[58]

[58] A simple non-trivial fibre bundle is the Möbius strip, viewed as a bundle over the circle $\mathcal{S}^1$ with fibre $[0,1]$. The trivial bundle $\mathcal{S}^1 \times [0,1]$ describes a strip without a twist.

However, it is always possible trivialise *locally*. That is, for any base point $x \in \mathscr{M}$, there exists a neighbourhood $x \in U \subseteq \mathscr{M}$ for which the subbundle $F \hookrightarrow \pi^{-1}(U) \overset{\pi}{\twoheadrightarrow} U$ admits a (global) trivialisation. Hence, it is always possible to assign *local* coordinates $\{x^\mu, x^a\}$ to the bulk of a fibre bundle, where x^μ are coordinates on the base and x^a are coordinates on the fibres, such that x^μ do not vary along the fibres. In other words, local trivialisations are equivalent to local coordinates.

Sections of fibre bundles

In the language of fibre bundles, a field $f : \mathscr{M} \to F$ is replaced by a *section*, which is a "vertical" map $f : \mathscr{M} \to \mathscr{F}$ into the bulk $\mathscr{F}$ such that $f(x) \in F_x$.

Definition 31. *A* SECTION f *of a fibre bundle* $F \hookrightarrow \mathscr{F} \overset{\pi}{\twoheadrightarrow} \mathscr{M}$ *is a right-inverse of* π. *The space of sections is denoted*

$$\Gamma(\mathscr{F}) = \{f : \mathscr{M} \to \mathscr{F} \mid \pi \circ f = \mathrm{id}\}.$$

(Again, sections $f \in \Gamma(\mathscr{F})$ are assumed continuous, and SMOOTH SECTIONS are sections of smooth fibre bundles for which f is smooth.)

For example, the instantaneous fluid velocity $\boldsymbol{u}$ on a sphere $\mathcal{S}^2$ is a section $\boldsymbol{u} \in \Gamma(\mathrm{T}\,\mathcal{S}^2)$ of the tangent bundle, with a single vector at $x \in \mathcal{S}^2$ is denoted $\boldsymbol{u}|_x \in \mathrm{T}_x\,\mathcal{S}^2$.

7.2.1. Exterior differential forms on manifolds

Section 2.2.2 defined exterior forms $\Omega(V, A)$ as alternating multilinear maps from the fixed vector space V into A. Similarly, the differential of a map $\mathrm{d}f : \mathrm{T}\,\mathcal{M} \to \mathrm{T}\,\mathcal{N}$ is an object that takes a vector argument $\boldsymbol{u} \in \mathrm{T}\,\mathcal{M}$ — just like an exterior 1-form, except that the entire tangent bundle $\mathrm{T}\,\mathcal{M}$ is not itself a vector space.

Exterior forms, which are alternating maps from a fixed space $V^{\otimes}$, can be extended to exterior *differential* forms, which exist on manifolds and define alternating maps from $(\mathrm{T}_x\,\mathcal{M})^{\otimes}$ at each $x \in \mathcal{M}$.

Although the entire bundle $\mathrm{T}\,\mathcal{M}$ is not a vector space, the space of vector sections $\Gamma(\mathrm{T}\,\mathcal{M})$ is. Hence, when viewed as a map $\mathrm{d}f : \Gamma(\mathrm{T}\,\mathcal{M}) \to \Gamma(\mathrm{T}\,\mathcal{N})$ the differential $\mathrm{d}f \in \Omega^1(\Gamma(\mathrm{T}\,\mathcal{M}), \Gamma(\mathrm{T}\,\mathcal{N}))$ is a $\Gamma(\mathrm{T}\,\mathcal{N})$-valued exterior 1-form (by definition 11). This mouthful may be eased by defining the notation

$$\Omega(\mathcal{M}, \mathcal{E}) \coloneqq \Omega(\Gamma(\mathrm{T}\,\mathcal{M}), \Gamma(\mathcal{E}))$$

for some vector bundle $\mathcal{M} \hookrightarrow \mathcal{E} \twoheadrightarrow V$. As with exterior forms, the wedge product is defined as in eq. (2.6).

An element of $\Omega^k(\mathcal{M}, \mathcal{E})$ is called an $\mathcal{E}$-VALUED EXTERIOR DIFFERENTIAL k-FORM, where 'differential' distinguishes it as an object on a manifold. Scalar-valued exterior differential forms are elements of $\Omega^k(\mathcal{M}) \coloneqq \Omega^k(\mathcal{M}, \mathcal{M} \times \mathbb{R})$, where $\mathcal{M} \hookrightarrow \mathcal{M} \times \mathbb{R} \overset{\pi}{\twoheadrightarrow} \mathbb{R}$ is the trivial line bundle with projection $\pi(x, \lambda) = \lambda$.

We sometimes use the notation $\underset{\sim}{\alpha}$ to emphasise that α is an exterior differential form. {TO SELF: Not necessary for differentials $\mathrm{d}f$ since is clear.}

{TO DO: Formula for exterior derivative.}

7.3. Lie Derivatives

In general, the derivative of a section of a fibre bundle is not defined, because there is no way of comparing fibres without additional structure (such as a *connection*; see chapter 8). For some kinds of object, however, it is possible to define transport between fibres using the *flow* of a tangent vector section. We call objects for which this is possible FLOWABLE.

Thus, the value of a flowable object at a point x to may be directly compared to its value at some other point y by flowing the y-value back to the x-fibre. This enables the definition of a kind of derivative with respect to the flow — a construction called the *Lie derivative*.

Definition 32. *The FLOW of $\boldsymbol{u}$ is the map $\mathrm{fl}_{\boldsymbol{u}}^t : \mathscr{M} \to \mathscr{M}$ satisfying*

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{fl}_{\boldsymbol{u}}^t(x)\bigg|_y = \boldsymbol{u}|_y$$

for all values of t.

Definition 33. *The LIE DERIVATIVE of a flowable object A along a tangent section $\boldsymbol{u} \in \Gamma(\mathrm{T}\mathscr{M})$ is*

$$\pounds_{\boldsymbol{u}} A := \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{fl}_{\boldsymbol{u}}^{-t} A\bigg|_{t=0}.$$

Scalar sections $f : \mathscr{M} \to \mathbb{R}$ are flowable by defining $\mathrm{fl}_{\boldsymbol{u}}^t f := e^{-t\boldsymbol{u}} f$. For example, in one dimension, $\mathrm{fl}_{\partial_x}^t f = e^{-t\partial_x} f(x) = f(x-t)$ is the Taylor series of f translated by t. Tangent vectors $\boldsymbol{v} \in \Gamma(\mathrm{T}\mathscr{M})$ are also flowable, using the differential of a flow $\mathrm{d}\left(\mathrm{fl}_{\boldsymbol{u}}^t\right) : \mathrm{T}\mathscr{M} \to \mathrm{T}\mathscr{M}$. Thus, we also define

$$\mathrm{fl}_{\boldsymbol{u}}^t \boldsymbol{v} := \mathrm{d}\left(\mathrm{fl}_{\boldsymbol{u}}^t\right)(\boldsymbol{v})$$

in terms of the flow of points.[59] Other flowable objects include structures built from the tangent bundle, e.g., tangent tensors $(\mathrm{T}\mathscr{M})^{\otimes}$ or multivectors $\mathscr{G}(\mathrm{T}\mathscr{M}, \eta)$. {TO DO: This is the first appearance of $\mathscr{G}(\mathrm{T}\mathscr{M}, \eta)$...}

[59] Risking overloaded notation, fl_u^t on the left-hand side indicates the flow of vectors, while on the right-hand side the flow of points.

Lemma 20. *The Lie derivative on scalars is* $\pounds_{\boldsymbol{u}} f = \boldsymbol{u}(f)$*, and on tangent vectors is the Lie bracket,* $\pounds_{\boldsymbol{u}} \boldsymbol{v} = [\boldsymbol{u}, \boldsymbol{v}] \coloneqq \boldsymbol{u} \circ \boldsymbol{v} - \boldsymbol{v} \circ \boldsymbol{u}$.

Proof. For scalars, the result follows from $\pounds_{\boldsymbol{u}} f = \frac{\mathrm{d}}{\mathrm{d}t} e^{-t\boldsymbol{u}} f \big|_{t=0} = \boldsymbol{u}(f)$.

For tangent vectors, unpacking definition 33 for a vector argument, and then using definition 28 to rewrite the pushforward, we have

$$(\pounds_{\boldsymbol{u}} \boldsymbol{v}) f|_x = \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{d}\left(\mathrm{fl}_{\boldsymbol{u}}^{-t}\right) \left(\boldsymbol{v}|_{\mathrm{fl}_{\boldsymbol{u}}^{t}(x)}\right) f \bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{v}(f \circ \mathrm{fl}_{\boldsymbol{u}}^{-t})|_{\mathrm{fl}_{\boldsymbol{u}}^{t}(x)} \bigg|_{t=0}.$$

By the product rule over the two appearances of t, this is equal to

$$\boldsymbol{v}\left(\frac{\mathrm{d}}{\mathrm{d}t} f \circ \mathrm{fl}_{\boldsymbol{u}}^{-t} \bigg|_{t=0}\right)\bigg|_x + \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{v}(f)|_{\mathrm{fl}_{\boldsymbol{u}}^{t}(x)} \bigg|_{t=0}. \tag{7.4}$$

Using the chain rule (lemma 19) and definition 32, we have $\frac{\mathrm{d}}{\mathrm{d}t} g \circ \mathrm{fl}_{\boldsymbol{u}}^{t} \big|_{t=0} = \mathrm{d}g(\boldsymbol{u}) = \boldsymbol{u}(g)$. Taking g to be f and $\boldsymbol{v}(f)$ for the left- and right-hand terms of eq. (7.4) respectively, we find

$$(\pounds_{\boldsymbol{u}} \boldsymbol{v}) f = -\boldsymbol{v}(\boldsymbol{u}(f)) + \boldsymbol{u}(\boldsymbol{v}(f))$$

which is the lie bracket acting on the arbitrary scalar section f. □

7.3.1. On tensors and differential forms

By requiring $\pounds_{\boldsymbol{u}}$ to be a derivation, we can deduce from $\pounds_{\boldsymbol{u}}\, \varphi(\boldsymbol{v}) = (\pounds_{\boldsymbol{u}} \varphi)(\boldsymbol{v}) + \varphi(\pounds_{\boldsymbol{u}} \boldsymbol{v})$ the form of the Lie derivative on a covector φ. Continuing in this way, it follows that the Lie derivative of a general tensor $T = T^{\mu_1 \cdots \mu_p}{}_{\nu_1 \dots \nu_q} \boldsymbol{e}_{\mu_1} \otimes \cdots \otimes \boldsymbol{e}_{\mu_p} \otimes \boldsymbol{e}^{\nu_1} \otimes \cdots \otimes \boldsymbol{e}^{\nu_q}$ is

$$\pounds_{\boldsymbol{u}} T^{\mu_1 \cdots \mu_p}{}_{\nu_1 \dots \nu_q} = u^\lambda \partial_\lambda T^{\mu_1 \cdots \mu_p}{}_{\nu_1 \dots \nu_q} - \sum_{i=1}^{p} T^{\mu_1 \dots \lambda \dots \mu_p}{}_{\nu_1 \dots \nu_q} \partial_\lambda u^{\mu_i} + \sum_{i=1}^{q} T^{\mu_1 \cdots \mu_p}{}_{\nu_1 \dots \lambda \dots \nu_q} \partial_{\nu_i} u^\lambda.$$

This sets the stage for how much simpler the form taken by the Lie derivative is on exterior differential forms and multivectors.

On exterior differential forms $\underset{\sim}{\varphi}$, the Lie derivative may be expressed in a basis-free fashion using Cartan's "magic formula"[60]

$$£_{\boldsymbol{u}}\underset{\sim}{\varphi} = \boldsymbol{u} \rfloor \mathrm{d}\underset{\sim}{\varphi} + \mathrm{d}\big(\boldsymbol{u} \rfloor \underset{\sim}{\varphi}\big), \tag{7.5}$$

which employs the INTERIOR DERIVATIVE OR HOOK PRODUCT $\boldsymbol{u} \rfloor : \Omega^k(V) \to \Omega^{k-1}(V)$ defined by $(\boldsymbol{u} \rfloor \underset{\sim}{\varphi})(\boldsymbol{u}_2 \otimes \cdots \otimes \boldsymbol{u}_k) = \underset{\sim}{\varphi}(\boldsymbol{u} \otimes \boldsymbol{u}_2 \otimes \cdots \otimes \boldsymbol{u}_k)$. Cartan's magic formula is the statement that the Lie derivative on forms is the anti-commutator of the exterior and interior derivatives.

60 *Sketch proof.* d and $\boldsymbol{u} \rfloor$ are anti-derivations, so their anti-commutator is a derivation. Derivations agreeing on scalars and exact 1-forms (which generate the exterior algebra) are equal. Indeed, $\boldsymbol{u} \rfloor \mathrm{d}f = \boldsymbol{u}(f) = £_{\boldsymbol{u}} f$ while $\mathrm{d}(\boldsymbol{u} \rfloor f) = 0$; and for exact 1-forms, $\boldsymbol{u} \rfloor \mathrm{d}\underset{\sim}{\varphi} = 0$ while $\mathrm{d}\big(\boldsymbol{u} \rfloor \underset{\sim}{\varphi}\big) = \mathrm{d}\underset{\sim}{\varphi}(\boldsymbol{u}) = £_{\boldsymbol{u}}\underset{\sim}{\varphi}$.

7.3.2. The geometric Lie bracket and derivative

Similar to Cartan's formula eq. (7.5) for exterior differential forms, the Lie derivative admits a simple form when applied to tangent multivectors, i.e., elements of the geometric algebra $\mathscr{G}(\mathrm{T}\mathscr{M}, \eta)$ over the tangent bundle. Insight begins with the generalisation of the vector Lie bracket $[\boldsymbol{u}, \boldsymbol{v}] = \boldsymbol{u} \circ \boldsymbol{v} - \boldsymbol{v} \circ \boldsymbol{u}$ to multivectors.

Definition 34. *The* GEOMETRIC LIE BRACKET *of two multivectors $A, B \in \mathscr{G}(\mathrm{T}\mathscr{M}, \eta)$ is*

$$[A, B] \coloneqq (A \lfloor \boldsymbol{\partial}) \wedge B - (B \lfloor \boldsymbol{\partial}) \wedge A,$$

where $\boldsymbol{\partial}$ acts on the multivector to its immediate right.

When acting on vectors, definition 34 reduces to the standard vector Lie bracket,[61]

$$(\boldsymbol{u} \lfloor \boldsymbol{\partial}) \wedge \boldsymbol{v} - (\boldsymbol{v} \lfloor \boldsymbol{\partial}) \wedge \boldsymbol{u} \equiv \boldsymbol{u} \cdot \boldsymbol{\partial}\boldsymbol{v} - \boldsymbol{v} \cdot \boldsymbol{\partial}\boldsymbol{u} = [\boldsymbol{u}, \boldsymbol{v}],$$

so the use of the same notation $[\ ,\ ]$ is appropriate. However, definition 34 is a significant generalisation of the vector Lie bracket, applicable to multivectors of arbitrary grade.

61 $\boldsymbol{u} \lfloor \boldsymbol{\partial} = \boldsymbol{u} \cdot \boldsymbol{\partial} = \partial_{\boldsymbol{u}}$ are scalar operators, so the wedge product becomes scalar multiplication.

Theorem 4. *Let $A \in \mathscr{G}(\mathrm{T}\mathscr{M}, \eta)$ be a multivector and $\boldsymbol{u} \in \mathrm{T}\mathscr{M}$ a tangent vector. The Lie derivative of A along $\boldsymbol{u}$ is*

$$£_{\boldsymbol{u}} A = [\boldsymbol{u}, A]. \tag{7.6}$$

This is a remarkably elegant result: it applies to multivectors of any kind (vectors, k-blades, even inhomogeneous rotors) and the Lie derivative has the same simple form. {TO DO: Contrast this to the Lie derivative of a general tensor, eq. (?).}

Proof. Since $\pounds_{\boldsymbol{u}}$ is linear, it suffices to prove the case where $A = \boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{a}_k$ is a k-blade. Because $\pounds_{\boldsymbol{u}}$ is a derivation, we must have the result that

$$\pounds_{\boldsymbol{u}}(\boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{a}_k) = \sum_{i=1}^{k} \boldsymbol{a}_1 \wedge \cdots \wedge [\boldsymbol{u}, \boldsymbol{a}_i] \wedge \cdots \wedge \boldsymbol{a}_k \tag{7.7}$$

since $\pounds_{\boldsymbol{u}} \boldsymbol{a}_i = [\boldsymbol{u}, \boldsymbol{a}_i]$ is the vector Lie bracket. Expanding the right-hand side of eq. (7.6), we have, by definition 34

$$[\boldsymbol{u}, A] = \boldsymbol{u} \cdot \boldsymbol{\partial} A - (A \lfloor \boldsymbol{\partial}) \wedge \boldsymbol{u}.$$

We will expand the two terms on the right-hand side.

The first term is

$$\boldsymbol{u} \cdot \boldsymbol{\partial} A = \boldsymbol{u} \cdot \boldsymbol{\partial}(\boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{a}_k) = \sum_{i=1}^{k} \boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{u} \cdot \boldsymbol{\partial} \boldsymbol{a}_i \wedge \cdots \wedge \boldsymbol{a}_k \tag{7.8}$$

since $\boldsymbol{u} \cdot \boldsymbol{\partial} \equiv \partial_{\boldsymbol{u}}$ is a scalar derivation.

The second term is $(A \lfloor \boldsymbol{\partial}) \wedge \boldsymbol{u}$. Recall that contraction by a vector is an anti-derivation {TO DO: show this}. Thus, for some vector $\boldsymbol{v}$,

$$\boldsymbol{v} \rfloor A = \boldsymbol{v} \rfloor (\boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{a}_k) = \sum_{i=1}^{k} (-1)^{i-1} \boldsymbol{a}_1 \wedge \cdots \wedge (\boldsymbol{v} \cdot \boldsymbol{a}_i) \wedge \cdots \wedge \boldsymbol{a}_k.$$

Wedging this with a vector $\boldsymbol{u}$ produces

$$\boldsymbol{u} \wedge (\boldsymbol{v} \rfloor A) = \sum_{i=1}^{k} \boldsymbol{a}_1 \wedge \cdots \wedge (\boldsymbol{a}_i \cdot \boldsymbol{v}) \boldsymbol{u} \wedge \cdots \wedge \boldsymbol{a}_k, \tag{7.9}$$

where the factor of $(-1)^{i-1}$ is cancelled by anticommuting $\boldsymbol{u}$ to the ith position. Now, note that A, $\boldsymbol{v} \rfloor A$ and $\boldsymbol{u} \wedge (\boldsymbol{v} \rfloor A)$ are of grades k, $k-1$ and k, respectively, allowing us to exploit reversion to obtain

$$\boldsymbol{u} \wedge (\boldsymbol{v} \rfloor A) = s_k (\boldsymbol{v} \rfloor A)^\dagger \wedge \boldsymbol{u}^\dagger = s_k (A^\dagger \lfloor \boldsymbol{v}^\dagger) \wedge \boldsymbol{u} = (A \lfloor \boldsymbol{v}) \wedge \boldsymbol{u}. \tag{7.10}$$

The notation on the right-hand side lends itself better to the case where $\boldsymbol{v}$ is instead the vector derivative $\boldsymbol{\partial}$ acting on $\boldsymbol{u}$, since $\boldsymbol{u}$ is then to its immediate right. Thus, with eqs. (7.9) and (7.10) we have shown that

$$(A \lfloor \boldsymbol{\partial}) \wedge \boldsymbol{u} = \sum_{i=1}^{k} \boldsymbol{a}_1 \wedge \cdots \wedge (\boldsymbol{a}_i \cdot \boldsymbol{\partial u}) \wedge \cdots \wedge \boldsymbol{a}_k. \tag{7.11}$$

Combining eqs. (7.8) and (7.11) yields

$$[\boldsymbol{u}, A] = \boldsymbol{u} \cdot \boldsymbol{\partial} A - (A \lfloor \boldsymbol{\partial}) \wedge \boldsymbol{u} = \sum_{i=1}^{k} \boldsymbol{a}_1 \wedge \cdots \wedge (\boldsymbol{u} \cdot \boldsymbol{\partial a}_i - \boldsymbol{a}_i \cdot \boldsymbol{\partial u}) \wedge \cdots \wedge \boldsymbol{a}_k$$

whose right-hand side is equal to eq. (7.7). □

{TO DO: Disambiguate $\boldsymbol{u} \cdot \boldsymbol{\partial} A = (\boldsymbol{u} \cdot \boldsymbol{\partial}) A$ from $\boldsymbol{u} \cdot (\boldsymbol{\partial} A)$.}

Chapter 8.

Connections on Fibre Bundles

We have seen that it is more natural to describe physical fields in the language of fibre bundles rather than simply as maps into a fixed codomain. However, with a field $f \in \Gamma(\mathscr{F})$ now formulated as a section of a fibre bundle, it no longer makes sense to directly compare values $f|_x$ at different points $x \in \mathscr{M}$, since each value exists in its own fibre. But the ability to compare across fibres is desirable, particularly because a notion of derivative requires comparing values across 'infinitesimally neighbouring' fibres. To accomplish this, the additional structure of a *connection* on the fibre bundle is required; this then defines the *covariant derivative* of a section.

A trivial example of a connection is the one associated with (the tangent bundle of) Euclidean space. In this case, tangent vectors at a base point may be *parallel transported* (i.e., translated irrotationally) to any other base point in a well-defined, path-independent way.[62] This defines an isomorphism between every tangent space and tangent space at the origin, which is a connection on $\mathrm{T}\,\mathbb{R}^n$.

[62] Any tangent vectors $\boldsymbol{u}_p \in \mathrm{T}\,\mathbb{R}^n \cong \mathbb{R}^n \oplus \mathbb{R}^n$ are compared by translating them to the origin (or discarding the base point) $\boldsymbol{u}_p \equiv (p, \boldsymbol{u}) \mapsto \boldsymbol{u} \in \mathbb{R}^n$.

We may try to define connections on general fibre bundles in this way, by choosing an isomorphism from every fibre to a single 'reference' fibre. This is the same as choosing a trivialisation $\mathscr{F} \to \mathscr{M} \times F$, which identifies every fibre with F (equivalent to prescribing global coordinates on the bundle).

However, defining a connection by a trivialisation like this is a needlessly strict requirement, and is of course impossible to do globally on non-trivial bundles. For instance, the tangent bundle of the sphere $\mathrm{T}\,\mathcal{S}^2$ is non-trivial, so it is impossible to give a globally smooth identification of tangent spaces.[63] However, it is always possible to define a connection *locally* on the sphere, since local trivialisations always exist. In other words, tangent vectors on the sphere can be parallel transported over sufficiently short paths, since locally the sphere looks like the Euclidean plane.

[63] To see this, consider a point on the globe. Given a trivialisation of $\mathrm{T}\,\mathcal{S}^2$, the northward vector is extended to a vector field on the sphere. The hairy ball theorem implies the field vanishes at some point, at which the trivialisation fails.

This generalises to all smooth manifolds: To define a connection, it is only necessary to specify how values are parallel-transported in an infinitesimal manner.

8.1. Connections on General Fibre Bundles

The most general kind of smooth bundle $\mathscr{F}$ is one where the fibres are each diffeomorphic to a manifold F and have no further assumed structure. A point $p \in \mathscr{F}$ in the bundle belongs to the fibre $F_{\pi(p)}$ rooted at the base point $\pi(p) \in \mathscr{M}$. If the point p is moved within its fibre, the base point remains fixed and the motion is said to be "vertical". The tangent space $\mathrm{T}_p\, F_{\pi(p)}$ of the fibre (in isolation from the bulk) consists of those displacement vectors which define vertical motion.

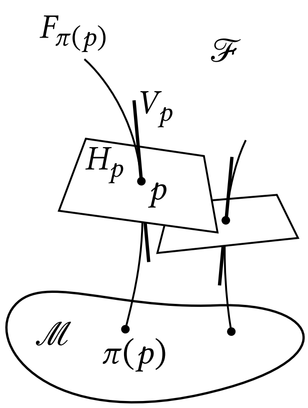


Figure 8.1.: Illustration of an Ehresmann connection.

Definition 35. *The* VERTICAL BUNDLE *of a smooth fibre bundle* $F \hookrightarrow \mathscr{F} \twoheadrightarrow \mathscr{M}$ *is a smooth* $(\dim F)$*-dimensional tangent subbundle* $\mathrm{V}\,\mathscr{F} \subseteq \mathrm{T}\,\mathscr{F}$ *defined by* $\mathrm{V}_p\,\mathscr{F} = \mathrm{T}_p\, F_p$ *for each point* $p \in \mathscr{F}$.

In other words, the tangent bundles of all the fibres taken together form the vertical bundle.

On the other hand, a *connection* specifies how the value $p \in \mathscr{F}$ changes when the base point $\pi(p) \in \mathscr{M}$ moves, if p is to be considered to move "horizontally", i.e., if p is to undergo parallel transport.

Definition 36. *A* HORIZONTAL BUNDLE *or* (EHRESMANN) CONNECTION H *on a smooth fibre bundle* $F \hookrightarrow \mathscr{F} \twoheadrightarrow \mathscr{M}$ *is a smooth* $(\dim \mathscr{M})$*-dimensional* *tangent subbundle $H \subseteq \mathrm{T}\mathscr{F}$ which is complementary to the vertical bundle $V \subseteq \mathrm{T}\mathscr{F}$, in the sense that $\mathrm{T}_p\mathscr{F} = \mathrm{V}_p\mathscr{F} \oplus H_p$ for each point $p \in \mathscr{F}$.*

Note that while the tangent and vertical bundles $\mathrm{T}\mathscr{F}$ and $\mathrm{V}\mathscr{F}$ are canonical constructions, the choice of a horizontal bundle H is not canonical: there may be many distinct horizontal bundles.

The requirement that H be complimentary to $\mathrm{V}\mathscr{F}$ implies $H_p \cap \mathrm{V}_p\mathscr{F} = \{\mathbf{0}\}$ at each $p \in \mathscr{F}$. This means the restriction of $\mathrm{d}\pi : \mathrm{T}_p\mathscr{F} \hookrightarrow\!\!\!\!\rightarrow \mathrm{T}_{\pi(p)}\mathscr{M}$ to $H_p \subseteq \mathrm{T}_p\mathscr{F}$ is an isomorphism.[64] It therefore has an inverse,

$$\mathrm{d}\pi|_{H_p}^{-1} : \mathrm{T}_{\pi(p)}\mathscr{M} \hookrightarrow\!\!\!\!\rightarrow H_p, \tag{8.1}$$

which acts to "lift" tangent vectors from the base into the horizontal subbundle at p. This proves to be a useful construction:

[64] Using the fact that $\ker \mathrm{d}\pi = \mathrm{V}\mathscr{F}$, implying $\ker \mathrm{d}\pi|_{H_p} = \mathbf{0}$.

Definition 37. *Let $F \hookrightarrow \mathscr{F} \overset{\pi}{\twoheadrightarrow} \mathscr{M}$ be a fibre bundle with an Ehresmann connection $H \subseteq \mathrm{T}\mathscr{F}$. The* HORIZONTAL LIFT *to the point $p \in \mathscr{F}$ is the linear map*

$$\Gamma(p) \coloneqq -\mathrm{d}\pi|_{H_p}^{-1} : \mathrm{T}_{\pi(p)}\mathscr{M} \to H_p.$$

Also define the horizontal lift of a section $f \in \mathscr{F}$ at $x \in \mathscr{M}$ by

$$\Gamma(f)|_x \coloneqq -\mathrm{d}\pi|_{H_{f(x)}}^{-1}.$$

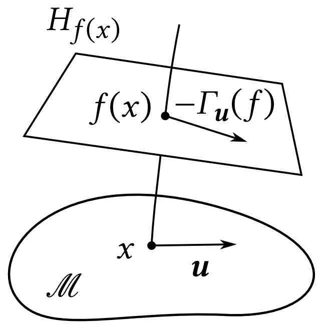


Figure 8.2.: The tangent vector $\boldsymbol{u}$ at x is lifted to the horizontal bulk vector $\Gamma_{\boldsymbol{u}}(f)$ at the point $f(x)$.

The horizontal lift of a section f is a horizontal-valued 1-form $\Gamma(f) \in \Omega^1(\mathscr{M}, H)$, whose action on tangent vectors $\boldsymbol{u}$ we may write as $\Gamma_{\boldsymbol{u}}(f) \coloneqq \Gamma(f)(\boldsymbol{u})$. This device is designed so that tangent vectors $\boldsymbol{u}$ are 'lifted' to horizontal bulk vectors $-\Gamma_{\boldsymbol{u}}(f)$ located on the section f (see fig. 8.2). 'Lifted' means $-\Gamma_{\boldsymbol{u}}(f)$ projects onto $\boldsymbol{u}$, so that we have $-\mathrm{d}\pi(\Gamma_{\boldsymbol{u}}(f)) = \boldsymbol{u}$. The minus sign is present to later align with the convention that a plus sign is present in the covariant derivative of a vector section.[65]

[65] E.g., "$\nabla_\mu X^a = \partial_\mu X^a + \Gamma_\mu{}^a{}_b X^b$".

8.1.1. Parallel transport

With a connection defined on a bundle, a value in the bulk may be moved between fibres so that the motion is always horizontal with respect to the connection. This is called PARALLEL TRANSPORTATION of the value along a path.

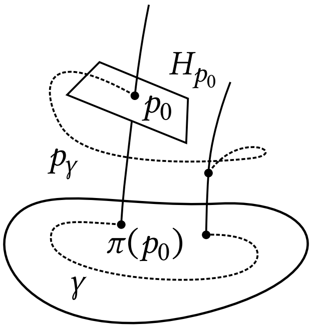


Figure 8.3.: The point p_0 and its parallel transport p_λ along a path γ.

More precisely, a path $\gamma : [0,1] \to \mathscr{M}$ representing the motion of a value $p_0 \in \mathscr{F}$ from $\gamma(0) = \pi(p_0)$ can be LIFTED to a horizontal path $p_0 : [0,1] \to \mathscr{F}$ in the bulk. This path is 'above' γ in the sense that $\pi(p_\gamma(\lambda)) = \gamma(\lambda)$, and 'horizontal' in the sense that $\mathrm{d}p_\gamma(\lambda) \in H_{p_\gamma(\lambda)}$ (see fig. 8.3). In other words, p_γ is a one-dimensional integral manifold of the connection H restricted to the 'wall' $\pi^{-1}(\gamma) \subset \mathscr{F}$.

It is useful to describe this path–lifting process as an operator, which associates maps between fibres to each path in $\mathscr{M}$.

Definition 38. *If $\gamma : [0,1] \to \mathscr{M}$ is a path, then the* TRANSPORT OPERATOR $\mathrm{trans}_\gamma : F_{\gamma(0)} \to F_{\gamma(1)}$ *is defined by* $\mathrm{trans}_\gamma\, p = p_\gamma(1)$ *for any point $p \in F_{\gamma(0)}$ where $p_\gamma : [0,1] \to \mathscr{F}$ is the lifted path satisfying*

$$\pi(p_\gamma(\lambda)) = \gamma(\lambda) \quad \textit{and} \quad \mathrm{d}p_\gamma(\lambda) \in H_{p_\gamma(\lambda)} \tag{8.2}$$

for all $\lambda \in [0,1]$.

The transport operator is invariant under path reparametrisation, since any path $\gamma'(\lambda) = \gamma(f(\lambda))$ where $f : [0,1] \to [0,1]$ is smooth also satisfies eq. (8.2) if γ does. Furthermore, the transport operator respects path concatenation $\gamma_2 * \gamma_1$ and inversion,

$$\underset{\gamma^{-1}}{\mathrm{trans}} = \underset{\gamma}{\mathrm{trans}}^{-1}, \qquad\qquad \underset{\gamma_2*\gamma_1}{\mathrm{trans}} = \underset{\gamma_2}{\mathrm{trans}} \circ \underset{\gamma_1}{\mathrm{trans}}.$$

This makes the transport operator a homomorphism from the groupoid of directed paths[66] (modulo reparametrisation) into the groupoid of fibre isomorphisms.

[66] where the partially-defined group operation is path concatenation

Parallel transport along a path involves integrating the connection, and the horizontal lift is the 'derivative' of the transport operator, in a way made precise in the following lemma.

Lemma 21. *The transport operator along a path γ satisfies the ordinary differential equation*

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \underset{\gamma(\lambda\leftarrow 0)}{\mathrm{trans}} = -\Gamma_{\dot{\gamma}(\lambda)} \circ \underset{\gamma(\lambda\leftarrow 0)}{\mathrm{trans}}, \tag{8.3}$$

where $\gamma(\lambda \leftarrow 0)$ denotes the sub-path of γ from $\gamma(0)$ to $\gamma(\lambda)$.

Proof. If $p \in F_{\gamma(0)}$ then we have $\operatorname{trans}_{\gamma(\lambda\leftarrow 0)} p = p_\gamma(\lambda)$ where p_γ is the lift of γ through p, satisfying the conditions in definition 38. Differentiating with respect to λ,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \operatorname*{trans}_{\gamma(\lambda\leftarrow 0)} p = \mathrm{d}p_\gamma(\lambda) \in H_{p_\gamma(\lambda)}, \tag{8.4}$$

which is the horizontal by eq. (8.2). Additionally, from $\pi \circ p_\gamma = \gamma$ we have $\mathrm{d}\pi \circ \mathrm{d}p_\gamma = \mathrm{d}\gamma$. Thus, we see that $\mathrm{d}p_\gamma(\lambda)$ is horizontal lift of $\mathrm{d}\gamma(\lambda)$ to the point $p_\gamma(\lambda)$,

$$\mathrm{d}p_\gamma(\lambda) = \mathrm{d}\pi|^{-1}_{H_{p_\gamma(\lambda)}}(\mathrm{d}\gamma(\lambda)) = -\Gamma_{\dot\gamma(\lambda)}(p_\gamma(\lambda)). \tag{8.5}$$

Finally, since $p_\gamma(\lambda) = \operatorname{trans}_{\gamma(\lambda\leftarrow 0)} p$, combining eqs. (8.4) and (8.5) we have the result. □

{TO DO: Show how Γ is a $\mathfrak{gl}(\mathscr{E})$-valued 1-form}

For a linear bundle (introduced in section 8.1.3) the composition in eq. (8.3) is just matrix multiplication, and the resulting linear differential equation can be solved explicitly.

Evaluating lemma 21 at $\lambda = 0$ yields the following useful result.

Corollary 2. *Let $\gamma : [0,1] \to \mathscr{M}$ be a path and let $p \in \mathscr{F}_{\gamma(0)}$.*

$$\left.\frac{\mathrm{d}}{\mathrm{d}\lambda} \operatorname*{trans}_{\gamma(\lambda\leftarrow 0)} p \right|_{\lambda=0} = -\Gamma_{\dot\gamma(0)}(p)$$

8.1.2. Covariant differentiation

We have seen that a choice of connection $H \subset \mathrm{T}\mathscr{F}$ determines which tangent vectors in the bulk of a bundle are horizontal. This in turn defines the coordinate-independent COVARAINT DERIVATIVE as the rate of change of a section with respect to the connection's horizontal.

To decompose vectors into horizontal and vertical components according to H, we employ the PROJECTION and REJECTION maps

$$\mathrm{proj}_{H_p} : \mathrm{T}_p\mathscr{F} \to H_p \quad \text{and} \quad \mathrm{rej}_{H_p} : \mathrm{T}_p\mathscr{F} \to \mathrm{V}_p\mathscr{F} \tag{8.6}$$

defined by $\mathrm{proj}_{H_p}\boldsymbol{u} + \mathrm{rej}_{H_p}\boldsymbol{u} = \boldsymbol{u} \in \mathrm{T}_p\mathscr{F}$ and idempotence.

Definition 39. *The* COVARIANT DERIVATIVE $\underset{\sim}{\nabla} f \in \Omega^1(\mathscr{M}, \mathrm{V}\mathscr{F})$ *of a section* $f \in \Gamma(\mathscr{F})$ *is defined by*

$$\underset{\sim}{\nabla} f = \mathrm{rej}_H \circ \mathrm{d}f. \tag{8.7}$$

Equation (8.7) is a vertical-valued 1-form, i.e., a linear map $\underset{\sim}{\nabla} f|_x : \mathrm{T}_x\mathscr{M} \to \mathrm{V}_{f(x)}\mathscr{F}$ defined at each $x \in \mathscr{M}$. Acting on a vector $\boldsymbol{u} \in \mathrm{T}_x\mathscr{M}$, this reads

$$\nabla_{\boldsymbol{u}} f \coloneqq \underset{\sim}{\nabla} f(\boldsymbol{u}) = \mathrm{rej}_{H_{f(x)}}\mathrm{d}f(\boldsymbol{u}) \in \mathrm{V}_{f(x)}\mathscr{F}.$$

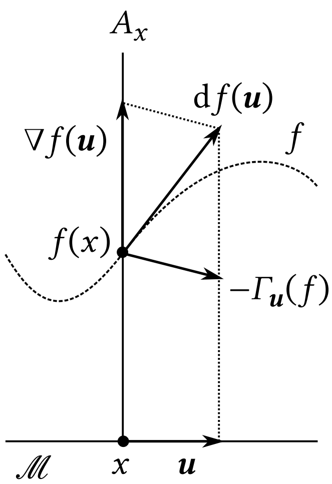


Figure 8.4.: Covariant derivative of f at $x \in \mathscr{M}$ along $\boldsymbol{u} \in \mathrm{T}_x\mathscr{M}$. The vector $-\Gamma_f(\boldsymbol{u}) = \mathrm{d}\pi|^{-1}_{H_{f(x)}}(\boldsymbol{u})$ indicates horizontal motion under the connection H, and $\nabla_{\boldsymbol{u}} f$ is the derivative relative to this horizontal.

This can be interpreted intuitively as follows. The true gradient vector $\mathrm{d}f(\boldsymbol{u}) \in \mathrm{T}_{f(x)}\mathscr{F}$ of the section f lies outside the fibre's tangent space $\mathrm{V}_{f(x)}\mathscr{F} \subseteq \mathrm{T}_{f(x)}\mathscr{F}$. However, we do not want to measure horizontal motion — just the *effective* vertical change of $f(x)$ within the fibre induced by moving x in the direction of $\boldsymbol{u}$. Thus, the covariant derivative $\nabla_{\boldsymbol{u}} f \in \mathrm{V}_{f(x)}\mathscr{F}$ is the vertical projection of $\mathrm{d}f(\boldsymbol{u})$ obtained by discarding its horizontal component, where 'horizontal' is of course specified by the connection (see fig. 8.4).

Lemma 22. *The covariant derivative as in definition 39 is equivalent to*

$$\nabla_{\boldsymbol{u}} f = \mathrm{d}f(\boldsymbol{u}) + \Gamma_{\boldsymbol{u}}(f),$$

where $\mathrm{d}f$ *is the push-forward of* $f \in \Gamma(\mathscr{F})$ *and* Γ *is the horizontal lift as in definition 37.*

Proof. By the defining property of the projection and rejection (8.6),

$$\mathrm{d}f = \mathrm{rej}_H \circ \mathrm{d}f + \mathrm{proj}_H \circ \mathrm{d}f$$

since $\mathrm{d}f : \mathrm{T}\mathscr{M} \to \mathrm{T}\mathscr{F}$ is linear. Therefore, rewriting definition 39,

$$\underset{\sim}{\nabla} f = \mathrm{rej}_H \circ \mathrm{d}f = \mathrm{d}f - \mathrm{proj}_H \circ \mathrm{d}f.$$

Using eq. (8.1), the projection operator at $p \in \mathscr{F}$ can be written as

$$\mathrm{proj}_{H_p} = \mathrm{d}\pi|_{H_p}^{-1} \circ \mathrm{d}\pi.$$

Finally, because f is a section, $\pi \circ f = \mathrm{id}$ and so $\mathrm{d}\pi \circ \mathrm{d}f = \mathrm{id}$ by the chain rule (lemma 19). Thus, acting on a base vector $\boldsymbol{u} \in \mathrm{T}_x\mathscr{M}$,

$$\begin{aligned}\nabla_{\boldsymbol{u}} f &= \mathrm{d}f(\boldsymbol{u}) - \mathrm{d}\pi|_{H_{f(x)}}^{-1} \circ \mathrm{d}\pi \circ \mathrm{d}f(\boldsymbol{u}) \\ &= \mathrm{d}f(\boldsymbol{u}) - \mathrm{d}\pi|_{H_{f(x)}}^{-1}(\boldsymbol{u}),\end{aligned}$$

which by definition 37 gives the result. □

Coordinate representation

At this point, we may introduce component forms of the above devices for a general fibre bundle. Let $\{x^\mu\}$ be local coordinates on $\mathscr{M}$ and $\{x^a\}$ local coordinates of the fibres. Let capital Latin indices $\{x^A\} = \{x^\mu, x^a\}$ run over all coordinates so that $p \in \mathscr{F}$ has coordinates $(p^A) = (x^\mu, x^a)$. Vertical motion fixes the base coordinates, but the fibre coordinates x^a are *not* required to be constant under horizontal motion.

Denote the associated coordinate basis of $\mathrm{T}\mathscr{F}$ by $(\boldsymbol{\partial}_A) = (\boldsymbol{\partial}_\mu, \boldsymbol{\partial}_a)$. Recall that $\Gamma(f) \in \Omega^1(\mathscr{M}, H)$ is a 1-form, and hence is linear in its tangent vector argument $\boldsymbol{u} \in \Gamma(\mathrm{T}\mathscr{M})$. Thus, we define the components

$$\Gamma_\mu := \Gamma_{\boldsymbol{\partial}_\mu}$$

so that $\Gamma_{\boldsymbol{u}}(f) = u^\mu \Gamma_\mu(f)$. Since $\Gamma_{\boldsymbol{u}}(f)|_x \in H_{f(x)}$ is a horizontal vector, we may also define the 2-component form ${\Gamma_\mu}^A$ by

$$\Gamma_\mu(f) = {\Gamma_\mu}^A(f)\,\boldsymbol{\partial}_A.$$

Note that horizontal vectors have both fibre *and* base components,

$${\Gamma_\mu}^A\,\boldsymbol{\partial}_A = {\Gamma_\mu}^\nu\,\boldsymbol{\partial}_\nu + {\Gamma_\mu}^a\,\boldsymbol{\partial}_a.$$

Indeed, the same applies to the push-forward $\mathrm{d}f$,

$$\mathrm{d}f = \mathrm{d}f^\mu\,\boldsymbol{\partial}_\mu + \mathrm{d}f^a\,\boldsymbol{\partial}_a$$

since $\mathrm{d}f$ is not vertical (refer back to fig. 8.4) — this is the problem the covariant derivative addresses. However, since $\nabla_\mu f \in \mathrm{V}\mathscr{F}$ as a whole *is* vertical, the base components $\mathit{\Gamma}_\mu{}^\nu$ and $\mathrm{d}f^\nu$ must cancel.

This is verified by noting that

$$\mathrm{d}\pi(\mathrm{d}f(\boldsymbol{u})) = \boldsymbol{u} \quad \text{and} \quad \mathrm{d}\pi(-\mathit{\Gamma}_{f(x)}(\boldsymbol{u})) \equiv \mathrm{d}\pi(\mathrm{d}\pi|^{-1}_{H_{f(x)}}(\boldsymbol{u})) = \boldsymbol{u} \tag{8.8}$$

are equal. In effect, $\mathrm{d}\pi$ projects onto components of the base, $\mathrm{d}\pi(X^A\boldsymbol{\partial}_A) = X^\nu\boldsymbol{\partial}_\nu$, and so eq. (8.8) implies $\mathrm{d}f^\nu(\boldsymbol{u}) = -u^\mu\mathit{\Gamma}_\mu{}^\nu$. Therefore, in components, the covariant derivative of a section is

$$\nabla_\mu f^a = \partial_\mu f^a + \mathit{\Gamma}_\mu{}^a(f), \tag{8.9}$$

with base components of $\mathrm{d}f(\boldsymbol{u})$ and $\mathit{\Gamma}_{\boldsymbol{u}}(f)$ suppressed.[67] Note that f need not be a vector section of a linear bundle — eq. (8.9) is general to smooth fibre bundles.

[67] In practice, one usually works with a (local) trivialisation in which $f : \mathscr{M} \to F$ is given as a field. Then, $\mathrm{d}f = \mathrm{d}f^a\,\boldsymbol{\partial}_a$ has no base components anyway, so we take $\mathit{\Gamma}_\mu(f) = \mathit{\Gamma}_\mu{}^a(f)\,\boldsymbol{\partial}_a$.

8.1.3. Structured connections

So far, we have treated connections in the setting of a general fibre bundle, in which fibres have the minimal structure of a smooth manifold. We now consider connections and their associated covariant derivatives on *vector* bundles $V \hookrightarrow \mathscr{V} \twoheadrightarrow \mathscr{M}$, with more or less additional structure.

On vector bundles

In general, the transport operator over a path is an invertible map between the start and end fibres. For a vector bundle, we require this to be a *linear* map. By lemma 21, this means the horizontal lift is also linear in its fibre argument, $\underset{\sim}{\mathit{\Gamma}}(\lambda^i X_i) = \lambda^i\underset{\sim}{\mathit{\Gamma}}(X_i)$, so we may regard $\mathit{\Gamma}_{\boldsymbol{u}}$ as a matrix and $\underset{\sim}{\mathit{\Gamma}}$ as a matrix-valued 1-form, acting on vectors $X \in \mathscr{V}$ by matrix multiplication, $\underset{\sim}{\mathit{\Gamma}}X \coloneqq \underset{\sim}{\mathit{\Gamma}}(X)$.

If $\{\boldsymbol{e}_a\}$ is a basis for some vector bundle $V \hookrightarrow \mathscr{V} \twoheadrightarrow \mathscr{M}$, then we may introduce the 3-component CONNECTION COEFFICIENTS,

$$\Gamma_\mu{}^a{}_b := \Gamma_\mu{}^a \boldsymbol{e}_b.$$

We may write expressions in both basis-free and component forms;

$$\Gamma_{\boldsymbol{u}} X = u^\mu \, \Gamma_\mu{}^a{}_b \, X^b \, \boldsymbol{e}_a.$$

Linearity also allows the covariant derivative to be expressed as the limit of a difference, similar to the usual analytical definition of the derivative of a real function.

Lemma 23. *If $\gamma : [0,1] \to \mathscr{M}$ is a path and $X \in \Gamma_\gamma(\mathscr{V})$ is a smooth vector section defined on γ, then*

$$\begin{aligned}\nabla_{\dot{\boldsymbol{\gamma}}(0)} X|_{\gamma(0)} &= \lim_{\varepsilon \to 0} \frac{X|_{\gamma(\varepsilon)} - \operatorname{trans}_{\gamma(\varepsilon \leftarrow 0)} X|_{\gamma(0)}}{\varepsilon} \\ &= \frac{\mathrm{d}}{\mathrm{d}\lambda} \left(X|_{\gamma(\lambda)} - \underset{\gamma(\lambda \leftarrow 0)}{\operatorname{trans}} X|_{\gamma(0)} \right) \Bigg|_{\lambda = 0}.\end{aligned}$$

Proof. Using corollary 2, the right-hand side is equal to

$$\mathrm{d}X(\dot{\boldsymbol{\gamma}}(0)) + \Gamma_{\dot{\boldsymbol{\gamma}}(0)} X,$$

which by lemma 22 is equal to $\nabla_{\dot{\boldsymbol{\gamma}}(0)} X|_{\gamma(0)}$. □

Metric compatibility

A linear connection on a vector bundle $V \hookrightarrow \mathscr{V} \twoheadrightarrow \mathscr{M}$ is called METRIC COMPATIBLE if for any vectors $X, Y \in \mathscr{V}$,

$$\langle \operatorname{trans} X, \operatorname{trans} Y \rangle = \langle X, Y \rangle$$

where the transport operators are over some common path.

Lemma 24. *A metric compatible connection satisfies*

$$\left\langle \underset{\sim}{\Gamma} X, Y \right\rangle = -\left\langle X, \underset{\sim}{\Gamma} Y \right\rangle \quad \text{or} \quad \Gamma_{\mu ab} = -\Gamma_{\mu ba}$$

where $\Gamma_{\mu ab} = \eta_{ac} \Gamma_\mu{}^c{}_b$.

Proof. Consider transport along a path $\gamma(\lambda \leftarrow 0)$, and abbreviate $T_\lambda \coloneqq \operatorname{trans}_{\gamma(\lambda\leftarrow 0)}$. Since $\langle T_\lambda X, T_\lambda Y\rangle = \langle X, Y\rangle$ is constant with respect to λ, its λ-derivative vanishes. But we also have

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}\lambda} \langle T_\lambda X, T_\lambda Y\rangle\bigg|_{\lambda=0} &= \left\langle \frac{\mathrm{d}}{\mathrm{d}\lambda} T_\lambda X\,|_{\lambda=0}, Y\right\rangle + \left\langle X, \frac{\mathrm{d}}{\mathrm{d}\lambda} T_\lambda Y\,|_{\lambda=0}\right\rangle \\ &= -\left\langle \mathit{\Gamma}_{\dot{\boldsymbol{\gamma}}(0)} X, Y\right\rangle - \left\langle X, \mathit{\Gamma}_{\dot{\boldsymbol{\gamma}}(0)} Y\right\rangle. \end{aligned}$$

Since γ is arbitrary, we have $\langle \mathit{\Gamma}_{\boldsymbol{u}} X, Y\rangle + \langle X, \mathit{\Gamma}_{\boldsymbol{u}} Y\rangle = 0$ for all $\boldsymbol{u} \in \mathrm{T}\mathscr{M}$.

Writing this in component form,

$$\eta_{ab}\, {\mathit{\Gamma}_\mu}^a{}_c\, X^c\, Y^b = -\eta_{ab}\, X^a\, {\mathit{\Gamma}_\mu}^b{}_c\, Y^c$$

which implies $\eta_{ab}\, {\mathit{\Gamma}_\mu}^a{}_c = -\eta_{ab}\, {\mathit{\Gamma}_\mu}^b{}_c$ since X and Y are arbitrary. □

Algebra-preserving connections

Vector bundles may be further equipped with an associative product, forming an algebra bundle. We require the product $\circledast : V_x \times V_x \to V_x$ to vary smoothy with $x \in \mathscr{M}$, so that $X \circledast Y \in \Gamma(\mathscr{V})$ is a smooth section whenever X and Y are. Requiring the transport operator to respect this product means enforcing

$$(\operatorname{trans} X) \circledast (\operatorname{trans} Y) = \operatorname{trans}(X \circledast Y), \tag{8.10}$$

which places further constraints on the connection.

Lemma 25. *Let $\circledast$ be a bilinear associative product on a vector bundle $\mathscr{V}$ which is respected by parallel transport as in eq. (8.10). Then,*

$$\underset{\sim}{\nabla}(X_1 \circledast \cdots \circledast X_k) = \underset{\sim}{\mathrm{d}}(X_1 \circledast \cdots \circledast X_k) + \sum_{i=1}^k X_1 \circledast \cdots \circledast \underset{\sim}{\mathit{\Gamma}} X_i \circledast \cdots \circledast X_k$$

for $X_i \in \Gamma(\mathscr{V})$.

Proof. Let $T_\lambda \coloneqq \operatorname{trans}_{\gamma(\lambda\leftarrow 0)}$ for some path γ. Using lemma 23, we have

$$\nabla_{\dot{\boldsymbol{\gamma}}(0)}(X_1 \circledast \cdots \circledast X_k) = \underset{\sim}{\mathrm{d}}(X_1 \circledast \cdots \circledast X_k)(\dot{\boldsymbol{\gamma}}(0)) - \frac{\mathrm{d}}{\mathrm{d}\lambda} \operatorname*{trans}_{\gamma(\lambda\leftarrow 0)}(X_1 \circledast \cdots \circledast X_k)\bigg|_{\lambda=0}.$$

Since $\circledast$ is respected by parallel transport, and by bilinearity and associativity, the rightmost term is

$$-\frac{\mathrm{d}}{\mathrm{d}\lambda}T_\lambda X_i \circledast \cdots \circledast T_\lambda X_i\bigg|_{\lambda=0} = -\sum_{i=1}^{k} X_1 \circledast \cdots \circledast \frac{\mathrm{d}}{\mathrm{d}\lambda}T_\lambda X_i\big|_{\lambda=0} \circledast \cdots \circledast X_k,$$

which by corollary 2 gives the result, after removing reverence to the arbitrary vector $\dot{\boldsymbol{\gamma}}(0)$. □

{TO DO: Show $\underset{\sim}{\nabla}$ is derivation.} A consequence of lemma 25 is that a linear connection on a vector bundle $\mathscr{V}$ induces a unique $\circledast$-respecting connection on the algebra bundle generated by $\circledast$. In particular, such a connection obeys the product rule:

$$\begin{aligned}\underset{\sim}{\nabla}(X \circledast Y) &= \underset{\sim}{\mathrm{d}}X \circledast Y + X \circledast \underset{\sim}{\mathrm{d}}Y + \underset{\sim}{\Gamma}X \circledast Y + X \circledast \underset{\sim}{\Gamma}Y \\ &= \underset{\sim}{\nabla}X \circledast Y + X \circledast \underset{\sim}{\nabla}Y\end{aligned}$$

For example, for the tensor bundle $\mathscr{V}^{\otimes}$, we obtain the product rule written in component form,

$$\nabla_\mu X^a Y^b = \partial_\mu(X^a Y^b) + \Gamma_\mu{}^a{}_c X^c Y^b + X^a \Gamma_\mu{}^b{}_c Y^c, \tag{8.11}$$

and if the connection is also compatible with a metric used to raise and lower indices, we derive the familiar formula for covariant differentiation of general type-(p,q) tensors,

$$\nabla_\mu T^{a_1\cdots a_p}{}_{b_1\cdots b_q} = \partial_\mu T^{a_1\cdots a_p}{}_{b_1\cdots b_q} + \sum_{i=1}^{p} \Gamma_\mu{}^{a_i}{}_c T^{a_1\cdots c\cdots a_p}{}_{b_1\cdots b_q} - \sum_{j=1}^{q} \Gamma_\mu{}^c{}_{b_j} T^{a_1\cdots a_p}{}_{b_1\cdots c\cdots b_q}. \tag{8.12}$$

On geometric algebra bundles

The covariant derivative assumes an elegant form when expressed as a bivector operator in the geometric algebra framework. A geometric algebra $\mathscr{G}(V,\eta)$ may be defined on a manifold by taking V to be the vector space of *sections* $\Gamma(\mathscr{V})$. We write $\mathscr{G}(\mathscr{V},\eta)$ as shorthand for $\mathscr{G}(\Gamma(\mathscr{V}),\eta)$ to indicate this construction.

Consider the covariant derivative of a vector $\boldsymbol{X} \in \mathscr{G}_1(\mathscr{V},\eta)$,

$$\nabla_\mu \boldsymbol{X} = (\partial_\mu X^a + \Gamma_\mu{}^a{}_b X^b)\boldsymbol{e}_a.$$

Rewrite the non-derivative term as

$$\begin{aligned}\Gamma_\mu{}^a{}_b\, \boldsymbol{e}_a X^b &= \Gamma_{\mu ab}\, \boldsymbol{e}^a(\boldsymbol{e}^b \cdot \boldsymbol{X}) \\ &= \frac{1}{2}\Gamma_{\mu ab}\left(\boldsymbol{e}^a(\boldsymbol{e}^b \cdot \boldsymbol{X}) - (\boldsymbol{e}^a \cdot \boldsymbol{X})\boldsymbol{e}^b\right)\end{aligned}$$

using the fact that $\Gamma_{\mu ab} = -\Gamma_{\mu ba}$ for a metric compatible connection, and that $\boldsymbol{e}^a \cdot \boldsymbol{X}$ is a scalar commuting with $\boldsymbol{e}^b$. Then, since for vectors the inner product is $\boldsymbol{X} \cdot \boldsymbol{Y} = \frac{1}{2}(\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{Y}\boldsymbol{X})$, this is b

$$\frac{1}{4}\Gamma_{\mu ab}\left(\boldsymbol{e}^a\boldsymbol{e}^b\boldsymbol{X} + \boldsymbol{e}^a\boldsymbol{X}\boldsymbol{e}^b - \boldsymbol{e}^b\boldsymbol{X}\boldsymbol{e}^a - \boldsymbol{X}\boldsymbol{e}^b\boldsymbol{e}^a\right) = \frac{1}{2}\Gamma_{\mu ab}\left(\boldsymbol{e}^a\boldsymbol{e}^b\boldsymbol{X} - \boldsymbol{X}\boldsymbol{e}^a\boldsymbol{e}^b\right).$$

In the right-hand side, the scalar parts from the products between $\boldsymbol{e}^a$ and $\boldsymbol{e}^b$ cancel, leaving a commutator product of the bivector $\boldsymbol{e}^a \wedge \boldsymbol{e}^b$ with $\boldsymbol{X}$,

$$\Gamma_{\mu ab}\,(\boldsymbol{e}^a \wedge \boldsymbol{e}^b) \times \boldsymbol{X} = \omega_\mu \times \boldsymbol{X},$$

where we define the CONNECTION BIVECTORS in the basis $\{\boldsymbol{e}_a\}$ by

$$\omega_\mu \coloneqq \Gamma_{\mu ab}\, \boldsymbol{e}^a \wedge \boldsymbol{e}^b.$$

Thus, we may write the covariant derivative of $\boldsymbol{X}$ as

$$\nabla_\mu \boldsymbol{X} = \partial_\mu \boldsymbol{X} + \omega_\mu \times \boldsymbol{X}. \tag{8.13}$$

The connection bivectors are especially useful for writing covariant expressions in the geometric algebra, because the form of eq. (8.13) is in fact general to all multivectors — in sharp contrast to, e.g., eqs. (8.11) and (8.12) in terms of connection coefficients. For sake of basis-independent notation, we may define the CONNECTION BIVECTOR 1-FORM $\underset{\sim}{\omega}$ by

$$\underset{\sim}{\omega}(\boldsymbol{u}) \equiv \omega_{\boldsymbol{u}} \coloneqq u^a \omega_a$$

expressed in any basis.

Lemma 26. *The covariant derivative of any multivector $A \in \mathscr{G}(\mathscr{V}, \eta)$ is*

$$\underset{\sim}{\nabla} A = \underset{\sim}{\mathrm{d}} A + \underset{\sim}{\omega} \times A.$$

Proof. The covariant derivative is a derivation if the connection respects the geometric product. Therefore, the covariant derivative of a product of k-many vectors is

$$\begin{aligned}\utilde{\nabla}(\boldsymbol{u}_1 \cdots \boldsymbol{u}_k) &= \sum_{i=1}^{k} \boldsymbol{u}_1 \cdots (\utilde{\mathrm{d}}\boldsymbol{u}_i + \utilde{\omega} \times \boldsymbol{u}_i) \cdots \boldsymbol{u}_k \\ &= \utilde{\mathrm{d}}(\boldsymbol{u}_1 \cdots \boldsymbol{u}_k) + \utilde{\omega} \times (\boldsymbol{u}_1 \cdots \boldsymbol{u}_k),\end{aligned}$$

using eq. (8.13) and the fact that commutation by a bivector is a derivation (lemma 9). Since all multivectors are linear combinations of products of vectors, the general result follows. □

A rapid alternative derivation of lemma 26 starts from the observation that parallel transport along a path may be written as

$$\operatorname*{trans}_{\gamma} A = RAR^{\dagger},$$

since any transformation continuously connected to the identity which preserves the geometric product belongs to the rotor group, Spin^{+} (see section 3.3). Any such rotor is of the form $R = e^{\sigma/2}$ for a bivector σ, so we have

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \operatorname*{trans}_{\gamma(\lambda \leftarrow 0)} A = \frac{1}{2} R(\sigma A - A\sigma)R^{\dagger}$$

where $\sigma = \sigma(\lambda)$ and hence R are functions of the path parameter. At $\lambda = 0$, the rotor is trivial, so by corollary 2 we find

$$\left.\frac{\mathrm{d}}{\mathrm{d}\lambda} \operatorname*{trans}_{\gamma(\lambda \leftarrow 0)} A\right|_{\lambda=0} = -\Gamma_{\dot{\boldsymbol{\gamma}}(0)}(A) = \sigma(0) \times A.$$

Thus, the horizontal lift is given by commutation with a specified bivector. Since this holds for arbitrary multivectors A, by lemma 22 we have the universally applicable formula for the covariant derivative of a multivector

$$\nabla_{\boldsymbol{u}} A = \partial_{\boldsymbol{u}} A + \omega_{\boldsymbol{u}} \times A$$

where $\omega_{\boldsymbol{u}}$ is the required bivector.

8.2. Covariant Algebraic Derivatives

The covariant derivative $\nabla_{\boldsymbol{u}}$ is analogous to the directional derivative $\partial_{\boldsymbol{u}}$ of section 6.1, but defined for sections on a manifold $F \in \Gamma(\mathscr{A})$ instead of fields on a vector space $F : V \to A$. In identical vein to definition 24, it is useful to define a 'total' derivative operator $\mathscr{D}$ which is independent of a direction $\boldsymbol{u} \in \mathrm{T}\mathscr{M}$. Like the algebraic derivative D of section 6.1, $\mathscr{D}$ is defined whenever an inclusion $\iota : \mathrm{T}^*\mathscr{M} \to \mathscr{A}$ is given (but usually left implicit) enabling tangent vectors to be multiplied by elements in the algebra.

Definition 40. *Let $A \hookrightarrow \mathscr{A} \twoheadrightarrow \mathscr{M}$ be an algebra bundle with product $\circledast$ and connection $\underset{\sim}{\nabla}$, equipped with an inclusion $\iota : \mathrm{T}^*\mathscr{M} \to \mathscr{A}$. The* COVARIANT ALGEBRAIC DERIVATIVE *of a section $F \in \Gamma(\mathscr{A})$ is*

$$\mathscr{D}F \coloneqq \iota(\boldsymbol{e}^a) \circledast \nabla_{\boldsymbol{e}_a} F$$

(summation on a) where $\{\boldsymbol{e}_a\} \subset \Gamma(\mathrm{T}\mathscr{M})$ and $\{\boldsymbol{e}^a\} \subset \Gamma(\mathrm{T}^*\mathscr{M})$ are dual bases of tangent sections.

8.2.1. Covariant vector derivative

The geometric algebra $\mathscr{G}(V, \eta)$ may be defined on a manifold by taking V to be the vector space of tangent sections $\Gamma(\mathrm{T}\mathscr{M})$. Write $\mathscr{G}(\mathscr{M}, \eta)$ as shorthand for $\mathscr{G}(\Gamma(\mathrm{T}\mathscr{M}), \eta)$ to indicate this construction.

On a geometric algebra bundle $\mathscr{G}(\mathscr{M}, \eta)$, definition 40 is the COVARIANT VECTOR DERIVATIVE $\boldsymbol{\mathscr{D}}$ and has similar properties to the vector derivative $\boldsymbol{\partial}$ on $\mathscr{G}(V, \eta)$. Given a basis $\{\boldsymbol{e}^a\}$ (of $\Gamma(\mathrm{T}^*\mathscr{M})$ and $\mathscr{G}_1(\mathscr{M}, \eta)$ alike), the covariant vector derivative is the operator

$$\boldsymbol{\mathscr{D}} = \boldsymbol{e}^a \nabla_a,$$

sharing the algebraic properties of a grade-1 vector.[68] In particular, for a k-vector section $F \in \mathscr{G}_k(\mathscr{M}, \eta)$, there is a decomposition into $(k \pm 1)$-grade parts,

$$\boldsymbol{\mathscr{D}}F = \boldsymbol{\mathscr{D}} \rfloor F + \boldsymbol{\mathscr{D}} \wedge F.$$

[68] Hence its bold symbol, like $\boldsymbol{\partial}$.

Chapter 9.

Curvature

Given a connection on a fibre bundle, values in the bulk may be parallel transported along a curve in the base manifold. If the curve is a closed loop, then values are not necessarily mapped back onto themselves. The action of parallel transport around a loop known as its HOLONOMY, and its deviation from the identity operator measures the connection's *curvature*.

Curvature may be restated as the obstruction to the *integrability* of the connection. Therefore, the curvature of a connection may be derived by finding the integrability condition of the parallel transport equations, which is most easily done via Frobenius' theorem [16, §6].

9.1. Integrability and Frobenius' Theorem

A vector field may be *integrated* by finding integral curves which are everywhere tangent to the vector field. This notion can be generalised to higher-dimensional analogues of vector fields which associate to each point a vector *subspace*, instead of merely a vector.

Definition 41. *A k-dimensional* TANGENT SUBBUNDLE $\mathscr{D} \subseteq \mathrm{T}\,\mathscr{M}$ *is a vector bundle* $\mathbb{R}^k \hookrightarrow \mathscr{D} \overset{\pi_S}{\twoheadrightarrow} \mathscr{M}$ *where each fibre* $\mathscr{D}|_x \cong \mathbb{R}^k$ *is a k-dimensional subspace of* $\mathrm{T}_x\,\mathscr{M}$.

Similarly, the notion of an integral curve to a vector field may be generalised to a tangent subbundles.

Definition 42. *A submanifold $\mathscr{I} \subseteq \mathscr{M}$ is called an* INTEGRAL MANIFOLD *of a tangent subbundle $\mathscr{D}$ if $\mathrm{T}_x \mathscr{I} \subseteq \mathscr{D}|_x$ for all $x \in \mathscr{I}$. The subbundle $\mathscr{D}$ is called* INTEGRABLE *if there exist integral manifolds through each point.*

For example, an integral curve of a vector field $\boldsymbol{u}$ through a point may be viewed as the 1-dimensional integral manifold of the 1-dimensional tangent subbundle described by $\boldsymbol{u}$. In higher dimensions, any embedded submanifold is a maximal integral manifold of its own tangent space, viewed as a tangent subbundle in the ambient space.

An integral manifold is MAXIMAL if $\mathrm{T}_x \mathscr{I} = \mathscr{D}|_x$, meaning the manifold dimension of $\mathscr{I}$ is the dimension of $\mathscr{D}$. Indeed, any tangent subbundle admits 1-dimensional integral curves, but is not maximally integrable in general. The existence of maximal integral surfaces requires a special property known as *involutivity*.

Definition 43. *A tangent subbundle $\mathscr{D}$ is* INVOLUTIVE *if $[\mathscr{D}, \mathscr{D}] \subseteq \mathscr{D}$. That is, if for any two sections $\boldsymbol{u}, \boldsymbol{v} \in \Gamma(\mathscr{D})$ in the subbundle, their Lie bracket $[\boldsymbol{u}, \boldsymbol{v}] \in \Gamma(\mathscr{D})$ also lies in the subbundle.*

The importance of involutivity as the integrability condition for a tangent subbundle is the content of Frobenius' theorem:

Theorem 5 (Frobenius'). *If $\mathscr{D}$ is a tangent subbundle, then*

$$\mathscr{D} \text{ is integrable} \quad \Longleftrightarrow \quad \mathscr{D} \text{ is involutive.}$$

Frobenius' theorem can be dualised into a statement involving exterior forms instead of vector subbundles, which can be more useful for calculation. This stems from the observation that a vector subspace $U \subseteq V$ may be represented by the subspace Ω of dual vectors with U contained in their kernels,

$$\Omega = \{\omega \in V^* \mid \omega(\boldsymbol{u}) = 0, \forall \boldsymbol{u} \in U\} \subseteq V^*.$$

The original subspace U is recovered as $U = \bigcap_{\omega \in \Omega} \ker \omega$.

Definition 44. *The dual representation I of a tangent subbundle $\mathscr{D}$ is the ideal*[69] *generated by the* 1*-form annihilators of $\mathscr{D}$,*

$$I = \left\{\!\!\left\{\omega \in \Omega^1(\mathscr{M}) \,\middle|\, \omega(\boldsymbol{u}) = 0, \forall \boldsymbol{u} \in \Gamma(\mathscr{D})\right\}\!\!\right\}.$$

69 Recall from definition 3 that an ideal (of forms) is closed under addition and satisfies $\alpha \wedge \omega \in I$ whenever $\omega \in I$, for *any* α.

The following lemma shows how the condition that $\mathscr{I}$ is an integral manifold translates between tangent subbundles and ideals.

Lemma 27. *Let $\mathscr{D}$ be a tangent subbundle and I is its associated ideal. Suppose $\mathscr{I}$ is a submanifold with the inclusion map $\iota : \mathscr{I} \to \mathscr{M}$. Then,*

$$\mathscr{D}|_p = \mathrm{T}_p\mathscr{I} \quad \iff \quad \mathscr{I} \textit{ is an integral manifold} \quad \iff \quad \iota^* I = 0.$$

Proof. The first equivalence is by definition, included for readability. For the second equivalence, assume $\mathscr{I}$ is an integral manifold. Then, if $\boldsymbol{u} \in \mathrm{T}\mathscr{I}$ then the inclusion $\mathrm{d}\iota(\boldsymbol{u}) \in \mathscr{D}$ lies in the tangent subbundle. Suppose $\omega \in I$ so that $\omega(\boldsymbol{v}) = 0$ for all $\boldsymbol{v} \in \mathscr{D}$. The restriction of ω to $\mathscr{I}$ via the pullback $\iota^*\omega$ is identically zero, because

$$(\iota^*\omega)(\boldsymbol{u}) \equiv \omega(\mathrm{d}\iota(\boldsymbol{u})) = 0.$$

Since $\boldsymbol{u}$ and $\omega \in I$ are arbitrary, we write $\iota^* I = 0$. □

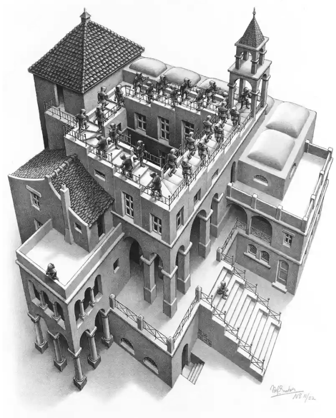


Figure 9.1.: *"Ascending and Descending" by M. C. Escher, 1960* — perhaps the most famous illustration of an inexact 2-form (the slope of the stairs) and its inconsistent 'integral' (the impossible staircase).

We can also translate the involutivity condition from tangent subbundles to ideals.

Theorem 6. *If $\mathscr{D} \subseteq \mathrm{T}\mathscr{M}$ is a tangent subbundle and $I \subseteq \Omega^1(\mathscr{M})$ is its associated ideal, then*

$$[\mathscr{D}, \mathscr{D}] \subseteq \mathscr{D} \quad \iff \quad \mathscr{D} \textit{ is involutive} \quad \iff \quad \mathrm{d}I \subseteq I.$$

Proof. The first equivalence is by definition, included for readability. For the second, note that the ideal I is generated by 1-forms ω which vanish on $\mathscr{D}$. That is, $\omega(\boldsymbol{u}) = 0$ for all $\boldsymbol{u} \in \Gamma(\mathscr{D})$, so if $\boldsymbol{u}, \boldsymbol{v} \in \Gamma(\mathscr{D})$ then

$$\begin{aligned} \mathrm{d}\omega(\boldsymbol{u}, \boldsymbol{v}) &= \boldsymbol{u}(\omega(\boldsymbol{v})) - \boldsymbol{v}(\omega(\boldsymbol{u})) - \omega([\boldsymbol{u}, \boldsymbol{v}]) \\ &= -\omega([\boldsymbol{u}, \boldsymbol{v}]), \end{aligned}$$

since $\omega(\boldsymbol{u}) = \omega(\boldsymbol{v}) = 0$. If $\mathscr{D}$ is involutive then $[\boldsymbol{u}, \boldsymbol{v}] \in \Gamma(\mathscr{D})$ and $\mathrm{d}\omega(\boldsymbol{u}, \boldsymbol{v}) = 0$. Thus, $\mathrm{d}\omega \in I$ if and only if $\mathscr{D}$ is involutive. □

Hence, by theorems 5 and 6, a tangent subbundle admits maximal integral surfaces if and only if its associated ideal I is closed under exterior differentiation, $\mathrm{d}I \subseteq I$.

Stokes' theorem 8 states that a differential form φ is integrable if it is exact (i.e., if $\varphi = \mathrm{d}\phi$). On a contractible domain, this is equivalent to φ being closed, by Poincaré's lemma. In the same vein, theorem 6 states that an exterior differential system is integrable over a contractible domain if and only if its associated ideal is closed.

9.1.1. Curvature as an obstruction to integrability

We may employ Frobenius' theorem to find the integrability condition for the connection on a vector bundle $V \hookrightarrow \mathscr{V} \twoheadrightarrow \mathscr{M}$. A linear Ehresmann connection H is integrable if there exist maximal integral manifolds $f \in \Gamma(\mathscr{F})$ which are everywhere horizontal, $\mathrm{T}_p f = H_p$. This means that $\nabla f = 0$ everywhere, that parallel transport is path-independent, and that loop holonomy is always trivial.

Elaborating the condition $\nabla f = 0$, we have

$$\nabla_{\boldsymbol{u}} X = \boldsymbol{u}(X) + \Gamma(\boldsymbol{u})X = 0 \quad \text{or} \quad \partial_\mu X^a = -\Gamma_\mu{}^a{}_b X^b \tag{9.1}$$

everywhere for all $\boldsymbol{u} \in \mathrm{T}\mathscr{M}$. These equations describe the tangent subbundle H. To express this, introduce coordinates $\{x^\mu\}$ of $\mathscr{M}$ and linear coordinates $\{x^a\}$ of V with respect to some basis. A point $X \in \mathscr{V}$ is a base point $\pi(X) \equiv (X^\mu) \in \mathscr{M}$ together with a fibre value $(X^a) \in V$, having total coordinates $X = (X^\mu, X^a)$. Similarly, a vector in $\mathrm{T}_X\mathscr{V}$ has components $\delta X = (\delta X^\mu, \delta X^a)$.

Such a vector $\delta X \in \mathrm{T}_X\mathscr{V}$ satisfies eq. (9.1) if $\delta X^a/\delta X^\mu = -\Gamma_\mu{}^a{}_b X^b$, and hence the Ehresmann connection may be expressed as

$$H_X = \operatorname{span}\Big\{(\delta X^\mu, -\Gamma_\mu{}^a{}_b X^b \delta X^\mu) \,\Big|\, (\delta X^\mu) \in \mathrm{T}_X\mathscr{M}\Big\} \tag{9.2}$$

for each $X \in \mathscr{V}$. Geometrically, this describes the change in vector components δX^a induced by a nudge in the base point δX^μ if X is constrained to move along H.

To employ Frobenius' theorem, we will find a dual representation of eq. (9.2) in terms of forms. Any $X \in H$ is of the form

$$X = \delta X^\mu (\boldsymbol{\partial}_\mu - \Gamma_\mu{}^a{}_b X^b \boldsymbol{\partial}_a).$$

If I is the ideal associated to H, then any 1-form $\underset{\sim}{\omega} \in I$ satisfies

$$\underset{\sim}{\omega}(X) = \delta X^\mu \left(\omega_\mu - \Gamma_\mu{}^a{}_b X^b \omega_a \right) = 0$$

where $\omega_A := \underset{\sim}{\omega}(\boldsymbol{\partial}_A)$, implying $\omega_\mu = \Gamma_\mu{}^a{}_b X^b \omega_a$ at X. Written in the coordinate dual basis $\{\mathrm{d}\underset{\sim}{X}^\mu, \mathrm{d}\underset{\sim}{X}^a\} \subset \mathrm{T}^* \mathscr{V}$,

$$\underset{\sim}{\omega} = \omega_a \left(\mathrm{d}\underset{\sim}{X}^a + \Gamma_\mu{}^a{}_b X^b \mathrm{d}\underset{\sim}{X}^\mu \right) \tag{9.3}$$

where ω_a are free scalar parameters. Here, we adopt the notation '$\underset{\sim}{\ }$' to label differential forms for clarity. Since eq. (9.3) is a general 1-form of the ideal I, we can see that I is generated by the 1-forms

$$\underset{\sim}{\Omega}^a = \mathrm{d}\underset{\sim}{X}^a + \underset{\sim}{\Gamma}^a{}_b X^b, \tag{9.4}$$

where we define the connection 1-forms $\underset{\sim}{\Gamma}^a{}_b := \Gamma_\mu{}^a{}_b \, \mathrm{d}\underset{\sim}{X}^\mu$.

The dual formulation of Frobenius' theorem (theorem 6) states that the tangent subbundle H is involutive if and only if the ideal I is closed. This means that $\mathrm{d}\underset{\sim}{\Omega}^a \in \mathrm{d}I$ for every generator, which is equivalent to the condition $\mathrm{d}\underset{\sim}{\Omega}^a = \underset{\sim}{\alpha}_a \wedge \underset{\sim}{\Omega}^a$ for arbitrary 'component 1-forms' $\underset{\sim}{\alpha}_a$. By direct calculation,

$$\begin{aligned} \mathrm{d}\underset{\sim}{\Omega}^a &= \mathrm{d}^2 \underset{\sim}{X}^a + \mathrm{d}\underset{\sim}{\Gamma}^a{}_b X^b - \underset{\sim}{\Gamma}^a{}_b \wedge \mathrm{d}\underset{\sim}{X}^b \\ &= (\mathrm{d}\underset{\sim}{\Gamma}^a{}_b + \underset{\sim}{\Gamma}^a{}_c \wedge \underset{\sim}{\Gamma}^c{}_b) X^b - \underset{\sim}{\Gamma}^a{}_b \wedge \underset{\sim}{\Omega}^a \end{aligned}$$

where we substitute eq. (9.4) on the second line. Therefore, $\mathrm{d}\underset{\sim}{\Omega}^a \in I$ if and only if the residual term, called the CONNECTION 2-FORM

$$\underset{\sim}{R}^a{}_b := \mathrm{d}\underset{\sim}{\Gamma}^a{}_b + \underset{\sim}{\Gamma}^a{}_c \wedge \underset{\sim}{\Gamma}^c{}_b, \tag{9.5}$$

vanishes. These $\underset{\sim}{R}^a{}_b$ measure the obstruction to integrability of the covariant derivative, and are identified as the primary object describing the connection's curvature.

9.2. Stokes' Theorem for the Curvature 2-form

Another way of showing that parallel transport is path-independent if and only if the curvature (9.5) vanishes is by relating the holonomy of a loop to the curvature across a surface bounded by the loop.

9.2.1. Path-ordered exponentiation

An initial value problem of the form

$$\frac{\mathrm{d}U(t)}{\mathrm{d}t} = A(t)U(t) \tag{9.6}$$

with $U(0)$ given has the solution

$$U(t) = e^{\int_0^t d\tau A(\tau)}U(0)$$

provided that $A(t)$ commutes with itself at all other times, $[A(t), A(s)] = 0$. If $A(t)$ is not necessarily commutative, then the solution may still be written formally in the following way.

By a first-order Taylor expansion, the value after an infinitesimal time-step dt step is

$$U(dt) = U(0) + \partial_t U(0)dt = (1 + A(0)dt)U(0) = e^{A(0)dt}U(0).$$

The value at a finite time t is then recovered by composing steps as above, forming the PATH-ORDERED EXPONENTIAL

$$U(t)U^{-1}(0) = \overleftarrow{\mathbb{P}}_\tau \exp\int_0^t d\tau A(\tau) := \lim_{dt\to 0}\prod_{t_i}^{t\leftarrow 0} e^{A(t_i)dt},$$

where the product $\prod_{t_i}^{t\leftarrow 0}$ is over values $t \geq t_i \geq 0$ in steps of dt where each exponential factor appears *right-to-left* in order of increasing t_i.

From the observation that $\partial_t(U(t)U^{-1}(t)) = 0$ we obtain the 'inverse' of the original differential equation,

$$\partial_t U(t)^{-1} = -U(t)^{-1}A(t), \tag{9.7}$$

which is identical to (9.6) only transposed and substituting $U(t)^T \mapsto U(t)^{-1}$ and $A(t)^T \mapsto -A(t)$. Hence, (9.7) has solution

$$\begin{aligned} U(t)^{-1} &= U(0)^{-1}\vec{\mathbb{P}}_\tau \exp \int_0^t d\tau(-A(\tau)) \\ &= U(0)^{-1}\overleftarrow{\mathbb{P}}_\tau \exp \int_t^0 d\tau A(\tau). \end{aligned}$$

Hence, the left-to-right ordered exponential $\vec{\mathbb{P}}$ exp is the same as a right-to-left $\overleftarrow{\mathbb{P}}$ exp if the endpoints $0 \leftrightarrow t$ are swapped and the integrand $d\tau \mapsto -d\tau$ flips sign.

The transport operator as a path-ordered exponential

The transport operator may be solved explicitly in terms of a path-ordered exponential. If $\gamma : \mathbb{R} \to \mathscr{M}$ is a curve, then a vector $X|_{\gamma(0)}$ is parallel-transported along γ if $\nabla_{\dot\gamma(s)} X^a|_{\gamma(s)} = 0$ for all s. This is equivalent to the partial differential system

$$\frac{\partial X^a}{\partial \dot\gamma^\mu} = -\Gamma_\mu{}^a{}_b X^a$$

restricted to the curve $\gamma(s)$. By the chain rule, multiplying by $\dot\gamma^\mu(s)$ produces an ordinary differential equation in s,

$$\frac{\mathrm{d}X^a}{\mathrm{d}s} = -\Gamma_\mu{}^a{}_b \dot\gamma^\mu(s) X^a = -\underset{\sim}{\Gamma}{}^a{}_b(\dot{\boldsymbol{\gamma}}(s))\, X^a \tag{9.8}$$

with both sides evaluated at $\gamma(s)$. Equation (9.8) can be integrated from the initial data $X|_{\gamma(0)} = X_0$ to give

$$X^a|_{\gamma(s)} = \underset{\gamma(s\leftarrow 0)}{\mathrm{trans}}{}^a{}_b\, X_0^b$$

where $\gamma(b \leftarrow a)$ is the path along γ from $\gamma(a)$ to $\gamma(b)$, and

$$\underset{\gamma(s\leftarrow 0)}{\mathrm{trans}} = \overleftarrow{\mathbb{P}} \exp \int_\gamma (-\underset{\sim}{\Gamma}) = \vec{\mathbb{P}} \exp \int_s^0 ds\, \underset{\sim}{\Gamma}(\dot{\boldsymbol{\gamma}}(s)). \tag{9.9}$$

Equation (9.9) is the solution to eq. (8.3) along γ with the initial data $\mathrm{trans}_{\gamma(0\leftarrow 0)} = \mathrm{id}$.

9.2.2. Surface-ordered exponentiation

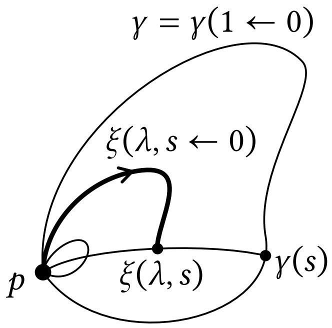


Figure 9.2.: The curve γ and the surface of homotopy ξ. The bold curve represents the portion of $h_\lambda \circ \gamma$ from parameter value 0 to s.

Theorem 7 (Stokes theorem for curvature 2-forms). *Let $\gamma : [0,1] \to \mathscr{M}$ be a contractable loop with start and end point p. Let h_λ be a contraction homotopy with $\lambda \in [0,1]$ so that $h_0(x) = p$ and $h_1(x) = x$. Define $\xi(\lambda, s) := h_\lambda(\gamma(s))$ as the surface swept out by γ under the contraction.*

Let $\underset{\sim}{\Gamma}$ be a connection 1-form and let $U(\lambda, s) := \mathrm{trans}_{\xi(\lambda, s\leftarrow 0)}$ be the group element resulting from parallel transport along the path $\xi(\lambda, s \leftarrow 0)$. Then,

$$\underset{\gamma}{\mathrm{trans}} = \overleftarrow{\mathbb{P}}_s \exp \int_\gamma (-\underset{\sim}{\Gamma})$$

$$= \overrightarrow{\mathbb{P}}_\lambda \exp \int_0^1 d\lambda \int_0^1 ds\, U^{-1}\, \underset{\sim}{R}(\partial_s \xi, \partial_\lambda \xi)\, U,$$

where $\underset{\sim}{R} = \mathrm{d}\underset{\sim}{\Gamma} + \underset{\sim}{\Gamma} \wedge \underset{\sim}{\Gamma}$ is the curvature 2-form. Note that $U \equiv U(\lambda, s)$ and $\xi \equiv \xi(\lambda, s)$.

{TO SELF: This awkwardly uses different path orderings. Bralić's is left-to-right.}

Proof. Define the abbreviations

$$\Gamma_\lambda := \underset{\sim}{\Gamma}(\partial_\lambda \xi) \qquad \text{and} \qquad \Gamma_s := \underset{\sim}{\Gamma}(\partial_s \xi),$$

noting that λ and s are *not* indices. In full component form, these would be written, e.g.,

$$(\Gamma_\lambda)^a{}_b\,|_{\xi(\lambda,s)} = \Gamma_\mu{}^a{}_b\,|_{\xi(\lambda,s)} \frac{\partial \xi^\mu(\lambda, s)}{\partial \lambda}.$$

From corollary 2, we have

$$\left.\frac{\partial U(\lambda, s)}{\partial s}\right|_{s=0} = \left.\frac{\mathrm{d}}{\mathrm{d}s} \underset{\xi(\lambda, s\leftarrow 0)}{\mathrm{trans}}\right|_{s=0} = -\underset{\sim}{\Gamma}(\partial_s \xi)$$

where $\xi \equiv \xi(\lambda, s)$, which implies

$$\partial_s U = -\Gamma_s U \qquad \text{and} \qquad \partial_s U^{-1} = U^{-1} \Gamma_s$$

where $U \equiv U(\lambda, s)$. From these two relations it follows easily that

$$\begin{aligned} &\partial_s\big(U^{-1}\partial_\lambda U\big) = U^{-1}(\Gamma_s\partial_\lambda U + \partial_\lambda\partial_s U) = -U^{-1}(\partial_\lambda \Gamma_s)\,U \\ \text{and}\quad &\partial_s(U^{-1}\Gamma_\lambda U) = U^{-1}(\Gamma_s\Gamma_\lambda + \partial_s\Gamma_\lambda - \Gamma_\lambda\Gamma_s)U. \end{aligned}$$

The sum of the two equations above is

$$\partial_s(U^{-1}(\partial_\lambda + \Gamma_\lambda)U) = U^{-1}(\partial_s\Gamma_\lambda + \Gamma_s\Gamma_\lambda - (s \leftrightarrow \lambda))\,U.$$

Note that $\partial_s\Gamma_\lambda = \partial_s(\Gamma_\mu(\partial_\lambda\xi)) = (\partial_s\Gamma_\mu)\partial_\mu\xi^\mu + \Gamma_\mu\,\partial_s\partial_\lambda\xi^\mu$ and similarly for $\partial_\lambda\Gamma_s$, so that mixed partial derivatives cancel, leaving

$$\partial_s\Gamma_\lambda - \partial_\lambda\Gamma_s = (\partial_s\underset{\sim}{\Gamma})(\partial_\lambda\xi) - (\partial_\lambda\underset{\sim}{\Gamma})(\partial_s\xi).$$

Putting this together, we have

$$\begin{aligned} \partial_s(U^{-1}(\partial_\lambda + \Gamma_\lambda)U) &= U^{-1}\big((\partial_s\underset{\sim}{\Gamma})(\partial_\lambda\xi) + \underset{\sim}{\Gamma}(\partial_s\xi)\underset{\sim}{\Gamma}(\partial_\lambda\xi) - (s \leftrightarrow \lambda)\big)\,U \\ &= U^{-1}(\mathrm{d}\underset{\sim}{\Gamma} + \underset{\sim}{\Gamma}\wedge\underset{\sim}{\Gamma})(\partial_s\xi, \partial_\lambda\xi)\,U \\ &= U^{-1}\underset{\sim}{R}(\partial_s\xi, \partial_\lambda\xi)\,U. \end{aligned} \tag{9.10}$$

Recall that U and U^{-1} are the group elements which parallel transport vectors along $\xi(\lambda, s \leftarrow 0)$ and back again, respectively. Also, note that $\underset{\sim}{R}$ is a $\mathfrak{gl}(\mathscr{V})$-valued 2-form, which acts to infinitesimally transform vectors in $\mathscr{V}$. With these in mind, it is clear that eq. (9.10) is an infinitesimal linear map from the fibre $\mathscr{V}_p$ to itself.[70] Thus, it is well-defined to integrate eq. (9.10) with respect to s, to obtain a finite linear transformation on $\mathscr{V}_p$.

[70] Imagine the right-hand side of eq. (9.10) acting on a vector X. First, X is transported by U from $\xi(\lambda, 0) = p$ to $\xi(\lambda, s)$, then transformed infinitesimally by $\underset{\sim}{R}$, and finally transported back to the fibre at p by U^{-1}.

Integrating the left-hand side of eq. (9.10) yields

$$\int_0^1 ds\,U^{-1}(\lambda, 1)(\partial_\lambda + \Gamma_\lambda)U(\lambda, 1) = U^{-1}(\lambda, 1)\partial_\lambda U(\lambda, 1) \tag{9.11}$$

since $\Gamma_\lambda = \underset{\sim}{\Gamma}(\partial_\lambda\xi(\lambda, s))$ vanishes at $s \in \{0, 1\}$ because $\xi(\lambda, 0) = \xi(\lambda, 1) = p$ is constant. Thus, integrating both sides yields

$$U^{-1}(\lambda, 1)\partial_\lambda U(\lambda, 1) = \int_0^1 ds\,U^{-1}\underset{\sim}{R}(\partial_s\xi, \partial_\lambda\xi)\,U.$$

This is an initial value problem of the form $\partial_\lambda U(\lambda, 1) = U(\lambda, 1)A(\lambda)$, whose solution at $\lambda = 1$ may be given as the path-ordered exponential

$$U(1,1) = U(1,0)\vec{\mathbb{P}} \exp \int_0^1 d\lambda \, A(\lambda)$$

where $A(\lambda)$ is the right-hand side of eq. (9.11). Since $U(1,1) = \mathrm{trans}_\gamma$ and $U(1,0) = \mathrm{id}$, this shows the right-hand side of the theorem. □

Corollary 3. *Parallel transport is path-independent if and only if curvature vanishes.*

Proof. If the curvature vanishes everywhere, then by theorem 7 the holonomy around any loop is trivial, implying the transport operator between two fixed points is path-independent.

Conversely, if parallel transport is path-independent, then the transport operator around any loop γ is the identity. By theorem 7, this implies that the total curvature on a surface bounded by γ is zero. But since the surface and loop are arbitrary, the curvature must vanish everywhere. □

Appendix A.

Integral Theorems

A.1. Stokes' theorem for exterior calculus

Theorem 8 (Stokes' theorem in $\mathbb{R}^n$). *If $R \subseteq \mathbb{R}^n$ is a compact k-dimensional hypersurface with boundary ∂R, then a smooth differential form $\omega \in \Omega^{k-1}(R)$ satisfies*

$$\int_R \mathrm{d}\omega = \int_{\partial R} \omega. \tag{A.1}$$

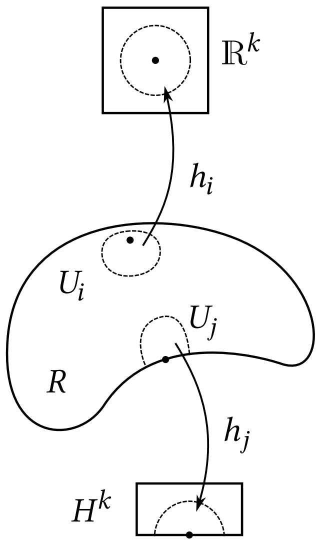


Figure A.1.: Neighbourhoods in R are diffeomorphic either to interior balls or boundary half-balls.

Proof. Since R is a k-dimensional region with boundary, every point $x \in R$ has a neighbourhood diffeomorphic to a neighbourhood of the origin in either $\mathbb{R}^k$ or $H^k := [0, \infty) \oplus \mathbb{R}^{k-1}$, depending on whether x is an interior point or a boundary point, respectively.

Let $\{U_i\}$ be a cover of R consisting of such neighbourhoods. Since R is compact, we may assume $\bigcup_{i=1}^N \{U_i\} = R$ to be a finite covering. Thus, we have finitely maps $h_i : U_i \to X$ where X is either $\mathbb{R}^k$ or the half-space H^k, where $U_i \cong h_i(U_i)$ are diffeomorphic (see fig. A.1).

Finally, let $\{\phi_i : R \to [0, 1]\}$ be a partition of unity subordinate to $\{U_i\}$, so that $\{x \in R \mid \phi_i(x) > 0\} \subseteq U_i$ and $\omega = \sum_{i=1}^N \phi_i \omega$. We need only prove the equality (A.1) for each $\omega_i := \phi_i \omega$, and the full result follows be linearity.

The form $h_i^* \omega_i \in \Omega^{k-1}(X)$ can be written with respect to canonical coordinates of X as

$$h_i^* \omega_i = \sum_{j=1}^{k} f_j (-1)^{j-1} \mathrm{d}x^{1\cdots\hat{j}\cdots k}$$

using the multi-index notation $\mathrm{d}x^{i_1\cdots i_k} \equiv \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$, where the hat denotes an omitted term. The factor of $(-1)^{j-1}$ gives the $(k-1)$-form the boundary orientation induced by the volume form $\mathrm{d}x^{1\cdots k}$ for convenience. Since pullbacks commute with d,

$$h^* \mathrm{d}\omega_i = \mathrm{d}\big(h_i^* \omega_i\big) = \sum_{j=1}^{k} \frac{\partial f_j}{\partial x^j} \mathrm{d}x^{1\cdots n}.$$

There are then two cases to consider.

- *Interior case.* If $h_i : U_i \to \mathbb{R}^k$, then the right-hand side of eq. (A.1) vanishes because ω_i is zero outside the neighbourhood $U_i \subset R$ which nowhere meets the boundary ∂R.

  $$\int_{\partial R} \omega_i = \int_{\partial U_i} \omega_i = \int_{\varnothing} \omega_i = 0$$

  The left-hand side evaluates to

  $$\begin{aligned}\int_R \mathrm{d}\omega_i = \int_X \mathrm{d}\big(h_i^* \omega_i\big) = \int_{\mathbb{R}^k} \sum_{j=1}^{k} \frac{\partial f_j}{\partial x^j} \mathrm{d}x^{1\cdots n} \\ = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{k} \sum_{j=1}^{k} \frac{\partial f_j}{\partial x^j} dx^1 \cdots dx^k \\ = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{k-1} \sum_{j=1}^{k} f_j \Big|_{x^j=-\infty}^{+\infty} (-1)^{j-1} dx^1 \cdots \widehat{dx^j} \cdots dx^k = 0,\end{aligned}$$

  which vanishes because $h_i^* \omega_i$, and hence the f_j, vanish outside the neighbourhood $h_i(U_i) \subset \mathbb{R}^k$.

- *Boundary case.* If $h_i : U_i \to H^k$, then the boundary $\partial U_i \subset \partial R$ is mapped onto the hyperplane $\partial H^k = \left\{(0, x^2, \dots, x^k) \,\middle|\, x^j \in \mathbb{R}\right\}$. Thus, $dx^1 = 0$ on this boundary, and the right-hand side of eq. (A.1) becomes

$$\int_{\partial R} \omega_i = \int_{\partial U_i} h_i^* \omega_i = - \int_{\mathbb{R}^{k-1}} f_1 dx^2 \cdots dx^k$$
$$= - \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{k-1} f_1(0, x^2, \dots, x^k) dx^2 \cdots dx^k.$$

The factor of -1 comes from the induced orientation of the boundary ∂H^k, which is outward-facing, so in the *negative* x^1 direction. For the left-hand side of eq. (A.1),

$$\int_R \mathrm{d}\omega_i = \int_{H^k} h_i^* \mathrm{d}\omega_i = \int_0^\infty \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \sum_{j=1}^{k} \frac{\partial f_j}{\partial x^j} dx^1 \cdots dx^k$$

All terms $\frac{\partial f_j}{\partial x^j} dx^j$ in the sum for $j > 1$ integrate to boundary terms $x_j \to \pm\infty$ where f_j vanishes. This leaves the single term from the integration of dx^1,

$$= - \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_1 \Big|_{x^1=0}^{\infty} dx^2 k \cdots k dx^{\cdot}$$

Thus, we have equality for all ω_i, so

$$\int_R \mathrm{d}\omega = \sum_{i=1}^{N} \int_R \mathrm{d}\omega_i = \sum_{i=1}^{N} \int_{\partial R} \omega_i = \int_{\partial R} \omega$$

by linearity. □

A.2. Fundamental theorem of geometric calculus

Theorem 9. *Let $f(\boldsymbol{x})$ be a multivector field. The vector derivative is*

$$\partial f(\boldsymbol{x}) = \lim_{|\mathscr{R}|\to 0} \frac{1}{|\mathscr{R}|\mathbb{I}} \oint_{\partial\mathscr{R}} \mathbf{d}S f,$$

where $\mathscr{R}$ is a volume containing $\boldsymbol{x}$ with boundary $\partial\mathscr{R}$ and volume $|\mathscr{R}| = \int_{\mathscr{R}} \mathbf{d}V$. The limit is taken as the volume $\mathscr{R}$ shrinks to the point $\boldsymbol{x}$.

Note that the integrand $\mathbf{d}Sf$ is the geometric product between the hypersurface element and the field.

Proof. It will suffice to prove the case where $\mathscr{R}$ is a box shape; arbitrary regions can be approximated via tessellation in the limit of vanishing voxel size.

Let $B_\varepsilon = \{x^i\boldsymbol{e}_i \mid x^i \in [-\varepsilon, +\varepsilon]\}$ denote the n-dimensional cube of diameter 2ε centred at the origin. If the surface ∂B_ε is oriented outward, then the face in the $+\boldsymbol{e}^k$ direction is orientated like the $(n-1)$-blade $\mathbb{I}\boldsymbol{e}^k = (-1)^{n-k}\boldsymbol{e}_1 \wedge \cdots \wedge \widehat{\boldsymbol{e}_k} \wedge \cdots \wedge \boldsymbol{e}_n$. Upon this face the infinitesimal surface element is

$$\mathbf{d}^{(k)}x = \mathbb{I}\boldsymbol{e}^k dx^1 \cdots \widehat{dx^k} \cdots dx^n,$$

while the opposing face has surface element $-\mathbf{d}^{(k)}x$.

Consider the integral of f over the surface ∂B_ε, split into a sum of n surface integrals over each pair of opposing faces. The kth pair are the surfaces $\{x^i\boldsymbol{e}_i \pm \varepsilon\boldsymbol{e}_k \mid x^i \in [-\varepsilon, +\varepsilon], i \neq k\}$ where i sums over axes *other* than k. Hence, we have

$$\oint_{\partial B_\varepsilon} \mathbf{d}Sf = \sum_{k=1}^{n} \int_{[-\varepsilon,+\varepsilon]^{n-1}} \mathbf{d}^{(k)}x \left(f(x^i\boldsymbol{e}_i + \varepsilon\boldsymbol{e}_k) - f(x^i\boldsymbol{e}_i - \varepsilon\boldsymbol{e}_k)\right), \quad (i \neq k).$$

By series expanding f in each x^i, and then in ε, obtain

$$f(x^i\boldsymbol{e}_i \pm \varepsilon\boldsymbol{e}_k) = f(\pm\varepsilon\boldsymbol{e}_k) + x^i\partial_{\boldsymbol{e}^i}(f(0) \pm \varepsilon\partial_{\boldsymbol{e}^k} f(0)).$$

Since $|x^i| \leq \varepsilon$, the last term is $\mathcal{O}(\varepsilon^2)$, and difference in the integrand is hence

$$\begin{aligned} f(x^i\boldsymbol{e}_i + \varepsilon\boldsymbol{e}_k) - f(x^i\boldsymbol{e}_i - \varepsilon\boldsymbol{e}_k) &= f(\varepsilon\boldsymbol{e}_k) - f(-\varepsilon\boldsymbol{e}_k) + \mathcal{O}(\varepsilon^2) \\ &= 2\varepsilon\partial_{\boldsymbol{e}^k} f(0) + \mathcal{O}(\varepsilon^2). \end{aligned}$$

Therefore, after pulling constants outside the integrals, we have

$$\oint\limits_{\partial B_\varepsilon} \mathbf{d}Sf \approx \sum_{k=1}^{n} 2\varepsilon\, \partial_{\boldsymbol{e}^k} f(0) \int\limits_{[-\varepsilon,+\varepsilon]^{n-1}} \mathbf{d}^{(k)}x$$

to order $\mathcal{O}(\varepsilon^2)$. The integrands each evaluate to the area $(2\varepsilon)^{n-1}$, giving

$$\oint\limits_{\partial B_\varepsilon} \mathbf{d}Sf \approx (2\varepsilon)^n \mathbb{I}\boldsymbol{e}^k \partial_{\boldsymbol{e}^k} f(0) = |B_\varepsilon|\mathbb{I}\, \boldsymbol{\partial} f(0),$$

to order $\mathcal{O}(\varepsilon^2)$, which becomes exact in the limit,

$$\boldsymbol{\partial} f(0) = \lim_{\varepsilon\to 0} \frac{1}{|B_\varepsilon|\mathbb{I}} \oint\limits_{\partial B_\varepsilon} \mathbf{d}Sf. \tag{A.2}$$

By translation, $f(\boldsymbol{x}) \mapsto f'(\boldsymbol{x}) = f(\boldsymbol{x} - \boldsymbol{u})$, we obtain the integral form of $\boldsymbol{\partial} f(\boldsymbol{u})$ evaluated at an arbitrary position $\boldsymbol{u}$. □

Theorem 10. *For an n-dimensional region $\mathscr{R}$ with boundary $\partial\mathscr{R}$, and a multivector field $f(\boldsymbol{x})$,*

$$\int\limits_{\mathscr{R}} \mathbf{d}V\, \boldsymbol{\partial} f = \oint\limits_{\partial\mathscr{R}} \mathbf{d}Sf,$$

where $\mathbf{d}V$ denotes an n-blade volume element, and $\mathbf{d}S$ an $(n-1)$-blade surface element, and where juxtoposition is the geometric product.

Proof. An arbitray volume $\mathscr{R}$ with boundary $\partial\mathscr{R}$ can be approximated as tessellated boxes of arbitrily small size.[71] Suppose $\mathscr{R}$ is approximated by a regular lattice of N boxes of radius ε. Consider the sum of $\boldsymbol{\partial} f$ over

[71] It is not neccesary that the surface area of the approximation converge to $|\partial\mathscr{R}|$.

the lattice points, weighted by volume. From eq. (A.2) this can be written in terms a sum of surface integrals,

$$\sum_{i=1}^{N} |B_i| \mathbb{I}\, \boldsymbol{\partial} f(\boldsymbol{x}_i) = \sum_{i=1}^{N} \oint_{\partial B_i} \mathbf{d}S\, f.$$

Note that interior faces of the boxes come in oppositely-oriented pairs, so that surface integrals over interior faces cancel. Therefore, the result is obtained in the continuous limit $N \to \infty$. □

{TO DO: Comment on how this generalizes Stokes' theorem.}

Bibliography

[1] Einstein, A. *On the electrodynamics of moving bodies.* Ann. Phys., 17(10):891–921 (**Jun. 1905**).

[2] Klein, F. *A comparative review of recent researches in geometry.* Bull. Amer. Math. Soc., 2(10):215–249 (**1893**). ISSN 0273-0979, 1088-9485. doi:10.1090/s0002-9904-1893-00147-x.

[3] Lipschitz, R. *Principes d'un calcul algébrique qui contient, comme espèces particulières, le calcul des quantités imaginaires et des quaternions : (extrait d'une lettre adressée à M. Hermite).* Gauthier-Villars (**1880**).

[4] Clifford, P. *Applications of grassmann's extensive algebra.* Am. J. Math., 1(4):350 (**1878**). ISSN 0002-9327. doi:10.2307/2369379.

[5] Chappell, J. M., Iqbal, A., Hartnett, J. G. and Abbott, D. *The vector algebra war: A historical perspective.* #IEEE_O_ACC#, 4:1997–2004 (**2016**). ISSN 2169-3536. doi:10.1109/access.2016.2538262.

[6] Silberstein, L. *LXXVI. quaternionic form of relativity.* The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 23(137):790–809 (**May 1912**). ISSN 1941-5982, 1941-5990. doi:10.1080/14786440508637276.

[7] De Leo, S. *Quaternions and special relativity.* J. Math. Phys., 37(6):2955–2968 (**Jun. 1996**). ISSN 0022-2488, 1089-7658. doi:10.1063/1.531548.

[8] Dirac, P. *Application of quaternions to Lorentz transformations.* Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences, 50:261–270 (**1944**). ISSN 00358975.

[9] Berry, T. and Visser, M. *Relativistic combination of non-collinear 3-velocities using quaternions.* Universe, 6(12):237 (**Dec. 2020**). ISSN 2218-1997. doi:10.3390/universe6120237.

[10] Berry, T. and Visser, M. *Lorentz boosts and Wigner rotations: Self-adjoint complexified quaternions.* Physics, 3(2):352–366 (**May 2021**). ISSN 2624-8174. doi:10.3390/physics3020024.

[11] Altmann, S. L. *Hamilton, rodrigues, and the quaternion scandal.* Math. Mag., 62(5):291–308 (**Dec. 1989**). ISSN 0025-570X, 1930-0980. doi:10.1080/0025570x.1989.11977459.

[12] Doran, C. and Lasenby, A. *Geometric Algebra for Physicists.* Cambridge University Press (**May 2003**). ISBN 9780521480222, 9780521715959, 9780511807497. doi:10.1017/cbo9780511807497.

[13] Wilson, J. and Visser, M. *Explicit Baker–Campbell–Hausdorff–Dynkin formula for spacetime via geometric algebra.* Int. J. Geom. Methods Mod. Phys., 18(14) (**Oct. 2021**). ISSN 0219-8878, 1793-6977. doi:10.1142/s0219887821502261.

[14] Gallian, J. A. *Student Solutions Manual.* Textbooks in mathematics. Chapman and Hall/CRC (**Jun. 2021**). ISBN 9781003182306. doi:10.1201/9781003182306.

[15] Kobayashi, S. and Nomizu, K. *Foundations of differential geometry*, vol. 1. New York, London (**1963**).

[16] Spivak, M. *A comprehensive introduction to differential geometry*, vol. 5. Publish or Perish, Incorporated (**1975**).

[17] Misner, C., Thorne, K. and Wheeler, J. *Gravitation.* Princeton University Press (**1973**). ISBN 0716703440.

[18] Hestenes, D. and Ziegler, R. *Projective geometry with Clifford algebra.* Acta Appl Math, 23(1):25–63 (**Apr. 1991**). ISSN 0167-8019, 1572-9036. doi:10.1007/bf00046919.

[19] Vince, J. *Geometric Algebra for Computer Graphics.* Springer London (**2008**). ISBN 9781846289965, 9781846289972. doi:10.1007/978-1-84628-997-2.

[20] Flanders, H. *Differential Forms with Applications to the Physical Sciences*, vol. 11. Elsevier (**1963**). ISBN 9780122596506. doi:10.1016/s0076-5392(08)x6021-7.

[21] Lee, J. M. *Introduction to smooth manifolds.* Grad. Texts Math. (**2012**). ISSN 0072-5285. doi:10.1007/978-1-4419-9982-5.

[22] Hestenes, D. *A unified language for mathematics and physics.* In *Clifford Algebras and Their Applications in Mathematical Physics*, pp. 1–23. Springer Netherlands (**1986**). ISBN 9789401086028, 9789400947283. doi:10.1007/978-94-009-4728-3_1.

[23] Hestenes, D. *Multivector calculus.* J. Math. Anal. Appl., 24(2):313–325 (**Nov. 1968**). ISSN 0022-247X. doi:10.1016/0022-247x(68)90033-4.

[24] Dorst, L. *The inner products of geometric algebra.* In Dorst, L., Doran, C. and Lasenby, J., eds., *Applications of Geometric Algebra in Computer Science and Engineering*, pp. 35–46. Birkhäuser Boston, Boston, MA (**2002**). ISBN 9781461266068, 9781461200895. doi:10.1007/978-1-4612-0089-5_2.

[25] Lasenby, A. N. *Geometric algebra as a unifying language for physics and engineering and its use in the study of gravity.* Adv. Appl. Clifford Algebras, 27(1):733–759 (**Jul. 2016**). ISSN 0188-7009, 1661-4909. doi:10.1007/s00006-016-0700-z.

[26] Doran, C., Hestenes, D., Sommen, F. and Van Acker, N. *Lie groups as spin groups.* J. Math. Phys., 34(8):3642–3669 (**Aug. 1993**). ISSN 0022-2488, 1089-7658. doi:10.1063/1.530050.

[27] Gallier, J. *The Cartan–Dieudonné Theorem*, chap. The Cartan–Dieudonné Theorem, pp. 231–280. Springer New York, New York, NY (**2011**). ISBN 9781441999603, 9781441999610. doi:10.1007/978-1-4419-9961-0_8.

[28] Hestenes, D. *Spacetime physics with geometric algebra.* Am. J. Phys., 71(7):691–714 (**Jul. 2003**). ISSN 0002-9505, 1943-2909. doi:10.1119/1.1571836.

[29] Gull, S., Lasenby, A. and Doran, C. *Imaginary numbers are not real—The geometric algebra of spacetime.* Found Phys, 23(9):1175–1201 (**Sep. 1993**). ISSN 0015-9018, 1572-9516. doi:10.1007/bf01883676.

[30] Dressel, J., Bliokh, K. Y. and Nori, F. *Spacetime algebra as a powerful tool for electromagnetism.* Phys. Rep., 589:1–71 (**Aug. 2015**). ISSN 0370-1573. doi:10.1016/j.physrep.2015.06.001.

[31] Bialynicki-Birula, I. and Bialynicka-Birula, Z. *The role of the Riemann–Silberstein vector in classical and quantum theories of electromagnetism.*

J. Phys. A: Math. Theor., 46(5):053001 (**Jan. 2013**). ISSN 1751-8113, 1751-8121. doi:10.1088/1751-8113/46/5/053001.

[32] Achilles, R. and Bonfiglioli, A. *The early proofs of the theorem of campbell, baker, hausdorff, and dynkin.* Arch. Hist. Exact Sci., 66(3):295–358 (**Apr. 2012**). ISSN 0003-9519, 1432-0657. doi:10.1007/s00407-012-0095-8.

[33] Coll, B. and Martínez, F. S. J. *Composition of Lorentz transformations in terms of their generators.* Gen. Relat. Gravit., 34(9):1345–1356 (**Sep. 2002**). ISSN 0001-7701, 1572-9532. doi:10.1023/a:1020018616308.

[34] Coll, B. and San José, F. *On the exponential of the 2-forms in relativity.* Gen Relat Gravit, 22(7):811–826 (**Jul. 1990**). ISSN 0001-7701, 1572-9532. doi:10.1007/bf00764159.

[35] Lasenby, J. *Guide to Geometric Algebra in Practice.* Springer London (**2011**). ISBN 9780857298102, 9780857298119. doi:10.1007/978-0-85729-811-9.

[36] Rodrigues, O. *Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ces déplacements considérés indépendamment des causes qui peuvent les produire.* J. Math. Pures Appl. (9), 5(1):380–440 (**1840**).

[37] Baylis, W. E. and Jones, G. *The Pauli algebra approach to special relativity.* J. Phys. A: Math. Gen., 22(1):1–15 (**Jan. 1989**). ISSN 0305-4470, 1361-6447. doi:10.1088/0305-4470/22/1/008.

[38] Ungar, A. A. *The relativistic velocity composition paradox and the Thomas rotation.* Found Phys, 19(11):1385–1396 (**Nov. 1989**). ISSN 0015-9018, 1572-9516. doi:10.1007/bf00732759.

[39] Mocanu, C. I. *On the relativistic velocity composition paradox and the Thomas rotation.* Found Phys Lett, 5(5):443–456 (**Oct. 1992**). ISSN 0894-9875, 1572-9524. doi:10.1007/bf00690425.

[40] O'Donnell, K. and Visser, M. *Elementary analysis of the special relativistic combination of velocities, Wigner rotation and Thomas precession.* Eur. J. Phys., 32(4):1033–1047 (**Jun. 2011**). ISSN 0143-0807, 1361-6404. doi:10.1088/0143-0807/32/4/016.