Geometric Algebra for Physicists

Lecture script

Preliminary version!

Contents

Chapter 0

Introduction

If a man, holding a belief which he was taught in childhood or persuaded of afterwards, keeps down and pushes away any doubts which arise about it in his mind, purposely avoids the reading of books and the company of men that call in question or discuss it, and regards as impious those questions which cannot easily be asked without disturbing it - the life of that man is one long sin against mankind.

William Clifford, The Ethics of Belief (1879)

0.1 Overview

Physics is the science of trying to describe reality with mathematics. Over time, specific mathematical constructs like vectors, matrices, tensors or differential forms have been established as the standard mathematical language for physics. However, these concepts are not unique. Physicists do not do mathematics for the sake of mathematics, but to model reality - and the most important part of this is to find the best mathematical tools for the specific phenomenon we are trying to describe.

Sadly, it is not commonplace in physics to try to improve upon the existing mathematical tooling. Once some mathematical concept like the cross product is introduced, physicists generally don't try to improve it, but uncritically use it for all eternity - even if it is already long obsolete for mathematicians. This has led us into a difficult situation - the mathematical tooling we are being taught in our university classes is a messy, incoherent patchwork that grew over multiple centuries. The various inconsistencies that have crept into it over time have a very real effect on the everyday work of physicists - they inhibit intuitive understanding of physical concepts, slow down our calculations and lead to misunderstandings.

Geometric algebra (GA) is an attempt at clearing up some parts of this chaos. In particular, geometric algebra replaces, improves and unifies

- vector algebra,
- large parts of matrix algebra,
- antisymmetric tensor algebra,
- exterior algebra and differential forms,
- multivariate, vector and exterior calculus,

- the tetrad formalism in differential geometry, and
- spinor algebra

into a simple, elegant and natural framework. Many concepts from physics become a lot clearer once they are reformulated in GA. In particular, GA puts a focus on geometric intuition for what is going on behind the maths.

The core concept of geometric algebra is the so-called **geometric product**. It is both associative and invertible. These properties make practical calculations with it very simple. To make it work out, we need to define so-called **k-vectors** and **multivectors** - a generalization of the concept of a vector.

0.2 History of geometric algebra

The inventors of geometric algebra in the modern sense are Hermann Grassmann (1809 - 1877) and William Clifford (1845 - 1879).

Grassmann introduced the **wedge product** and the **exterior algebra** in 1844. At that time, however, the general concept of a vector still was in its infancy, and Grassmann's exterior algebra was regarded as incomprehensible by contemporary physicists and mathematicians. Also, it had several important issues - the wedge product was not invertible, which hindered practical computations. Grassmann's work was largely ignored during his lifetime.

The situation changed, however, when Clifford combined the wedge product with the inner product to form the so-called **geometric product**. Based on Grassman's exterior algebra, he introduced **geometric algebra** in his 1878 work "Applications of Grassmann's Extensive Algebra". The algebra he described was very elegant and much easier to work with than exterior algebra. Unfortunately, he died from tuberculosis several months later, so his ideas had little chance to spread. His work was only remembered by some mathematicians under the name "Clifford algebra".

Meanwhile, physicists were investigating electromagnetism. This is when the concept of vectors and vector fields became mainstream in physics. However, Josiah Gibbs and Oliver Heaviside, two chief investigators of electromagnetism at that time, did not know of the work of Clifford and instead invented the makeshift ad-hoc construct of the "cross product" in the 1880s. For physicists, it certainly represented an advance at that time, but it should never have become a permanent solution. Unfortunately, it did - Heaviside's formulation of the Maxwell equations is identical to the modern-day one being taught to undergraduate physics students.

The cross product quickly took its roots in physics in the early 20th century, and Clifford's geometric algebra was largely forgotten until the advent of quantum mechanics. While investigating electron spin, Wolfgang Pauli introduced the Pauli matrices in 1927, and Paul Dirac the Dirac matrices in 1928. However, they were regarded as something purely quantum-mechanical at that time - what Pauli and Dirac did not realize is that the algebra formed by their matrices is nothing but the geometric algebra, and that their matrices have a geometric interpretation.

Later, in the 1960s, theoretical physicist David Hestenes (1933 -) rediscovered Clifford's work and made the connection to the work of Pauli and Dirac. He realized that geometric algebra was a much more convenient, elegant and natural tool for performing spatial computations than the matrix-vector and tensor algebra prevailing at his time, and went on a crusade to convince other physicists to use it. He extended geometric algebra to the theory of special relativity, devising **spacetime algebra (STA)**.

However, geometric algebra did not gain traction until the 2000s and 2010s, when Chris Doran and Anthony Lasenby published the comprehensive work **Geometric Algebra for Physicists (2002)**[1]. In the following years, geometric algebra began to spread among physicists and computer scientists due to its many advantages in practical calculations.

[1] This book is a great resource on how to apply geometric algebra to various parts of physics. Among us GA connaisseurs, it is jokingly called the "GA bible" or the "Gabel".

0.2.1 "Unified language for physics" claims

Some proponents have claimed that geometric algebra is so versatile that it can be used as a unified mathematical language for the entirety of physics. For instance, there are GA reformulations of the entirety of Lie group theory - even of the groups that are not directly elements of a (real) geometric algebra. Perhaps the most vocal proponents of this claim are Doran and Lasenby - in their book, they try to use as few non-GA mathematical constructs as possible. Critics of this claim have argued that the usefulness of geometric algebra is limited to specific branches of mathematics, and that it should not be used everywhere regardless of its specific benefits.

0.3 This lecture

In this lecture, we put a focus on practical computations and geometric intuition instead of mathematical formalism. It is explicitly **not** a mathematics lecture, but a physics lecture. We roughly follow the path of Doran, Lasenby: Geometric Algebra for physicists - we first introduct the concept of multivectors and the geometric product, and then successively treat the 2D, 3D, and 1+3D geometric algebras and their applications to everyday physics. In contrast to Doran and Lasenby, we put an emphasis on showing how to directly translate between GA and conventional maths.

We are not going to concern ourselves with whether geometric algebra is a unified language for physics. Instead, we are just going to switch between conventional math and geometric algebra depending on whichever is most convenient for the problem at hand.

The conventions we use mostly come from either Doran and Lasenby or sudgylacmoe's videos on geometric algebra, or have been devised by ourselves to achieve maximum clarity. Most notably, we use

$$\tau = 2\pi. \tag{1}$$

Using Tau has become somewhat commonplace in geometric algebra, for instance in sudgylacmoe's videos. This convention makes rotations much clearer to think about - one full turn is equal to τ radians, one half of a turn is equal to $\tau/2$, a quarter-turn is equal to $\tau/4$, and so on.

We are going to use colored boxes for various purposes:

> This is a warning.

> This is a box for information that is only tangentially related to the topic at hand - hints at more advanced topics, trivia, fun facts, etc

> **Important equation**
>
> $$\partial F = j \tag{2}$$

0.3.1 Literature

For further study, we recommend:

- **Doran, Lasenby: Geometric Algebra for Physicists (2002).** Very comprehensive work on geometric algebra. Most notably, its treatment of rotors and spinors differs from ours.
- **sudgylacmoe's video series on geometric algebra** (available on YouTube). 3blue1brown-styled introduction to various fields of geometric algebra. Very good primer.

- **Hestenes: Spacetime Algebra (1966).** The original book that reintroduced geometric algebra to physics. It covers the material of the second half of this script, although it uses some conventions that have become antiquated.

Chapter 1

Basics of geometric algebra

Chapter summary

- k-vectors are a generalization of the concept of scalars and vectors.
- The exterior product can be used to construct k-vectors from vectors.
- The interior product "tears down" k-vectors.
- The geometric product is the combination between the interior and the exterior product.
- Multivectors are sums of k-vectors of different grades.
- The geometric product is associative and invertible, but not commutative.
- The geometric algebra of the 2D plane Cl(2) is the simplest example of a geometric algebra.
- We can use the 2D geometric algebra to reformulate complex numbers and 2D rotations.

1.1 k-vectors and the exterior product

Much of geometric algebra revolves around finding a **vector product** that is both **associative** and **invertible**. In this chapter, we start by examining the **exterior product** (also known as the **outer product** or **wedge product**), and the **interior product** (also known as the **dot product** or **inner product**). We are going to show why neither of them fulfills these requirements.

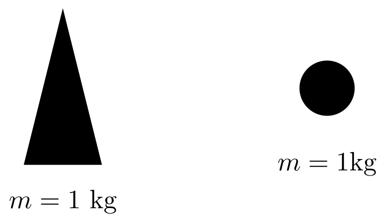
Then, we are going to combine the interior and exterior product into the **geometric product** and show that it does exactly what we want - the geometric product will be both associative and invertible.

1.1.1 Scalars and vectors

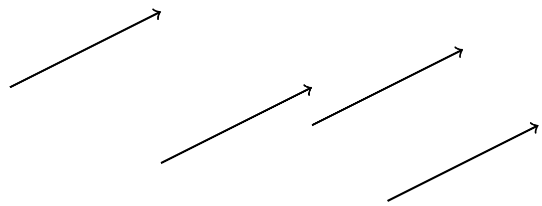
We start by introducing **k-vectors**, a core concept in GA. You already know two types of k-vectors:

- **scalars** (or 0-vectors): A quantity without dimension or orientation.
- **vectors** (or 1-vectors): A length with an orientation.

First of all, let's talk about scalars. Scalars represent simple quantities like mass or temperature which have no concept of direction or orientation. There's something important to note about scalars: If two scalars are "located" at two different positions in space - for instance the masses two distinct objects or two temperature measurements at different locations - we consider them equal iff their numerical values are equal, even though they are not at the same location. For instance, these two masses are considered equal:



Vectors, on the other hand, are a length with an orientation, like velocity, acceleration, or momentum. We consider two vectors equal to each other iff they have the same length and orientation. All of these vectors are considered equal to each other:



Mathematically, we describe vectors by first choosing a set of **basis vectors**, for instance in three dimensions:

$$e_1, \quad e_2, \quad e_3. \tag{1.1}$$

A specific vector is formed by taking a linear combination of them. For instance, the vectors in the above figure are given by

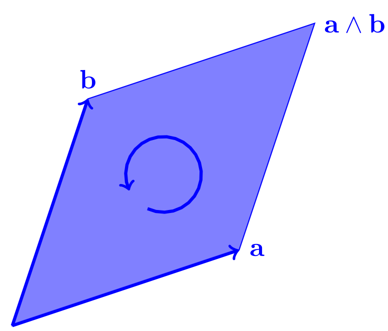
$$\mathbf{v} = 2e_1 + e_2. \tag{1.2}$$

1.1.2 Bivectors

The next grade of k-vectors we are going to examine is 2:

- **bivectors** (or 2-vectors): An area with an orientation.

To illustrate what that means, we are going to introduce the **wedge product** - also called **outer product** or **exterior product**: Let $\mathbf{a}, \mathbf{b}$ be two distinct vectors. They define a parallelogram:

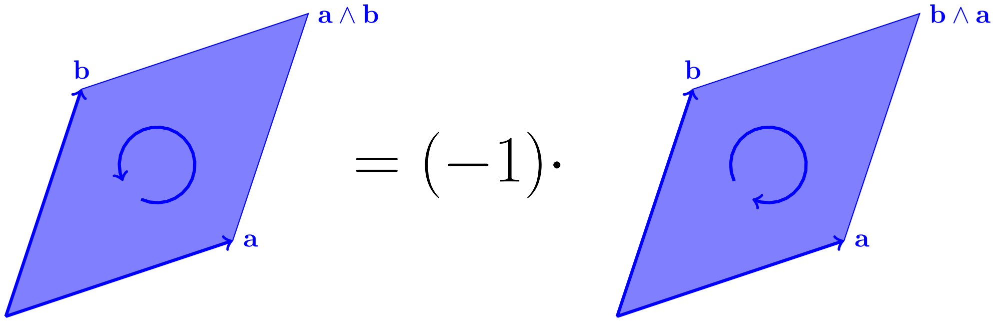


This parallelogram has an area and an orientation. Thus, we call it a bivector B. Its area can be calculated with the standard formula for parallelogram areas, while the orientation is given by the combined orientations of $\mathbf{a}$ and $\mathbf{b}$. Note that the order matters - we drew a "swirl" inside the bivector that marks the rotation from the first vector $\mathbf{a}$ to the second vector $\mathbf{b}$.

Mathematically, this bivector is given by the wedge product $\wedge$ between the two vectors:

$$B = \mathbf{a} \wedge \mathbf{b} \tag{1.3}$$

The wedge product is **anticommutative for vectors**, $\mathbf{a} \wedge \mathbf{b} = -\mathbf{b} \wedge \mathbf{a}$. The bivector given by going from $\mathbf{a}$ to $\mathbf{b}$ is exactly -1 times the bivector given by going from $\mathbf{b}$ to $\mathbf{a}$:

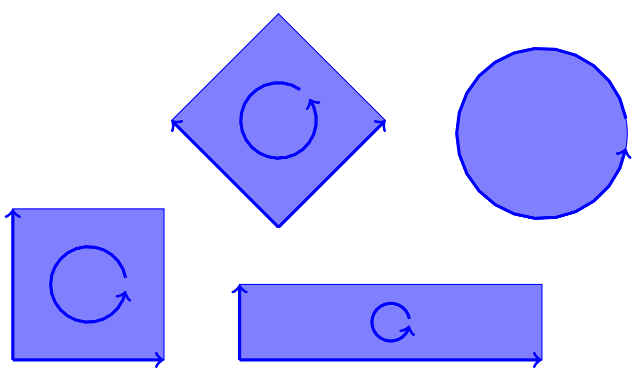


Note that the direction of the swirl in the second bivector was reversed. As a consequence, this means that for any vector,

$$\mathbf{a} \wedge \mathbf{a} = 0. \tag{1.4}$$

This makes intuitive sense - two times the same vector span no area between them.

Two bivectors are equal iff they have the same orientation and area, regardless of shape or position. This means that, for instance, the following bivectors are all equal to each other:



So how do we mathematically determine if two bivectors $B = \mathbf{a} \wedge \mathbf{b}$ and $C = \mathbf{c} \wedge \mathbf{d}$ are equal to each other? We make use of the associativity and distributivity of the wedge product. For instance, for:

$$\mathbf{a} = 2e_1 \tag{1.5}$$
$$\mathbf{b} = e_1 + 2e_2 \tag{1.6}$$
$$\mathbf{c} = 3e_1 - 4e_2 \tag{1.7}$$
$$\mathbf{d} = e_1 + e_3, \tag{1.8}$$

we insert the respective basis decompositions of the vectors into the wedge product:

$$B = \mathbf{a} \wedge \mathbf{b} \tag{1.9}$$
$$= 2e_1 \wedge (e_1 + 2e_2) \tag{1.10}$$
$$= 2e_1 \wedge e_1 + 4e_1 \wedge e_2 \tag{1.11}$$
$$= 4e_1 \wedge e_2, \tag{1.12}$$

and

$$
\begin{aligned}
C &= \mathbf{c} \wedge \mathbf{d} && (1.13)\\
&= (3e_1 + 4e_2) \wedge (e_1 + e_3) && (1.14)\\
&= 3e_1 \wedge e_1 + 3e_1 \wedge e_3 + 4e_2 \wedge e_1 + 4e_2 \wedge e_3 && (1.15)\\
&= -4e_1 \wedge e_2 + 4e_2 \wedge e_3 - 3e_3 \wedge e_1. && (1.16)
\end{aligned}
$$

We see that B and C clearly are not equal to each other. The form we brought them into right now is somewhat analogous to the basis decomposition for a vector - similarly to how we decomposed a vector **a** into a linear combination of $e_1, e_2, ...$, we decomposed the bivectors B and C into linear combinations of $e_1 \wedge e_2, e_2 \wedge e_3, ...$ to compare them.

Conventionally, we order the indices of our basis k-vectors in a cyclic way $(1 \to 2 \to 3 \to 1 \cdots)$.

1.1.3 Trivectors and more

The next type of k-vector on the line are:

- **trivectors** (or 3-vectors): A volume with an orientation.

We can form trivectors by taking the wedge product between three trivectors. For instance, the trivector

$$T = e_1 \wedge e_2 \wedge e_3 \tag{1.17}$$

is the parallelepiped spanned by the vectors e_1, e_2 and e_3.

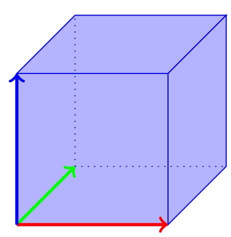


Figure 1.1: The trivector $e_1 \wedge e_2 \wedge e_3$.

As before, trivectors with the same volume and orientation but with different shapes are considered equal. In three dimensions, it is somewhat hard to see how different trivectors could have different orientations - after all, there is only one possible orientation for volumes in three dimensions. Consider, however, for a second, a four-dimensional Euclidean space with the basis vectors e_1, e_2, e_3, e_4. Then, e.g. the trivectors

$$e_1 \wedge e_2 \wedge e_3 \tag{1.18}$$

and

$$e_2 \wedge e_3 \wedge e_4 \tag{1.19}$$

would have the same volume, but differing orientations.

Again, these vectors need to be linearly independent. If they aren't, they don't span a volume, so the resulting trivector is zero. For instance, if we tried to take the wedge product between e_1, e_2 and e_2:

$$e_1 \wedge e_2 \wedge e_2 = 0. \tag{1.20}$$

We can also form trivectors by taking an already existing bivector and wedging another vector on top of it:

$$U = B \wedge e_3 = (4e_1 \wedge e_2) \wedge e_3 \tag{1.21}$$

The wedge product is **associative**, so we can drop the parentheses:

$$U = 4\; e_1 \wedge e_2 \wedge e_3 \tag{1.22}$$

In four-dimensional space, we will also get to know **tetravectors**, which represent four-dimensional hypervolumes. We can construct tetravectors either by wedging four vectors,

$$H = e_1 \wedge e_2 \wedge e_3 \wedge e_4, \tag{1.23}$$

a vector and a trivector,

$$H = (e_1 \wedge e_2 \wedge e_3) \wedge e_4 \tag{1.24}$$

or two bivectors:

$$H = (e_1 \wedge e_2) \wedge (e_3 \wedge e_4). \tag{1.25}$$

There is something important to take note of here: When we go to higher grades, the wedge product is **not always anticommutative**.

We could go to higher and higher grades of k-vectors, but in practice, tetravectors are going to be the highest grade of k-vectors we need.

1.1.4 Invertibility?

We have seen how we can build up bivectors, trivectors and tetravectors using the wedge product. But does it fulfill the conditions we set for a vector product? It is obviously associative.

The exterior product between a scalar λ and some other k-vector X is just the scalar multiplication:

$$\lambda \wedge X = \lambda X \tag{1.26}$$

In this special case, the exterior product is invertible. However, in general, it is not invertible at all for a simple reason:

We saw that the exterior product can be used to "span up" k-vectors from other k-vectors of lower grade. For instance, the wedge product between a vector and a bivector resulted in a trivector. For this to work out, the vector should not lie inside the bivector - if it does, the parallelipiped will be completely flat, and the trivector equal to zero. Only vectors with an orthogonal part w.r.t. the bivector produce non-zero trivectors.

This means that if we have a bivector B and a vector $\mathbf{v}$ with a part orthogonal and a part parallel to B,

$$\mathbf{v} = \mathbf{v}_\perp + \mathbf{v}_\parallel \tag{1.27}$$

the wedge product $B \wedge \mathbf{v}$ will only contain information about the orthogonal part:

$$B \wedge \mathbf{v} = B \wedge \mathbf{v}_\perp + B \wedge \mathbf{v}_\parallel \tag{1.28}$$

$$= B \wedge \mathbf{v}_\perp \tag{1.29}$$

This means that the information about the parallel part $\mathbf{v}_\parallel$ is lost. In other words, given the final trivector $T = B \wedge \mathbf{v}$ and the bivector B, there is no way we could reconstruct the original vector $\mathbf{v}$. Hence, the wedge product is not invertible.

1.2 The interior product

Next, we are going to talk about the **interior product** - also known as the **inner product** or the **dot product**. We already know the simplest case of the interior product - namely, the interior product between two vectors:

$$\mathbf{v} \cdot \mathbf{w} \tag{1.30}$$

The interior product is **commutative for vectors**. It measures the amount of overlap between $\mathbf{v}$ and $\mathbf{w}$. If $\mathbf{v}$ and $\mathbf{w}$ are parallel, we have $\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}||\mathbf{w}|$. If $\mathbf{v}$ and $\mathbf{w}$ are orthogonal, however, the interior product between them is zero.

Let's compare this with the exterior product between two vectors, $\mathbf{v} \wedge \mathbf{w}$. If the two vectors are orthogonal, the bivector has the area $|\mathbf{v}||\mathbf{w}|$, and if the vectors are parallel, the exterior product between them is zero. This is roughly the opposite of what the exterior product does.

1.2.1 Interior product between k-vectors and vectors

In conventional mathematics, the interior product is only defined for two vectors. It measures the amount of overlap between them. This concept extends neatly to k-vectors, though! For instance, we can take the interior product between a bivector $B = e_1 \wedge e_2$ and a vector e_1:

$$B \cdot e_1 = (e_1 \wedge e_2) \cdot e_1. \tag{1.31}$$

To evaluate this product, we first reorder the geometric product such that the vector standing next to the dot product is equal to the vector on the other side of the dot product:

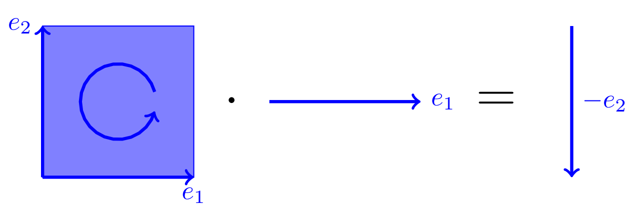
$$(e_1 \wedge e_2) \cdot e_1 = -(e_2 \wedge e_1) \cdot e_1. \tag{1.32}$$

If the two vectors next to this an interior product are **equal**, we can rewrite the parentheses to:

$$-(e_2 \wedge e_1) \cdot e_1 = -e_2(e_1 \cdot e_1) = -e_2 \tag{1.33}$$

This scheme only works if all of the vectors in the wedge product are orthogonal to each other. This is autmoatically the case if the k-vectors are in basis form, i.e. $e_i \wedge \cdot e_l$. If they aren't, we need to write out the wedge product into multiple terms like in (1.16).

The result of the interior product between a bivector and a vector is another vector. Intuitively, we can say that by dotting a vector onto the bivector, we "took away one grade" from the bivector and thus made it a vector. The resulting vector is what "remains" after "taking away" one grade from the bivector:



If the bivector and the vector are orthogonal to each other, for instance

$$(e_1 \wedge e_2) \cdot e_3 \tag{1.34}$$

the interior product between them is zero.

The same goes for trivectors and vectors. For instance, let $T = e_1 \wedge e_2 \wedge e_3$. If we want to take the interior product between T and e.g. e_2,

$$T \cdot e_2 = (e_1 \wedge e_2 \wedge e_3) \cdot e_2 \tag{1.35}$$

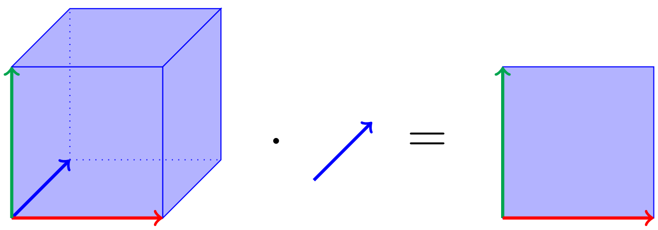
we first reorder the trivector such that the two e_2 stand next to each other:

$$(e_1 \wedge e_2 \wedge e_3) \cdot e_2 = -(e_1 \wedge e_3 \wedge e_2) \cdot e_2, \tag{1.36}$$

Now that the vectors left and right of the interior product are equal, we can rewrite the parentheses to:

$$-(e_1 \wedge e_3 \wedge e_2) \cdot e_2 = -(e_1 \wedge e_3)(e_2 \cdot e_2) \tag{1.37}$$
$$= -e_1 \wedge e_3 \tag{1.38}$$
$$= e_3 \wedge e_1. \tag{1.39}$$

The geometric picture is the same as previously - the grade of the trivector is lowered by one:



Again, the remaining bivector is the bivector orthogonal to e_2.

There is also a more elegant way to define the interior product between k-vectors - the Hodge adjoint of the exterior product. However, this definition requires too much abstract mathematics, so we won't treat it in this lecture.

1.2.2 Properties of the interior product

The interior product is not commutative in general. For instance, for a bivector and a vector, it is anticommutative:

$$(e_1 \wedge e_2) \cdot e_1 = -(e_2 \wedge e_1) \cdot e_1 = -e_2(e_1 \cdot e_1) = -e_2 \tag{1.40}$$
$$e_1 \cdot (e_1 \wedge e_2) = (e_1 \cdot e_1)e_2 = e_2 \tag{1.41}$$

In general, the interior product between a k-vector and a vector is commutative if k is odd, and anticommutative if k is even.

Just like the exterior product, the interior product is not invertible - but for a different reason: The exterior product throws the parallel part of the vector away, while the interior product throws the orthogonal part away. For instance, given a bivector $B = e_1 \wedge e_2$ and an interior product $B \cdot \mathbf{v} = e_1$, we can tell that the component of $\mathbf{v}$ parallel to B is e_2, but there is no way to tell if it also had an e_3 component - remember that

$$(e_1 \wedge e_2) \cdot e_3 = 0 \tag{1.42}$$

In the general case $\mathbf{v} = \mathbf{v}_\parallel + \mathbf{v}_\perp$, we can write that

$$B \cdot \mathbf{v} = B \cdot \mathbf{v}_\parallel. \tag{1.43}$$

Therefore, there is an inherent information loss in the interior product. It is not invertible.

In fact, the interior product is not even associative - consider for instance the two expressions

$$T \cdot (B \cdot \mathbf{w}) \tag{1.44}$$
$$(T \cdot B) \cdot \mathbf{w} \tag{1.45}$$

The first expression evaluates to a bivector, while the second expression evaluates to a a scalar. Therefore, it is important to always write down parentheses if we are handling multiple interior products.

1.3 Multivectors and the geometric product

In our quest for an invertible product, we have examined the interior and exterior product. We found that:

- The **exterior product** is used to construct k-vectors.
  - The exterior product between a k-vector and a vector is a $k+1$-vector.
  - The exterior product throws away the parallel parts.
- The **interior product** is used to tear down k-vectors.
  - The interior product between a k-vector and a vector is a $k-1$-vector.
  - The interior product throws away the orthogonal parts.

None of them are invertible. What if we combined them, though? Then we ought to have both pieces of information (the orthogonal and parallel part) we need for invertibility. For two vectors $\mathbf{a}$ and $\mathbf{b}$, we define the **geometric product**:

Geometric product between vectors

$$\mathbf{ab} = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b} \tag{1.46}$$

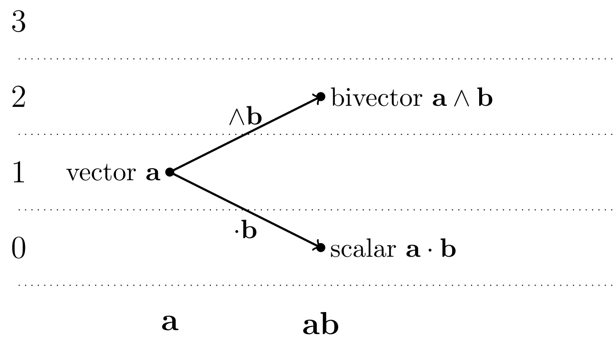
This is the sum of a scalar and a bivector. You are probably wondering how that is possible - how can we sum together two completely different things? The answer is that we just do it without evaluating the sum. Think of complex numbers - real and imaginary numbers are two completely different things, and yet we just sum them together. In geometric algebra, we will be constantly summing k-vectors of different grades together. A sum of k-vectors of different grades is called a **multivector**[1].

As we shall soon see, this geometric product is both **associative** and **invertible**. For now, let us calculate a few examples. For instance, let $\mathbf{a} = e_1$ and $\mathbf{b} = e_1 + e_2$. Then the geometric product between them is

$$\mathbf{ab} = e_1 \cdot (e_1 + e_2) + e_1 \wedge (e_1 + e_2) \tag{1.47}$$
$$= 1 + e_1 \wedge e_2, \tag{1.48}$$

a sum of a scalar and a bivector. We can use grade diagrams like this one to depict the geometric product:

[1] The mathematician's way of saying this is that the space of multivectors is the direct sum of all the spaces of k-vectors from $k = 0$ to $k = \infty$.



We can apply the geometric product to more than two vectors. For instance,

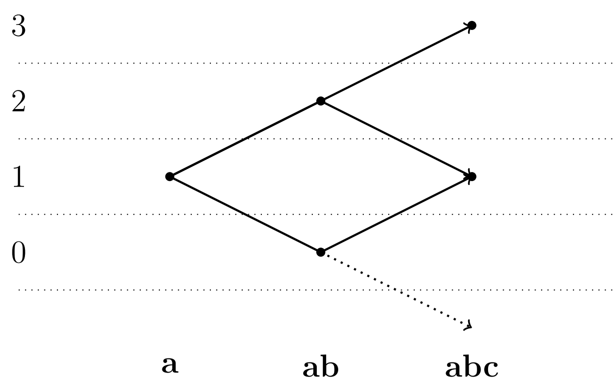
$$\mathbf{abc} \tag{1.49}$$

is a perfectly valid expression - the geometric product is associative, so we can drop the parentheses. It decomposes into four terms:

$$\mathbf{abc} = (\mathbf{a} \cdot \mathbf{b})\mathbf{c} + (\mathbf{a} \wedge \mathbf{b})\mathbf{c} \tag{1.50}$$

$$= (\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c} + (\mathbf{a} \wedge \mathbf{b}) \cdot \mathbf{c} + (\mathbf{a} \wedge \mathbf{b}) \cdot \mathbf{c} + (\mathbf{a} \wedge \mathbf{b}) \wedge \mathbf{c}. \tag{1.51}$$

First things first - in the first term, we are trying to take the interior product between a scalar (grade-0) and a vector (grade-1). This is supposed to lower the grade of the scalar by one - but there are no grade -1 elements. Thus , this interior product is equal to zero. The second and third terms are interior products between a bivector and a vector. They result in vectors. The fourth term is the wedge product between two vectors - this results in a trivector.



Therefore, the expression **abc** results in the sum of a vector and a trivector.

1.3.1 Grade projection

We can extract the real and imaginary parts of a complex numbers with the functions $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$. Similarly, given some arbitrary multivector M, we can extract the k-th grade component by using the so-called **grade projection**

$$\langle M \rangle_k . \tag{1.52}$$

For instance, the bivector part of M would be written $\langle M \rangle_2$. Applied to the previous example, it yields

$$\langle \mathbf{ab} \rangle_2 = \langle 1 + e_1 \wedge e_2 \rangle_2 = e_1 \wedge e_2. \tag{1.53}$$

Its scalar part is:

$$\langle \mathbf{ab} \rangle_0 = \langle 1 + e_1 \wedge e_2 \rangle_2 = 1. \tag{1.54}$$

We will also often drop the subscript to denote the scalar part, i.e.

$$\langle M \rangle := \langle M \rangle_0 \,. \tag{1.55}$$

1.3.2 Anticommutation relations

When the two vectors are orthogonal to each other, the interior product between them is zero. In such cases, the geometric product reduces to the scalar product:

$$e_i e_j = e_i \wedge e_j \qquad \text{for all } i \neq j. \tag{1.56}$$

Therefore, in such cases, the geometric product is anticommutative:

$$e_i e_j = -e_j e_i \qquad \text{for all } i \neq j. \tag{1.57}$$

Therefore, basis k-vectors like $e_1 \wedge e_2 \wedge e_3$ can be equally written with the geometric product, $e_1e_2e_3$. We can abbreviate this even further - to avoid having to write e's all the time in products like $e_1e_2e_3$, we just write e_{123}.

If the two vector operands of the geometric product are equal to each other, the exterior product is zero and only the interior product remains. In such a case, the geometric product is equal to the magnitude squared of the vector:

$$\mathbf{vv} = \mathbf{v}^2 = \mathbf{v} \cdot \mathbf{v} + \mathbf{v} \wedge \mathbf{v} \tag{1.58}$$
$$= \mathbf{v} \cdot \mathbf{v} \tag{1.59}$$
$$= |\mathbf{v}|^2 \tag{1.60}$$

Using the anticommutator brackets

$$\{A, B\} = AB + BA, \tag{1.61}$$

We can also denote these relations more succinctly as

$$\{e_i, e_j\} = 2\delta^{ij} \tag{1.62}$$

where δ^{ij} is the metric of our space - in the Euclidean case, just the Kronecker delta.

In fact, this relation is where mathematicians start when investigating Clifford algebras - they postulate this anticommutation relation and then derive all the other properties of geometric algebras from it. Obviously, we are not going to do that - we don't want mathematical rigour, but physical intuition.

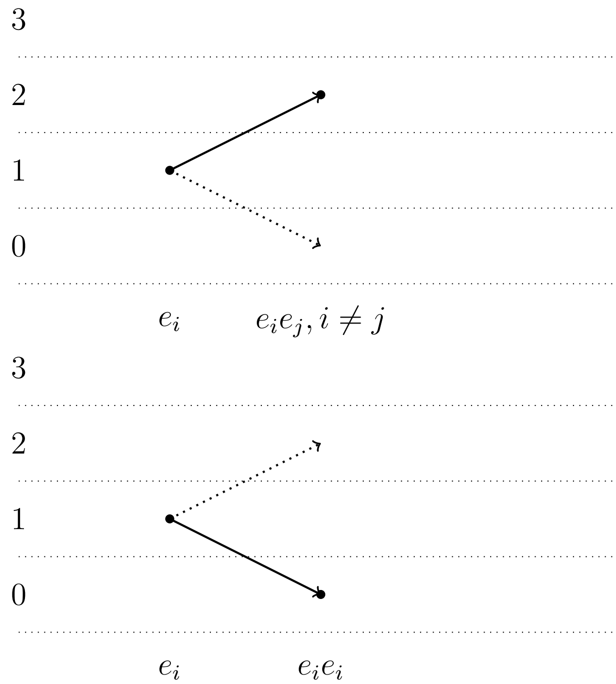


Figure 1.2: The grade diagram for the anticommutation relation $\{e_i, e_j\} = 2\delta^{ij}$.

1.3.3 Higher grades

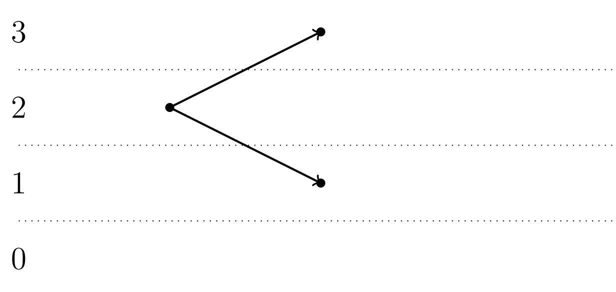
The most prominent advantage of the geometric product is that it is associative. So far, we've just been taking geometric products between vectors. However, we can just as well take the geometric product between arbitrary k-vectors. For instance, let B be the bivector

$$B = e_1 \wedge e_2 = e_1 e_2 = e_{12}. \tag{1.63}$$

Then, for another vector $\mathbf{v} = e_1 + e_3$, we can write down the geometric product

$$B\mathbf{v}. \tag{1.64}$$

Let's draw a grade diagram:



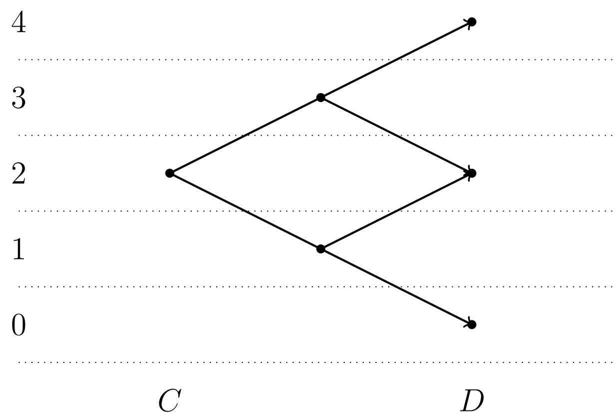
Before even evaluating it, we can tell that this product will consist of a vector and trivector part.

Let's test this:

$$B\mathbf{v} = e_{12}(e_1 + e_3) \tag{1.65}$$
$$= e_{12}e_1 + e_{12}e_3 \tag{1.66}$$
$$= e_{121} + e_{123} \tag{1.67}$$
$$= -e_{112} + e_{123} \tag{1.68}$$
$$= -e_2 + e_{123}, \tag{1.69}$$

a vector and a trivector. In general, we can calculate every geometric product between arbitrary multivectors by resolving the multivectors into products of vectors and then writing them out.

Let's try another example: Let $C = e_{12} + e_{23}$ and $D = e_{34} + e_{13}$ be two bivectors. Multiplying by a bivector is the same as consecutively multiplying by its constituent vectors. The grade diagram for the geometric product CD is:



We see that the result will be the sum of a scalar, a bivector and a tetravector. In fact,

$$CD = (e_{12} + e_{23})(e_{34} + e_{12}) \tag{1.70}$$
$$= e_{12}e_{34} + e_{23}e_{34} + e_{12}e_{12} + e_{23}e_{12} \tag{1.71}$$
$$= e_{1234} + e_{24} - e_{1221} - e_{2321} \tag{1.72}$$
$$= e_{1234} + e_{24} - e_{11} + e_{2231} \tag{1.73}$$
$$= e_{1234} + e_{24} - 1 + e_{31} \tag{1.74}$$
$$= -1 + e_{24} + e_{31} + e_{1234}. \tag{1.75}$$

We see that this is the sum of a scalar

$$\langle CD \rangle = -1, \tag{1.76}$$

a bivector

$$\langle CD \rangle_2 = e_{24} + e_{31}, \tag{1.77}$$

and a tetravector

$$\langle CD \rangle_4 = e_{1234}. \tag{1.78}$$

1.3.4 Invertibility

We call a vector $\mathbf{n}$ a **unit vector** if it squares to 1 with the geometric product:

$$\mathbf{n}^2 = \mathbf{nn} = \mathbf{n} \cdot \mathbf{n} = 1. \tag{1.79}$$

We are all used to reading this with respect to the scalar product - but the geometric product is more powerful than that. Most importantly, it is associative. If we multiply $\mathbf{n}$ onto an arbitrary multivector M,

$$M\mathbf{n}, \tag{1.80}$$

we can reverse this operation by multiplying $\mathbf{n}$ on top of it again:

$$(M\mathbf{n})\mathbf{n} = M\mathbf{n}\mathbf{n} = M(\mathbf{n}\mathbf{n}) = M. \tag{1.81}$$

Therefore, we can define the **vector inverse** for unit vectors:

$$\mathbf{n}^{-1} = \mathbf{n}, \tag{1.82}$$

or for arbitrary vectors $\mathbf{a}$:

$$\mathbf{a}^{-1} = \frac{\mathbf{a}}{\mathbf{a}^2}. \tag{1.83}$$

We can see that this is an inverse of $\mathbf{a}$ by calculating:

$$\mathbf{a}\mathbf{a}^{-1} = \mathbf{a}\frac{\mathbf{a}}{\mathbf{a}^2} = \frac{\mathbf{a}^2}{\mathbf{a}^2} = 1 \tag{1.84}$$

This is the true power of the geometric product - we can form the inverse and hence **divide by vectors**.

We can even form inverses for higher-grade k-vectors. For instance, the bivector e_{12} squares to:

$$(e_{12})^2 = e_{12}e_{12} = e_{1212} = -e_{1221} = -e_{11} = -1. \tag{1.85}$$

Therefore, its inverse is $(e_{12})^{-1} = -e_{12}$.

1.3.5 Interior and exterior product revisited

Previously, we defined the interior and exterior product for k-vectors of arbitrary grade and then combined them to form the geometric product. However, we can also do this the other way around. For a k-vector X and an l-vector Y with $k \geq l$, we define:

- The interior product $X \cdot Y$ is the lowest possible grade achievable by the geometric product XY:

$$X \cdot Y = \langle XY \rangle_{k-l} \tag{1.86}$$

- The exterior product $X \wedge Y$ is the highest possible grade achievable by the geometric product XY:

$$X \wedge Y = \langle XY \rangle_{k+l} \tag{1.87}$$

For instance, for two vectors $\mathbf{a}, \mathbf{b}$, we'd define:

$$\mathbf{a} \cdot \mathbf{b} = \langle \mathbf{ab} \rangle \tag{1.88}$$
$$\mathbf{a} \wedge \mathbf{b} = \langle \mathbf{ab} \rangle_2 . \tag{1.89}$$

For a bivector B and a vector $\mathbf{a}$, we'd define:

$$B \cdot \mathbf{a} = \langle B\mathbf{a} \rangle_1 \tag{1.90}$$
$$B \wedge \mathbf{a} = \langle B\mathbf{a} \rangle_3 . \tag{1.91}$$

These definitions provide us with a much easier way to evaluate the interior product: Instead of shuffling around the order of the basis k-vectors and then rearranging the parentheses, we can just calculate the geometric product between them and then pick out the $k-l$ grade.

There is also another way: We look at whether the grade operations are commutative or anticommutative. We already know that for two vectors $\mathbf{a}$ and $\mathbf{b}$, the inner product commutes:

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}, \tag{1.92}$$

and the wedge product anticommutes:

$$\mathbf{a} \wedge \mathbf{b} = -\mathbf{b} \wedge \mathbf{a}. \tag{1.93}$$

Therefore, we can tell that for two vectors:

$$\mathbf{a} \cdot \mathbf{b} = \frac{1}{2}(\mathbf{ab} + \mathbf{ba}) \tag{1.94}$$

$$\mathbf{a} \wedge \mathbf{b} = \frac{1}{2}(\mathbf{ab} - \mathbf{ba}). \tag{1.95}$$

Relations of this type are going to be very useful in the future.

1.4 The two-dimensional geometric algebra Cl(2)

A **geometric algebra** is the algebra of multivectors with the geometric product as its operation. We can build a geometric algebra by specifying its basis vectors and their squares. For instance, the three-dimensional euclidean geometric algebra is built from the basis vectors e_1, e_2, e_3 that all square to 1,

$$(e_1)^2 = (e_2)^2 = (e_3)^2 = 1. \tag{1.96}$$

We denote a geometric algebra with $p+q$ basis vectors of which p square to 1 and q square to -1 as $\mathrm{Cl}(p,q)$. If all the basis vectors square to 1, the space is **euclidean** and we denote its geometric algebra as $\mathrm{Cl}(p)$. If there are also vectors that square to -1, the space is **hyperbolic**. A prominent example of such a space is spacetime from special relativity, with one vector that squares to 1 and three others that square to -1. The spacetime algebra is denoted as $\mathrm{Cl}(1,3)$.

But for now, we are going to treat a much easier example: The two-dimensional geometric algebra $\mathrm{Cl}(2)$. It has two basis vectors:

$$e_1, e_2. \tag{1.97}$$

The only higher-grade k-vector we can construct from these vectors is the bivector e_12. Therefore, we only have four basis k-vectors:

- one scalar 1
- two vectors e_1, e_2
- one bivector e_{12}.

This makes the 2D geometric algebra $\mathrm{Cl}(2)$ particularly simple. In mathematical language, we can write this as:

$$\mathrm{Cl}(2) = \mathrm{span}\{1, e_1, e_2, e_{12}\} \tag{1.98}$$

In the following, we are going to examine its properties.

1.4.1 The even subalgebra $\mathrm{Cl}^+(2)$ and complex numbers

The even subalgebra $\mathrm{Cl}^+(p,q)$ of some geometric algebra $\mathrm{Cl}(p,q)$ is defined as the subalgebra that only contains k-vectors of even grade. For instance, the even subalgebra $\mathrm{Cl}^+(2)$ is composed only of the scalar and the bivector. Thus, any element $z \in \mathrm{Cl}^+(2)$ can be written as

$$z = a + be_{12}. \tag{1.99}$$

The unit bivector e_{12} squares to -1:

$$(e_{12})^2 = e_{1212} = -e_{1221} = -e_{11} = -1. \tag{1.100}$$

This is familiar to us - these elements z look just like complex numbers! Complex numbers are composed of scalars and "imaginary units",

$$z = a + bi \tag{1.101}$$

where i squares to -1:

$$i^2 = -1. \tag{1.102}$$

All complex numbers commute with each other. At first glance, it looks like this nice analogy between $\mathbb{C}$ and $\mathrm{Cl}^+(2)$ could fall apart because of that - after all, we have seen that the geometric product is not necessarily commutative. But in fact, we realize that:

- Scalars commute with themselves and with bivectors.
- The bivectors e_{12} commutes with itself and scalar multiples of itself.

In other words: All elements of $\mathrm{Cl}^+(2)$ commute with each other. This means that we can in fact translate any complex number $z = a + bi$ to an even multivector $z = a + be_{12}$ and do the exact same calculations with them. And we finally get a geometric interpretation for imaginary numbers - the imaginary unit i can be interpreted as the bivector e_{12}. Complex numbers can be interpreted as a sum of a scalar and a bivector.

We can do anything with even multivectors z. The analogue of complex conjugation is the **multivector reverse** $\widetilde{M}$. It reverses the order of all geometric products in the expression, but leaves it untouched otherwise:

Multivector reverse

$$\widetilde{1} = 1 \tag{1.103}$$
$$\widetilde{e_i} = e_i \tag{1.104}$$
$$\widetilde{e_{ij}} = e_{ji} = -e_{ij} \tag{1.105}$$
$$\widetilde{e_{ijk}} = e_{kji} = -e_{ijk} \tag{1.106}$$
$$\widetilde{e_{ijkl}} = e_{lkji} = -e_{ijkl} \tag{1.107}$$
$$\dots \tag{1.108}$$

If we apply the reverse operation on the even multivector $z = a + be_{12}$, we get:

$$\widetilde{z} = a + be_{21} = a - be_{12}. \tag{1.109}$$

This is exactly what we'd expect from the complex conjugate.

We can take the exponential of an imaginary number:

$$\exp(i\alpha) = \cos(\alpha) + i\sin(\alpha). \tag{1.110}$$

This works out because the series expansion of the exponential,

$$\exp(i\alpha) = \sum_{k=0}^{\infty} \frac{(i\alpha)^k}{k!} \tag{1.111}$$

splits into two parts: The even ones that are real-valued because $i^{2n} = (-1)^n$, and the odd ones that are imaginary-values because $i^{2n+1} = (-1)^n i$:

$$\exp(i\alpha) = \left(\sum_{l=0}^{\infty} \frac{(-1)^l \alpha^{2l}}{(2l)!}\right) + i\left(\sum_{l=0}^{\infty} \frac{(-1)^l \alpha^{2l+1}}{(2l+1)!}\right) \tag{1.112}$$

These are the respecive series expansions for $\cos(\alpha)$ and $\sin(\alpha)$. And we can do exactly the same for the bivector exponential

Bivector exponential

$$\exp(e_{12}\alpha) = \cos(\alpha) + e_{12}\sin(\alpha). \tag{1.113}$$

This bivector exponetial follows the exact same rules as the normal complex exponential:

$$\widetilde{\exp(e_{12}\alpha)} = \exp(-e_{12}\alpha) \tag{1.114}$$

$$\exp(e_{12}\alpha) + \exp(e_{12}\beta) = \exp(e_{12}(\alpha+\beta)) \tag{1.115}$$

1.4.2 Two-dimensional rotations

One of the primary uses of complex numbers in physics is to encode rotations through the complex plane. For instance, we might write down the formula

$$\mathbf{x}(t) = \mathbf{x}_0 e^{i\omega t} \tag{1.116}$$

with

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ -i \end{pmatrix} \tag{1.117}$$

to encode two-dimensional rotational motion with the angular frequency ω. The resulting expression would then be

$$\mathbf{x}(t) = \begin{pmatrix} \cos(\omega t) + i\sin(\omega t) \\ \sin(\omega t) - i\cos(\omega t) \end{pmatrix}. \tag{1.118}$$

This way of writing down oscillations is pretty ubiquitous in theoretical physics. When we do this, there is an implicit convention to only interpret the real part as something physical and to throw the imaginary part away. Therefore, we actually ought to write

$$\mathbf{x}(t) = \begin{pmatrix} \cos(\omega t) \\ \sin(\omega t) \end{pmatrix}. \tag{1.119}$$

But what *is* the imaginary part, then? It is a mathematical artifact to make our calculations somewhat easier. We've all gotten used to it by now, but wouldn't it be better if we found a way to encode rotations without having to resort to unphysical mathematical parts at all?

Luckily, the $\mathbb{C} \simeq \mathrm{Cl}^+(2)$ isomorphism provides us with the right tool we need to eliminate these unphysical imaginary parts. Let $\mathbf{v} = xe_1 + ye_2$ be a 2D vector. We can rotate it by a quarter-turn by writing

$$\mathbf{v}' = \mathbf{v}e_{12}. \tag{1.120}$$

We can show this by rotating the basis vectors

$$e_1 e_{12} = e_2 \tag{1.121}$$
$$e_2 e_{12} = -e_1. \tag{1.122}$$

Note that in contrast to complex numbers, this multiplication is **not** commutative! As you can check for yourself,

$$\mathbf{v} e_{12} = -e_{12} \mathbf{v}. \tag{1.123}$$

for all $\mathbf{v}$.

If we rotate by e_{12} two times, we get the original vector times -1.

$$\mathbf{v} e_{12} e_{12} = -\mathbf{v} \tag{1.124}$$

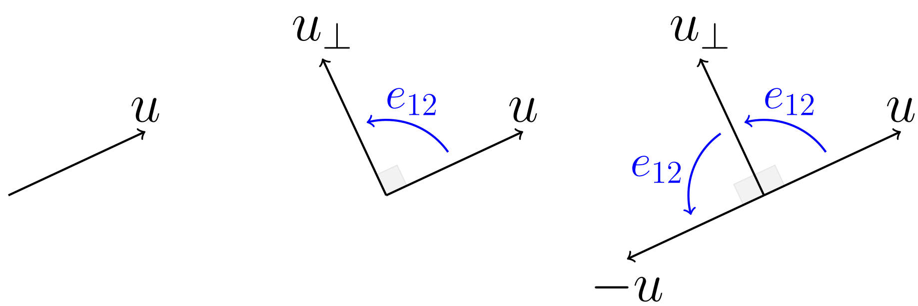


Figure 1.3: Applying the unit bivector e_{12} flips the vector $\mathbf{u}$ by a quarter-turn, $\frac{\tau}{4}$ radians. Two quarter-turns are equal to a sign flip $\mathbf{u} \to -\mathbf{u}$. This means that we can interpret the bivector e_{12} as a square root of -1.

We can also use the bivector exponential for this - if we want to rotate the 2D vector $\mathbf{v}$ by α degrees in the positive sense, we write:

2D rotations of vectors

$$\mathbf{v}' = \mathbf{v} \exp(e_{12}\alpha). \tag{1.125}$$

Again, keep in mind that we have to pay attention to the order - if we were to apply the bivector exponential from the other side, the vector would be rotated in the *negative* sense.

This approach has the advantage that contrary to complex numbers, it distinguishes between vectors and the objects rotating the vectors. This confusion ultimately is what led to the aforementioned problem with unphysical imaginary parts.

Also, the bivector $e_{12}\alpha$ is our first example of a so-called **rotation bivector** - a bivector whose orientation defines the plane we are rotating along, and whose magnitude defines the angle of the rotation. In two dimensions, there is only one possible plane we can rotate along, so all 2D rotation bivectors can be written $e_{12}\alpha$. In three dimensions, however, this will be more interesting, as we will see in the following chapter.

This rotation formula works only for vectors, and only in two dimensions. In the next chapter, we are going to learn about rotors and the rotor transformation law - they work for arbitrary multivectors and arbitrary dimensions.

Chapter 2

Geometric algebra of space

Chapter summary

- The geometric algebra of space Cl(3) describes objects in three-dimensional space.
- There is only one trivector in 3D: e_{123}. We call it the pseudoscalar.
- Every axial vector corresponds to a bivector. We call bivectors pseudovectors.
- The cross product has several undesirable properties. We can fully replace it with the wedge product and bivectors.
- We can perform 3D rotations by sandwiching two reflections. The resulting objects performing the rotation are called rotors.
- Rotation bivectors describe the plane and angle of a rotation. Rotors are formed by exponentiating rotation bivectors.
- Quaternions are left-handed 3D bivectors.
- We can perform infinitesimal rotations with the commutator brackets.

2.1 Overview

The **geometric algebra of space** Cl(3) is the geometric algebra describing three-dimensional space. Its vector basis consists of the three unit vectors

$$e_1, \quad e_2, \quad e_3. \tag{2.1}$$

This gives us the following possible basis k-vectors:

- 1 scalar
- 3 vectors e_1, e_2, e_3
- 3 bivectors e_{12}, e_{23}, e_{31},
- 1 trivector e_{123}.

$$1$$

$$e_1 \quad e_2 \quad e_3$$

$$e_{23} \quad e_{31} \quad e_{12}$$

$$e_{123}$$

Figure 2.1: The multivector pyramid of the space algebra.

2.1.1 3D pseudoscalars

Now that we have three dimensions instead of two, the maximum grade of our algebra is 3 - we get trivectors like e_{123}. They represent volume elements. There is only one possible orientation for trivectors, so every trivector we can write down will be proportional to e_{123}. For this reason, 3D trivectors are called pseudoscalars. We define the 3D unit pseudoscalar

$$I = e_{123}. \tag{2.2}$$

The exact order 123 is a matter of convention - we could've also defined $I' = e_{132} - I$ to be the unit pseudoscalar. By choosing either I or $-I$, we choose an orientation for our space - conventionally, we use right-handed coordinates, so we choose I. A left-handed space would have $-I$ as its unit pseudoscalar.

Just like in two dimensions, the unit pseudoscalar I squares to -1:

$$\begin{aligned} I^2 &= e_{123}^2 && (2.3)\\ &= e_{123123} && (2.4)\\ &= e_{112323} && (2.5)\\ &= e_{2323} && (2.6)\\ &= -e_{2233} && (2.7)\\ &= -e_{33} && (2.8)\\ &= -1. && (2.9) \end{aligned}$$

The 3D pseudoscalar trivially commutes with scalars and with itself, but also with vectors and bivectors:

$$e_i I = I e_i \tag{2.10}$$

$$e_{ij} I = I e_{ij} \tag{2.11}$$

We have seen how to replace complex numbers with the even subalgebra of Cl(2). This works just as well with the scalar-pseudoscalar subalgebra $\mathrm{Cl}^{1,I}(3)$ of the geometric algebra of space - that is, the subalgebra composed only of elements of the form

$$z = a + be_{123} = a + bI. \tag{2.12}$$

It largely depends on the context which one of these replacements is better suited. If we want to replace complex numbers with no geometric meaning, for instance in the context of complex analysis, we

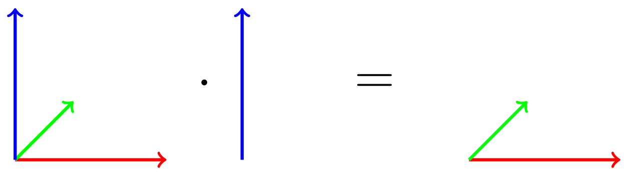


Figure 2.2: The equation $Ie_3 = e_{12}$ visualized.

often choose the Cl(2) scalar-pseudoscalar method. However, in other fields, complex numbers do carry geometric significance - for instance, in electromagnetism, we will replace the imaginary unit with the four-dimensional spacetime pseudoscalar.

The next thing we notice when taking a look at the multivector pyramid (Figure 2.1) is that there are just as many bivectors as vectors now. In two dimensions, there was only one possible orientation for an area element, but now, there are three - we can write an arbitrary bivector B as the linear combination

$$B = ae_{12} + be_{23} + ce_{31}. \tag{2.13}$$

This means that there is a one-to-one correspondence between vectors $\mathbf{v}$ and bivectors B - for every vector $\mathbf{v} = v^i e_i$, there is a corresponding bivector $B = I\mathbf{v}$. For instance, the bivector corresponding to the vector e_3 would be

$$Ie_3 = e_{123}e_3 = e_{12}. \tag{2.14}$$

The bivector $I\mathbf{v}$ describes the area orthogonal to the vector e_3 (see Figure 2.2). If we want to obtain the original vector from the bivector, we simply premultiply by $I^{-1} = -I$:

$$I^{-1}e_{12} = -Ie_{12} = -e_{123}e_{12} = -e_{11232} = -e_{232} = e_{223} = e_3, \tag{2.15}$$

or, more generally:

$$-IB = -I(I\mathbf{v}) = \mathbf{v}. \tag{2.16}$$

One might therefore ask why we need 3D bivectors at all, if the information of a bivector can also be encoded in a vector. In fact, this is the stance that most physicists implicitly take - in traditional vector algebra, areas are always described by their normal vectors. However, there is one important caveat here: Strictly speaking, 3D bivectors do not correspond to vectors - they correspond to *axial* vectors[1]. Axial vectors normally behave like normal vectors, but once reflections and parity flips are involved, things get messy.

For instance, the angular momentum axial vector $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ describes the circular motion of a rotating object. In order to find out which direction of rotation a given $\mathbf{L}$ describes, we use the right-hand rule: We point the thumb of our right hand in the direction of $\mathbf{L}$ and form a fist with the four other fingers.

Now, imagine that the whole scene is being reflected in a mirror parallel to the axis of rotation. While the plane of rotation described by the bivector $\mathbb{L} = \mathbf{x} \wedge \mathbf{p}$ behaves as we'd expect under this reflection, but the axial vector $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ includes an extra sign flip. Figure 2.3 shows this for the example $\mathbb{L} = Le_{12}$.

In mathematical terms, axial vectors are defined via their behaviour under parity flips. A parity flip P is the operation that flips the sign of all vectors:

Parity flip

$$P(\mathbf{x}) = -\mathbf{x} \tag{2.17}$$

[1] In traditional math, the terms "axial vector" and "pseudovector" are used synonymously. We call vectors that perform an extra sign flip under a reflection or parity flip "axial vectors", and bivectors "pseudovectors".

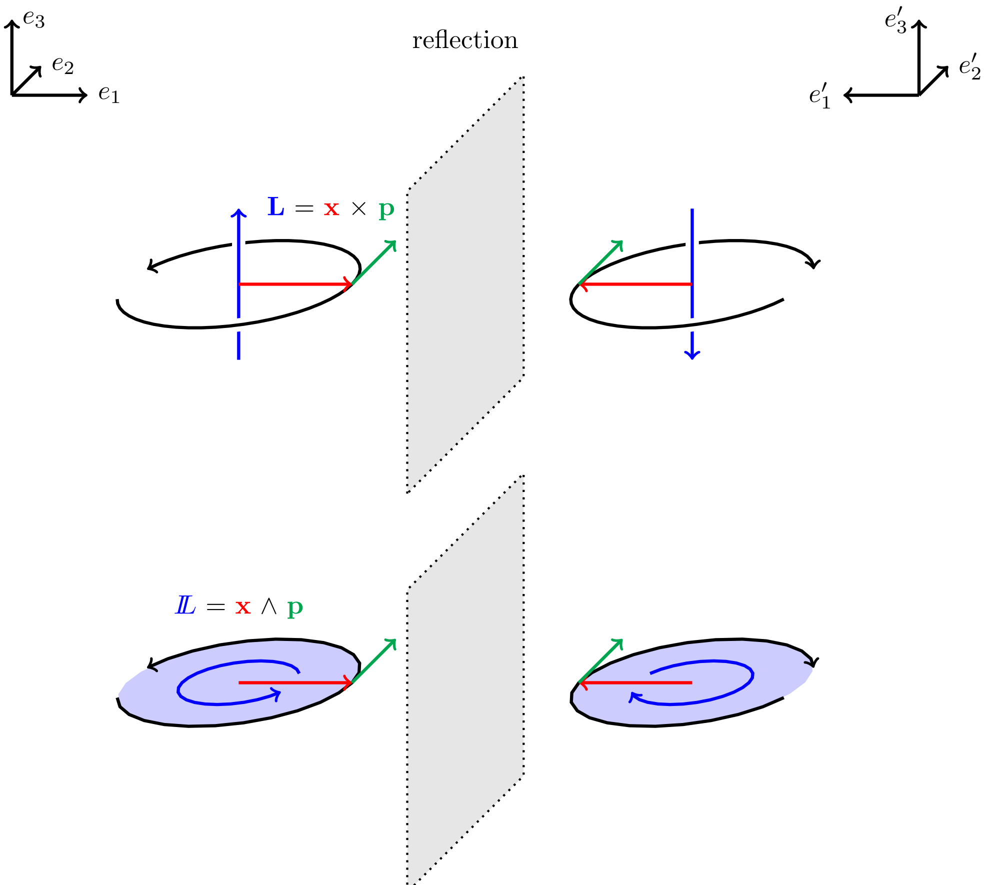


Figure 2.3: The behaviour of angular momentum under reflections. Evidently, it makes much more sense to use bivectors instead of axial vectors to describe rotational motion.

If we combine this parity flip with a rotation by 180°, we get a reflection as shown in Figure 2.3. Axial vectors $\mathbf{c}$, however, do not flip their sign under parity transforms. They just stay the same - which makes no sense at all, if you ask me. Hence, in geometric algebra, we generally banish all axial vectors like $\mathbf{L}$ in favour of bivectors like $\mathbb{L} = I\mathbf{L}$. For this reason, we also refer to 3D bivectors as **pseudovectors**. The reason why bivectors do not flip their sign under P is very intuitive -

$$
\begin{aligned}
P(\mathbf{a} \wedge \mathbf{b}) &= P(\mathbf{a}) \wedge P(\mathbf{b}) && (2.18)\\
&= (-\mathbf{a}) \wedge (-\mathbf{b}) && (2.19)\\
&= \mathbf{a} \wedge \mathbf{b}. && (2.20)
\end{aligned}
$$

On the other hand, the pseudoscalar I will flip its sign:

$$
\begin{aligned}
P(I) &= P(e_1 \wedge e_2 \wedge e_3) && (2.21)\\
&= P(e_1) \wedge P(e_2) \wedge P(e_3) && (2.22)\\
&= (-e_1) \wedge (-e_2) \wedge (-e_3) && (2.23)\\
&= -e_1 \wedge e_2 \wedge e_3 && (2.24)\\
&= -I. && (2.25)
\end{aligned}
$$

In other words - after a parity flip, our space goes from right-handed (I) to left-handed ($-I$).

2.1.2 Eliminating the cross product

Now that we have vowed to drive out all axial vectors of physics, the next logical step is to banish the cross product - the appearance of axial vectors in the cross product is unavoidable. In Figure 2.3, we have seen that

$$\textbf{vector} \times \textbf{vector} = \textbf{axial vector}. \qquad (2.26)$$

We can deduce that

$$
\begin{aligned}
\textbf{axial vector} \times \textbf{vector} &= \textbf{vector} && (2.27)\\
\textbf{vector} \times \textbf{axial vector} &= \textbf{vector} && (2.28)\\
\textbf{axial vector} \times \textbf{axial vector} &= \textbf{axial vector}. && (2.29)
\end{aligned}
$$

The cross product between two regular vectors $\mathbf{a} \times \mathbf{b}$ gives us the axial vector normal to the plane spanned by $\mathbf{a}$ and $\mathbf{b}$. In geometric algebra language, this means that we want to find the vector orthogonal to the bivector $\mathbf{a} \wedge \mathbf{b}$:

$$\mathbf{a} \times \mathbf{b} = -I(\mathbf{a} \wedge \mathbf{b}). \qquad (2.30)$$

But since we have decided that we want to replace all axial vectors $\mathbf{v}$ with their bivector counterparts, we are going to write:

Replacement for the cross product between two vectors

$$\mathbf{a} \wedge \mathbf{b} = I(\mathbf{a} \times \mathbf{b}). \qquad (2.31)$$

But what if one operand of the cross product is an axial vector $\mathbf{c} = IC$? In that case, we have to rewrite the cross product in terms of the bivector C, such that it becomes::

$$
\begin{aligned}
\mathbf{b} &= \mathbf{a} \times \mathbf{c} && (2.32)\\
&= -I(\mathbf{a} \wedge \mathbf{c}) && (2.33)\\
&= -I(\mathbf{a} \wedge (IC)) && (2.34)\\
&= -\frac{1}{2} I(\mathbf{a}IC - IC\mathbf{a}) && (2.35)\\
&= -\frac{1}{2} I(I\mathbf{a}C - IC\mathbf{a}) && (2.36)\\
&= \frac{1}{2}(\mathbf{a}C - C\mathbf{a}) && (2.37)\\
&= \mathbf{a} \cdot C. && (2.38)
\end{aligned}
$$

The cross product between a vector and an axial vector translates to the interior product between a vector and the bivector corresponding to the cross product.

Replacement for the cross product between a vector and an axial vector

$$\mathbf{a} \times \mathbf{c} = \mathbf{a} \cdot C \tag{2.39}$$

Similarly,

$$\mathbf{c} \times \mathbf{a} = C \cdot \mathbf{a} = -\mathbf{a} \cdot C \tag{2.40}$$

When we cross-multiply two axial vectors $\mathbf{c} = IC, \mathbf{d} = ID$, we get the axial vector:

$$
\begin{aligned}
\mathbf{c} \times \mathbf{d} &= -I(\mathbf{c} \wedge \mathbf{d}) && (2.41)\\
&= -\frac{1}{2} I((IC) \wedge (ID)) && (2.42)\\
&= -\frac{1}{2} I(ICID - IDIC) && (2.43)\\
&= -\frac{1}{2} I(I^2CD - I^2DC) && (2.44)\\
&= \frac{1}{2} I(CD - DC), && (2.45)
\end{aligned}
$$

which corresponds to the bivector

Replacement for the cross product between two axial vectors

$$I(\mathbf{c} \times \mathbf{d}) = \frac{1}{2}(DC - CD) = \langle DC \rangle_2 \tag{2.46}$$

One can geometrically imagine this product as follows: In 3D space, two bivectors will always intersect along a line. The line can be defined by a unit vector $\mathbf{n}$. This means that C and D can be described as the wedge between $\mathbf{n}$ and some other vectors $\mathbf{x}, \mathbf{y}$ orthogonal to $\mathbf{n}$:

$$
\begin{aligned}
C &= \mathbf{xn} && (2.47)\\
D &= \mathbf{yn} && (2.48)
\end{aligned}
$$

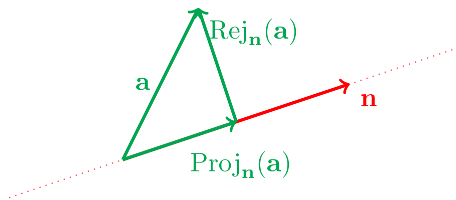


Figure 2.4: The projection and rejection of a vector $\mathbf{a}$ with respect to $\mathbf{n}$

The product $\langle DC \rangle_2$ can be described as:

$$\begin{aligned} \langle DC \rangle_2 &= \langle ynxn \rangle_2 && (2.49)\\ &= -\langle ynnx \rangle_2 && (2.50)\\ &= -\langle yx \rangle_2 && (2.51)\\ &= x \wedge y. && (2.52) \end{aligned}$$

These substitution rules allow us to fully eliminate the cross product from physics - and we will do so in GA. The widespread adoption of the concepts of axial vectors and the cross product was a grave mistake we intend to correct. This becomes particularly obvious in the field of electromagnetism.

2.2 Transformations

One of the main applications of the cross product is the mathematical description of rotational movement. Therefore, we are now going to investiage how to describe rotations.

The outline of this section is as follows: First, we will define the **projection** and **rejection** operations. Then, we will use them to define the reflection operation, which we will in turn use to define rotations and rotors.

2.2.1 Projection and rejection

Let $\mathbf{a}$ be some vector and $\mathbf{n}$ a normal vector, $\mathbf{n}^2 = 1$. We can decompose $\mathbf{a}$ into a parallel and orthogonal part with respect to $\mathbf{n}$ by writing:

$$\begin{aligned} \mathbf{a} &= \mathbf{a}\mathbf{n}^2 && (2.53)\\ &= \mathbf{a}\mathbf{n}\mathbf{n} && (2.54)\\ &= (\mathbf{a}\cdot\mathbf{n} + \mathbf{a}\wedge\mathbf{n})\mathbf{n} && (2.55)\\ &= (\mathbf{a}\cdot\mathbf{n})\mathbf{n} + (\mathbf{a}\wedge\mathbf{n})\mathbf{n} && (2.56) \end{aligned}$$

The interior product $a \cdot n$ measures how much $\mathbf{a}$ and $\mathbf{n}$ coincide, such that $(\mathbf{a}\cdot\mathbf{n})\mathbf{n}$ is the projection of a onto the line defined by n. Similarly, $\mathbf{a}\wedge\mathbf{n}$ measures the "orthogonality" of $\mathbf{a}$ with respect to $\mathbf{n}$, such that $(\mathbf{a}\wedge\mathbf{n})\mathbf{n}$ is the part of $\mathbf{a}$ that is orthogonal to $\mathbf{n}$. This is depicted in Figure 2.4. We call the former the **projection** and the latter the **rejection** and denote them as

Projection and rejection

$$\begin{aligned} \text{Proj}_{\mathbf{n}}(\mathbf{a}) &= (\mathbf{a}\cdot\mathbf{n})\mathbf{n} && (2.57)\\ \text{Rej}_{\mathbf{n}}(\mathbf{a}) &= (\mathbf{a}\wedge\mathbf{n})\mathbf{n}. && (2.58) \end{aligned}$$

2.2.2 Reflections

We now turn to the description of reflections by a mirror plane, like the one we have performed in Figure 2.3. We know this type of reflection from everyday life - it's what we see when we look in the mirror. This is why we are used to think about it as a reflection "by a mirror plane". But let's take a look at what actually happens with the vector coordinates when we perform this reflection, for instance by the y-z plane:

$$\mathbf{x} = (x^1, x^2, x^3) \rightarrow (-x^1, x^2, x^3). \tag{2.59}$$

Only the sign of the component of e_1 is flipped, while the two components associated with the plane we were reflecting by are unchanged. That seems a bit odd. In fact, it turns out that describing these reflections gets a lot easier if we do not think of them as reflections by a plane, but as reflections *along the normal vector* of the plane. For instance, we will now think about (2.59) as a reflection along the x-axis.

Let $\mathbf{n}$ be the normal vector of the plane of reflection, and $\mathbf{a}$ the vector we want to reflect along $\mathbf{n}$. Using the language of the previous section, we can split up $\mathbf{a}$ as follows:

$$\mathbf{a} = \text{Proj}_{\mathbf{n}}(\mathbf{a}) + \text{Rej}_{\mathbf{n}}(\mathbf{a}) \tag{2.60}$$

If we now want to reflect $\mathbf{a}$ along $\mathbf{n}$, we simply flip the sign of the projection:

$$\mathbf{a}' = -\text{Proj}_{\mathbf{n}}(\mathbf{a}) + \text{Rej}_{\mathbf{n}}(\mathbf{a}) \tag{2.61}$$

Inserting the respective definitions, we get

$$\begin{aligned} \mathbf{a}' &= -(\mathbf{a} \cdot \mathbf{n})\mathbf{n} + (\mathbf{a} \wedge \mathbf{n})\mathbf{n} && (2.62)\\ &= (-\mathbf{a} \cdot \mathbf{n} + \mathbf{a} \wedge \mathbf{n})\mathbf{n} && (2.63)\\ &= (-\mathbf{n} \cdot \mathbf{a} - \mathbf{n} \wedge \mathbf{a})\mathbf{n} && (2.64)\\ &= -\mathbf{nan}. && (2.65) \end{aligned}$$

This is the formula for reflections along a vector.

Reflection of a vector along a vector

$$\mathbf{a}' = -\mathbf{nan} \tag{2.66}$$

In fact, it works very similarly for higher grades, and not just for vectors. For instance, a bivector $B = \mathbf{ab}$ would reflect like

$$B' = \mathbf{a}'\mathbf{b}' = (-\mathbf{nan})(-\mathbf{nbn}) = \mathbf{nannbn} = \mathbf{nabn} = \mathbf{n}B\mathbf{n}. \tag{2.67}$$

2.2.3 Rotations

Now, we want to describe rotations. The first thing to realize is that every rotation can be decomposed into two successive reflections.

Two successive reflection along the same axis $\mathbf{n}$ do nothing to a vector $\mathbf{a}$:

$$\begin{aligned} \mathbf{a}'' &= -\mathbf{n}\mathbf{a}'\mathbf{n} = -\mathbf{n}(-\mathbf{nan})\mathbf{n} && (2.68)\\ &= \mathbf{n}^2\mathbf{a}\mathbf{n}^2 && (2.69)\\ &= \mathbf{a}. && (2.70) \end{aligned}$$

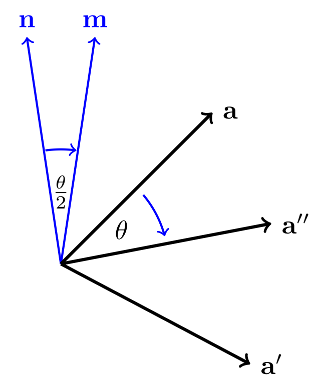


Figure 2.5: A rotation along an angle of θ can be performed by performing a double reflection along **n** and **m**. The angle between these two vectors has to be $\theta/2$.

However, if we first reflect along **n**, and then along another unit vector **m**, we get:

$$\mathbf{a}' = -\mathbf{nan} \tag{2.71}$$
$$\mathbf{a}'' = -\mathbf{ma}'\mathbf{m} \tag{2.72}$$
$$= \mathbf{mnanm}. \tag{2.73}$$

A single reflection is an orthonormal transformation i.e. an element of the group O(3). It has determinant -1. Two successive reflections, however, have determinant $(-1)^2 = 1$ - they are *special* orthogonal transformations in SO(3). Special orthogonal transformations in 3D are always rotations.

Taking a look at Figure 2.5, we can see that this double reflection is indeed a rotation. To be specific, it is a rotation along the plane defined by the bivector $\mathbf{n} \wedge \mathbf{m}$. The angle **a** is rotated by *twice* the angle between **n** and **m**. To make our life easier, we are going to define the so-called rotor

Constructing rotors from vectors

$$R = \mathbf{nm}, \tag{2.74}$$

that describes this rotation. Then, the formula for a rotation is:

Rotor law

$$\mathbf{a}'' = \tilde{R}\mathbf{a}R. \tag{2.75}$$

The tilde $\tilde{R}$ denotes the multivector reverse.

Analogously to the 2D case, we know that the interior and exterior products **nm** are equal to:

$$\mathbf{n} \cdot \mathbf{m} = \cos(\theta/2) \tag{2.76}$$
$$\mathbf{n} \wedge \mathbf{m} = \sin(\theta/2)B \tag{2.77}$$

where

$$B = \frac{\mathbf{n} \wedge \mathbf{m}}{\sqrt{-(n \wedge m)^2}} \tag{2.78}$$

is the unit bivector along which the rotation is performed. We can therefore say that

$$R = \cos(\theta/2) + B\sin(\theta/2). \tag{2.79}$$

This should look vaguely familiar - it looks a bit like the formula for complex exponentials,

$$\exp(i\theta/2) = \cos(\theta/2) + i\sin(\theta/2). \tag{2.80}$$

This formula arises because the polynomial expansion for the exponential splits up into a real and an imaginary part:

$$\exp(i\theta/2) = \sum_{k=0}^{\infty} \frac{i^k(\theta/2)^k}{k!} \tag{2.81}$$

$$= \sum_{l=0}^{\infty} \frac{i^{2l}(\theta/2)^{2l}}{(2l!)} + \sum_{l=0}^{\infty} \frac{i^{2l+1}(\theta/2)^{2l+1}}{(2l+1)!} \tag{2.82}$$

$$= \sum_{l=0}^{\infty} \frac{(-1)^l(\theta/2)^{2l}}{(2l!)} + i\sum_{l=0}^{\infty} \frac{(-1)^l(\theta/2)^{2l}}{(2l+1)!} \tag{2.83}$$

$$= cos(\theta/2) + i\sin(\theta/2). \tag{2.84}$$

This derivation relied on the fact that $i^2 = -1$. And we will in fact see that all 3D unit bivectors square to -1 too! For instance, $B = e_{31}$ describes the rotation along the zx plane:

$$(e_{31})^2 = e_{3131} = -e_{3311} = -1 \tag{2.85}$$

So we can write:

$$R = \exp(B\theta/2) = \cos(\theta/2) + B\sin(\theta/2), \tag{2.86}$$

the rotor that performs a rotation along the zx plane by an angle of θ.

This is an extremely useful fact. Given a **rotation bivector** Θ (capital theta) whose orientation describes the plane of rotation and whose magnitude describes the angle of rotation, we can simply find the corresponding rotor by calculating:

Constructing rotors from bivector exponentials

$$R = \exp(\Theta/2). \tag{2.87}$$

Both conceptually and practically, this is a lot simpler than the Euler angle approach we had to take for rotation matrices. For instance, it doesn't suffer from gimbal locking.

There is also an added benefit from using bivectors to describe rotations: In classical vector algebra, we use so-called "Euler" or "rotation vectors" to describe the axis and angle of the rotation we want to perform. For instance, a counterclockwise rotation around the z axis by a quarter-turn would correspond to the rotation "vector"

$$\boldsymbol{\theta} = \frac{\tau}{4}e_3. \tag{2.88}$$

It suffers from the same problems as other axial vectors. We will thus instead use rotation bivectors $\Theta = \boldsymbol{I\theta}$. The above rotation would be represented by the bivector

$$\Theta = \frac{\tau}{4}e_{12}. \tag{2.89}$$

Thus, in geometric algebra, we do not talk about rotations "around an axis", but rather about "**rotations along a plane**" - for instance, a rotation along the xy plane.

2.2.4 Some facts about rotors

Given a rotor $R = \exp(\Theta/2)$, the reverse operation evaluates to

$$\tilde{R} = \exp(-\Theta/2). \tag{2.90}$$

We can thus directly see that for any rotor,

$$R\tilde{R} = \tilde{R}R = 1. \tag{2.91}$$

The rotor law $M \to RM\tilde{R}$ is valid for any multivector grade. For instance, a bivector $D = \mathbf{ab}$ would rotate like

$$D' = \mathbf{a}'\mathbf{b}' = \tilde{R}\mathbf{a}R\tilde{R}\mathbf{b}R = \tilde{R}\mathbf{ab}R = \tilde{R}DR. \tag{2.92}$$

If we want to perform two successive rotations R_1, R_2, we just multiply the two rotors together:

$$R_{\text{total}} = R_1 R_2. \tag{2.93}$$

There is something important to note: Given a specific rotation R, the rotor $R' = -R$ has the exact same effect on vectors:

$$(-\tilde{R})\mathbf{x}(-R) = \tilde{R}\mathbf{x}R. \tag{2.94}$$

To understand this better, we take a look at the simplest pair of rotors where this peculiarity arises - $R = +1$ and $R' = -1$. Obviously, $R = 1$ corresponds to a rotation by an angle of zero, $\exp(0) = 1$. The second rotor, $R' = -1$ however, corresponds to a rotation by $\tau = 360°$:

$$R' = \exp\left(\frac{B}{2}\tau\right) = \cos(\tau/2) + B\sin(\tau/2) \tag{2.95}$$

$$= -1. \tag{2.96}$$

If we were to rotate by $2\tau = 720°$, we would obtain the original rotor again:

$$\exp\left(\frac{B}{2}2\tau\right) = \cos(2\tau/2) + B\sin(2\tau/2) \tag{2.97}$$

$$= 1. \tag{2.98}$$

This is a very interesting fact - in the rotor formalism, a rotation by 0° is not the same as a rotation by 360°. They correspond to two different rotors, but because of the geometric two-sided transformation law, they have the same effect on vectors. This is not some weird mathematical artifact of geometric algebra. It corresponds to an elementary geometric fact about rotations: The 3D group of rotations SO(3) is **not simply connected**. This means that if we trace out a path starting and ending at the same rotational state, it's not guaranteed that we will be able to smoothly deform this path into a point.

To elaborate on this, imagine a three-dimensional sphere S^2. If we draw a loop on its surface, we can always continuously shrink it into a point. On a 3D torus T^3, however, this is not always possible - if the loop makes one turn around the torus, we cannot deform it into a point without tearing it apart. We say that S^2 is simply connected, but T^3 isn't.
Doing the same thought experiment for elements of SO(3) - 9-dimensional matrices with some constraints - can be pretty difficult. Luckily, there is a trick devised by Paul Dirac to visualize this very intuitively - the so-called Dirac belt trick. In this trick, we use a belt to visualize a path we might take through SO(3). The two ends of the belt represent the start and end of our path through SO(3). The cross-section of the belt at any point can be pictured as a vector that represents the current state of rotation. If the orientation of both ends is the same, the belt describes a loop through SO(3). If the belt is just a straight line without twists, it describes a point in SO(3). Continuous deformations amount to shifting and twisting the belt around, while keeping the orientation of both ends fixed. You can check for yourself that if there is one τ twist in the belt, it is not possible to deform it back into a straight line with these costraints - but if there are two τ twists in the belts, it is possible.

2.2.5 A short note on quaternions

Quaternions are the go-to solution of many physicists and engineers to describe rotations without gimbal locking. They are an associative algebra of three complex numbers i, j, k satisfying

$$i^2 = j^2 = k^2 = ijk = -1. \tag{2.99}$$

They are something rather abstract normally, but with geometric algebra, they just translate to the unit bivectors:

$$i = e_{12} \tag{2.100}$$
$$j = e_{23} \tag{2.101}$$
$$k = -e_{31}. \tag{2.102}$$

Therefore, rotors can be formulated in quaternion algebra too. (In quaternion lingo, rotors are called "versors").

The even subalgebra $\mathrm{Cl}^+(3)$ only contains scalars and bivectors. Therefore, the even subalgebra of Cl(3) is isomorphic to the space of quaternions $\mathbb{H}$ - we can translate any quaternion

$$z = a + bi + cj + dk \tag{2.103}$$

to an element of the even subalgebra:

$$z = a + be_{12} + ce_{23} - de_{31}. \tag{2.104}$$

This is starting to look a lot like the isomorphism between complex numbers and the even subalgebra $\mathrm{Cl}^+(2)$. In two dimensions, we had one possible rotation and one bivector (e_{12}) that generated it. This bivector could be translated to the imaginary unit i. Now, in three dimensions, we have three rotations and three bivectors generating them (e_{12}, e_{23}, e_{31}). They can be translated to the unit quaternions $i, j, -k$[2].

However, both the complex numbers and the quaternions share the same issues: Neither has a concept of "vectors" in the proper sense. The object performing the rotation (the rotor, or "versor" in quaternion lingo) consists of the same basis elements as the object being rotated. This is very confusing, and it doesn't help that the quaternions themselves do not have any geometric interpretations. Instead, authors often

[2]In fact, this is how Quaternions originally arose - Hamilton was asking himself how the rotation math of complex numbers could be generalized to three dimensions.

resort to projective geometry black magic to explain them, and as a result, no one really understands quaternions. Luckily, we can now see that they are just bivectors, and get a geometric intuition for them!

Also, you might have asked yourself why k translates to $-e_{31}$ instead of e_{31}. This is because the definition $ijk = -1$ is wrong - it should have been $ijk = 1$. The former definition results in a left-handed space. These peculiarities smash all hopes of doing any halfway sensible geometry with quaternions. So let's turn our back on them and just use vectors and bivectors instead.

2.2.6 Infinitesimal rotations

We now investigate infinitesimal rotations - rotations whose angle very small, such that we can neglect second-order terms.

Let Θ be a rotation bivector. Its angle should be much smaller than one - $|\Theta| \ll 1$. We can then approximate the bivector exponential as

$$R = \exp(\Theta/2) \approx 1 + \Theta/2. \tag{2.105}$$

This means that the infinitesimal rotor transformation law is

$$\begin{aligned} M' = \tilde{R}MR &= (1 - \Theta/2)M(1 + \Theta/2) & (2.106) \\ &\approx M + M\frac{\Theta}{2} - \frac{\Theta}{2}M & (2.107) \\ &= M + \left[M, \frac{\Theta}{2}\right] & (2.108) \end{aligned}$$

This is the infinitesimal rotor transformation law for a multivector,

Infinitesimal rotor transformation law

$$\tilde{R}MR \approx M + \left[M, \frac{\Theta}{2}\right] \tag{2.109}$$

If M is a vector, $M = \mathbf{x}$, this commutator can also be expressed as:

$$\left[\mathbf{x}, \frac{\Theta}{2}\right] = \frac{1}{2}\left(\mathbf{x}\Theta - \Theta\mathbf{x}\right) = \mathbf{x} \cdot \Theta. \tag{2.110}$$

We are going to use the commutator notation $[\mathbf{x}, B/2]$ when we want to emphasize that we are dealing with an infinitesimal rotation - it is independent of the grade of the object being rotated. For instance, if we wanted to rotate a bivector C, we can equivalently write

$$C' = C + \left[C, \frac{\Theta}{2}\right]. \tag{2.111}$$

The translation of this commutator would be:

$$\left[C, \frac{\Theta}{2}\right] = \langle C\Theta \rangle_2 \,, \tag{2.112}$$

the bivector grade of the geometric product $C\Theta$. Draw a grade diagram - you will see that the geometric product between two bivectors will result in a scalar, bivector and tetravector grade. We use the interior product $C \cdot \Theta$ to get the scalar grade and the exterior product $C \wedge \Theta$ to get the tetravector grade. We don't have any shorthand notation for the bivector grade though, so we have to write $\langle C\Theta \rangle_2$.

The dot notation $\mathbf{x} \cdot \Theta$ for rotating vectors is primarily useful when we want to think about the geometric details - from our geometric intuition, we can directly tell that:

- The value of $\mathbf{x} \cdot \Theta$ is maximal if $\mathbf{x}$ fully lies in the plane of Θ. This makes intuitive sense - a vector will change fastest if it lies in the plane of rotation.
- The value of $\mathbf{x} \cdot \Theta$ is zero if $\mathbf{x}$ is orthogonal to the plane of Θ. This also makes intuitive sense - $\mathbf{x}$ will not change at all under the rotation if if is orthogonal to the plane of rotation.

In the language of commutation relations, we say that:

- If $\mathbf{x}$ lies in the plane of Θ, then $\mathbf{x}$ and Θ anticommute:

$$\mathbf{x}\Theta = -\Theta\mathbf{x} \tag{2.113}$$

- If $\mathbf{x}$ is orthogonal to the plane of Θ, then $\mathbf{x}$ and Θ commute:

$$\mathbf{x}\Theta = \Theta\mathbf{x} \tag{2.114}$$

This means that the bivector exponential $R = \exp(\Theta/2)$ will behave as follows: If $\mathbf{x}$ lies in Θ, we have

$$\mathbf{x}R = \mathbf{x}\exp(\Theta/2) = \exp(-\Theta/2)\mathbf{x} = \tilde{R}\mathbf{x}, \tag{2.115}$$

while if $\mathbf{x}$ is orthogonal to Θ, we have

$$\mathbf{x}R = R\mathbf{x}. \tag{2.116}$$

In the second case, the rotor transformation law simply reduces to

$$\tilde{R}\mathbf{x}R = \tilde{R}R\mathbf{x} = \mathbf{x}. \tag{2.117}$$

The vector is unchanged by a rotation along the plane orthogonal to it. In the first case, however, we get:

$$\tilde{R}\mathbf{x}R = \mathbf{x}RR = \mathbf{x}R^2 = \mathbf{x}\exp(\Theta). \tag{2.118}$$

We know this - this is the rotation law from the two-dimensional geometric algebra! This makes intuitive sense - in 2D, all vectors lie inside the plane of all bivectors, so we can always simplify the full two-sided rotor law into a one-sided one.

2.2.7 Angular velocity

The rotor $R = \exp(\Theta/2)$ describes a rotation along the direction and with the magnitude of Θ. Now, we can ask ourselves how we can describe a rotational motion, i.e. a motion where the angle depends on the time t. With geometric algebra, this is as simple as

$$R(t) = \exp(\Theta(t)/2). \tag{2.119}$$

The change in time of this rotor is:

$$\dot{R}(t) = \frac{\dot{\Theta(t)}}{2}\exp(\Theta(t)/2) = \exp(\Theta(t)/2)\frac{\dot{\Theta}(t)}{2}. \tag{2.120}$$

This is just the standard formula for taking the derivative of exponentials. Note that Θ and R commute. Commonly, the derivative of the angle bivector $\Theta(t)$ is called the “angular velocity bivector”:

$$\Omega(t) = \dot{\Theta}(t) \tag{2.121}$$

If the angular velocity is constant, the above equations simplify to

$$R(t) = \exp(\Omega t/2) \tag{2.122}$$

$$\dot{R}(t) = \exp(\Omega t/2)\frac{\Omega}{2} = \frac{\Omega}{2}\exp(\Omega t/2). \tag{2.123}$$

Now suppose we start with a vector $\mathbf{x}_0$ and use the rotor $R(t)$ to rotate it around. We obtain the vector

$$\mathbf{x}(t) = \tilde{R}(t)\mathbf{x}_0 R(t). \tag{2.124}$$

If we take the time derivative, we get:

$$\dot{\mathbf{x}} = (\tilde{R}\mathbf{x}_0 R)^{\cdot} \tag{2.125}$$

$$= \dot{\tilde{R}}\mathbf{x}_0 R + \tilde{R}\mathbf{x}_0 \dot{R} \tag{2.126}$$

$$= -\frac{\dot{\Theta}}{2}\tilde{R}\mathbf{x}_0 R + \tilde{R}\mathbf{x}_0 R\frac{\dot{\Theta}}{2} \tag{2.127}$$

$$= \left[\tilde{R}\mathbf{x}_0 R, \frac{\dot{\Theta}}{2}\right], \tag{2.128}$$

$$= \left[\mathbf{x}, \frac{\Omega}{2}\right] \tag{2.129}$$

We can therefore say that the velocity of an object at coordinate $\mathbf{x}$ being rotated with angular velocity Ω around the origin is

Rotation velocity

$$\dot{\mathbf{x}} = \left[\mathbf{x}, \frac{\Omega}{2}\right] \tag{2.130}$$

This formula should make perfect intuitive sense - the angular velocity bivector represents an infinitesimal rotation after an infinitesimal time (think $\mathrm{d}\Theta = \mathrm{d}\Omega\ \mathrm{d}t$), so it can be viewed as a generator acting on $\mathbf{x}$.

We can even go one step further and find the **centripetal acceleration** by taking another time derivative:

$$\ddot{\mathbf{x}} = \left[\dot{\mathbf{x}}, \frac{\Omega}{2}\right] \tag{2.131}$$

$$= \left[\left[\mathbf{x}, \frac{\Omega}{2}\right], \frac{\Omega}{2}\right] \tag{2.132}$$

$$= \frac{1}{4}\left(\mathbf{x}\Omega^2 - \Omega\mathbf{x}\Omega - \Omega\mathbf{x}\Omega + \Omega^2\mathbf{x}\right) \tag{2.133}$$

The quantity Ω^2 is a negative scalar. Therefore, it commutes with everything, and we can write:

Centripetal acceleration

$$\ddot{\mathbf{x}} = \frac{1}{2}\left(\Omega^2\mathbf{x} - \Omega\mathbf{x}\Omega\right) \tag{2.134}$$

If $\mathbf{x}$ lies inside the plane of the angular velocity Ω, they anticommute. In this case, we can write $\Omega\mathbf{x}\Omega = -\Omega^2\mathbf{x}$, and the above formula becomes particularly simple:

$$\ddot{\mathbf{x}} = \Omega^2\mathbf{x} \tag{2.135}$$

In 3D, bivectors square to negative values, i.e. $\Omega^2 = -\omega^2$. This automatically tells us that the centripetal force has to pull the object into the center. If we multiply with m, we thus get the well-known formula for the centripetal force:

$$\mathbf{F} = m\ddot{\mathbf{x}} = -m\omega^2\mathbf{x}. \tag{2.136}$$

2.3 Differential operators

The operator ∂_i is defined as the partial derivative of some value by the coordinate i:

$$\partial_i a := \frac{\partial a}{\partial x^i}. \tag{2.137}$$

In traditional vector calculus, we use it to define the gradient operator

$$\boldsymbol{\nabla} = \begin{pmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix}. \tag{2.138}$$

When we apply this operator to a scalar field a, it gives us the direction of the steepest increase of the field $\boldsymbol{\nabla} a$.

If we want to apply $\boldsymbol{\nabla}$ to a vector field $\mathbf{v}$, we could take the dot product between them to obtain the divergence:

$$\boldsymbol{\nabla} \cdot \mathbf{v} = \partial_1 v^1 + \partial_2 v^2 + \partial_3 v^3, \tag{2.139}$$

or the cross product to obtain the curl:

$$\boldsymbol{\nabla} \times \mathbf{v} = \begin{pmatrix} \partial_2 v^3 - \partial_3 v^2 \\ \partial_3 v^1 - \partial_1 v^3 \\ \partial_1 v^2 - \partial_2 v^1 \end{pmatrix}. \tag{2.140}$$

Needless to say, these expressions do not generalize to higher dimensions like 4D spacetime or behave sensibly under parity flips.

2.3.1 The geometric derivative

Instead, we define the so-called **geometric derivative** or **vector derivative**[3]

Geometric derivative

$$\boldsymbol{\partial} = e^i \partial_i. \tag{2.141}$$

This derivative is an element of the geometric algebra Cl(3) - it is a differential operator-valued vector. It is very versatile: When we apply it to a scalar field a, it acts like the gradient operator:

$$\boldsymbol{\partial} a = e^i \partial_i a. \tag{2.142}$$

Things start to get interesting once we apply it to a vector field $\mathbf{v} = v^i e_i$:

$$\boldsymbol{\partial} \mathbf{v} = e^i \partial_i v^j e_j \tag{2.143}$$

$$= e_i e_j \partial^i v^j \tag{2.144}$$

Note that the order of the vectors e_i, e_j needs to stay fixed. The objects ∂_i, v^j are scalars, so we can commute them around as we like.

The product between $\boldsymbol{\partial}$ and $\mathbf{v}$ is the normal geometric product. It consists of the interior product and the exterior product. We call $\boldsymbol{\partial} \cdot \mathbf{v}$ the **interior derivative** and $\boldsymbol{\partial} \wedge \mathbf{v}$ the **exterior derivative**:

$$\boldsymbol{\partial} \cdot \mathbf{v} = (e^i \cdot e^j) \partial_i v_j = \partial_i v^i \tag{2.145}$$

$$\boldsymbol{\partial} \wedge \mathbf{v} = (e^i \wedge e^j) \partial_i v_j = \frac{1}{2}(\partial^i v^j - \partial^j v^i) e_i \wedge e_j \tag{2.146}$$

[3] Also denoted as $\boldsymbol{\nabla}$ instead of $\boldsymbol{\partial}$.

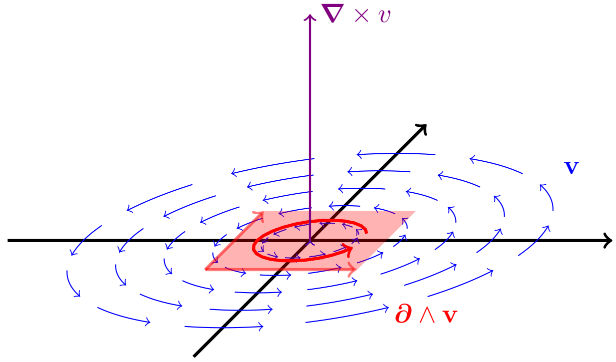


Figure 2.6: A sample vector field with its curl and exterior derivative.

These derivatives can be applied to any multivector field. The grade of the exterior derivative is one higher than the original field, and the grade of the interior derivative one lower.

When applied to a vector, the interior derivative does what the classical divergence does, and the exterior derivative does what the curl does. However, the curl was an axial vector - and the exterior derivative of a vector field is the bivector corresponding to that axial vector. Figure 2.6 illustrates this.

There is something important to note when taking the exterior and interior derivative of vector fields. Suppose we have an axial "vector" field $\mathbf{b}$, which corresponds to a bivector field $B = Ib$. Then, the curl $\nabla \times \mathbf{b}$ would translate to the *interior* derivative $\partial \cdot B$, and the divergence $\nabla \cdot \mathbf{b}$ would correspond to the *exterior* derivative $\partial \wedge B$.

Two successive applications of the interior derivative or the exterior derivative on any multivector are always zero:

Double interior and exterior derivatives

$$\partial \cdot (\partial \cdot M) = 0 \tag{2.147}$$
$$\partial \wedge \partial \wedge M = 0 \tag{2.148}$$

The second identity is quite easy to prove:

$$\partial \wedge \partial \wedge M = e^i \partial_i \wedge e^j \partial_j M \tag{2.149}$$
$$= \partial_i \partial_j (e^i \wedge e^j) M \tag{2.150}$$

The wedge product $e^i \wedge e^j$ is antisymmetric. However, partial derivatives commute - $\partial_i \partial_j = \partial_j \partial_i$. This means that the term is identically zero.

We can't quite prove the first identity yet without having to resort to a plethora of confusing $(-1)^{n(n-k)k}$ factors of the sort differential form advocates like to use to feel superior over us muggles. The good news is that we are going to develop the right tools for this proof in the next chapter, so stay tuned.

2.3.2 Identities involving the geometric derivative

The geometric derivative ∂ consists of the sum of the interior and exterior derivative:

$$\partial = \partial \cdot + \partial \wedge \tag{2.151}$$

It might seem like an unnecessary complication to sum them together. After all, don't we need the inner and outer derivative for two completely different purposes?

It turns out that this does indeed make sense - the geometric derivative $\boldsymbol{\partial}\phi$ "inherits" its associativity from the geometric product. Associativity gives us a lot more flexibility. For instance, suppose we have found a physical law saying that the geometric derivative of some multivector field M is always zero:

$$\boldsymbol{\partial} M = 0. \tag{2.152}$$

When we premultiply both sides with another $\boldsymbol{\partial}$, we can directly deduce:

$$\boldsymbol{\partial}^2 M = 0 \tag{2.153}$$

The operator $\boldsymbol{\partial}^2$ is a scalar. It is commonly known as the Laplacian $\Delta = \boldsymbol{\nabla}^2$. In traditional vector calculus, it would've been very difficult to deduce this identity at this level of generalization. This associative way of handling geometric derivatives will prove to be very useful in electrodynamics and quantum field theory.

The associativity is also very useful for some other purposes. For instance, let's try to decompose $\boldsymbol{\partial}^2$ into its constituents:

$$\begin{aligned} \boldsymbol{\partial}^2 M &= (\boldsymbol{\partial}\cdot + \boldsymbol{\partial}\wedge)(\boldsymbol{\partial}\cdot + \boldsymbol{\partial}\wedge)M & (2.154)\\ &= \boldsymbol{\partial}\cdot\boldsymbol{\partial}\cdot M + \boldsymbol{\partial}\cdot(\boldsymbol{\partial}\wedge M) + \boldsymbol{\partial}\wedge(\boldsymbol{\partial}\cdot M) + \boldsymbol{\partial}\wedge\boldsymbol{\partial}\wedge M & (2.155) \end{aligned}$$

The double-interior and double-exterior derivatives are zero:

$$\boldsymbol{\partial}^2 M = \boldsymbol{\partial}\cdot(\boldsymbol{\partial}\wedge M) + \boldsymbol{\partial}\wedge(\boldsymbol{\partial}\cdot M) \tag{2.156}$$

This is simple enough - a double geometric product decomposes into four parts, of which two are zero. Now let's assume that M is a vector field, i.e. $M = \mathbf{v}$. Then:

$$\boldsymbol{\partial}^2 \mathbf{v} = \boldsymbol{\partial}\cdot(\boldsymbol{\partial}\wedge\mathbf{v}) + \boldsymbol{\partial}\wedge(\boldsymbol{\partial}\cdot\mathbf{v}). \tag{2.157}$$

The rightmost term is the exterior derivative of the scalar $\boldsymbol{\partial}\cdot\mathbf{v}$. Scalars have grade 0, so the interior derivative of a scalar is always zero. Therefore, the exterior derivative of $\boldsymbol{\partial}\cdot\mathbf{v}$ is equal to its geometric derivative - aka its gradient, in conventional vector calculus. The exterior derivative $\boldsymbol{\partial}\wedge\mathbf{v}$ is a bivector, which would normally be an axial vector resulting from the curl. The subsequent interior derivative would also be translated as the curl, as it is being applied on a bivector field. Therefore, in conventional vector calculus, this identity reads:

$$\boldsymbol{\nabla}^2\mathbf{v} = -\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{v}) + \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{v}). \tag{2.158}$$

This identity is well-known from conventional vector calculus. However, our formulation (2.156) is both easier to derive and valid for all multivector grades instead of just vectors.

2.3.3 Product rule

The product rule is valid for the derivative. However, we have to be careful - the geometric derivative $\boldsymbol{\partial}$ is not commutative, so we cannot just write

$$\boldsymbol{\partial}(AB) = (\boldsymbol{\partial}A)B + A(\boldsymbol{\partial}B) \qquad \text{(INCORRECT!)} \tag{2.159}$$

Instead, we have do introduce **overdot notation**: In the second term in the above formula, $\boldsymbol{\partial}$ should still stand before A, but only act on B. We will denote this with an overdot over $\boldsymbol{\partial}$ and the part of the expression $\boldsymbol{\partial}$ is acting on:

Product rule for the geometric derivative

$$\boldsymbol{\partial}(AB) = \dot{\boldsymbol{\partial}}\dot{A}B + \dot{\boldsymbol{\partial}}A\dot{B}. \tag{2.160}$$

The dots mark which part of the multivector expression the operator is acting on. Written out, this would look like

$$e^i \partial_i (AB) = e^i (\partial_i A) B + e^i A \partial_i B. \tag{2.161}$$

The reason for this confusing change is simple: Previously, when we read an expression like $a\partial_i b$, we were used to the order indicating which part the differential acts on. But in multivector expressions, the order indicates the geometric product order - therefore, we need a new way to indicate the order of differentials, i.e. the overdots.

This might seem like a complication, but in fact, it gives us far greater liberty - for instance, we could also write down a "backwards derivative"

$$\dot{A}\dot{\boldsymbol{\partial}} = (\partial_i A) e^i. \tag{2.162}$$

This newfound freedom will prove especially useful in the area of geometric calculus. For instance, we have seen how we can write the interior and exterior product between vectors as the symmetric and antisymmetric part of the geometric product:

$$\mathbf{a} \cdot \mathbf{b} = \frac{1}{2} (\mathbf{ab} + \mathbf{ba}) \tag{2.163}$$

$$\mathbf{a} \wedge \mathbf{b} = \frac{1}{2} (\mathbf{ab} - \mathbf{ba}) \tag{2.164}$$

If we want to write the interior and exterior derivative this way, we need to use the overdot notation:

$$\boldsymbol{\partial} \cdot \mathbf{a} = \frac{1}{2} \left(\boldsymbol{\partial} \mathbf{a} + \dot{\mathbf{a}} \dot{\boldsymbol{\partial}} \right) \tag{2.165}$$

$$\boldsymbol{\partial} \wedge \mathbf{a} = \frac{1}{2} \left(\boldsymbol{\partial} \mathbf{a} - \dot{\mathbf{a}} \dot{\boldsymbol{\partial}} \right). \tag{2.166}$$

2.4 Pauli matrices

In quantum mechanics, we define the Pauli matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.167}$$

In practically all quantum mechanics courses, they just fall out of the sky. However, with geometric algebra, we're able to understand the understand the reasoning behind them. We first note that they all square to 1:

$$\sigma_1^2 = 1 \tag{2.168}$$

$$\sigma_2^2 = 1 \tag{2.169}$$

$$\sigma_3^2 = 1 \tag{2.170}$$

They anticommute with each other:

$$\sigma_i \sigma_j = -\sigma_j \sigma_i \qquad \text{for all } i \neq j. \tag{2.171}$$

This looks familiar - the basis vectors e_i of the geometric algebra of space behave exactly the same way. The matrix product between Pauli matrices takes the role of the geometric product. Mathematically, we say that the Pauli matrices σ_i are a **representation** of the geometric algebra of space Cl(3) - any calculation we can perform in Cl(3) can also be done with Pauli matrices if we make the substitutions

$$1 \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.172}$$

$$e_i \to \sigma_i \tag{2.173}$$

$$e_{ij} \to \sigma_i \sigma_j \tag{2.174}$$

$$e_{123} \to \sigma_{123} = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}. \tag{2.175}$$

Adding and multiplying these matrices yields the same result as adding and multiplying the underlying multivectors. We can even form matrix exponentials - for instance, a rotor generated by the rotation bivector

$$\Theta = \Theta^{12} e_{12} + \Theta^{23} e_{23} + \Theta^{31} e_{31} \tag{2.176}$$

would have the Pauli matrix representation

$$R = \exp\left(\Theta^{12}\sigma_1\sigma_2 + \Theta^{23}\sigma_2\sigma_3 + \Theta^{31}\sigma_3\sigma_1\right) \tag{2.177}$$

$$= \exp\begin{pmatrix} i\Theta^{12} & i\Theta^{23} + \Theta^{31} \\ i\Theta^{23} - \Theta^{31} & -i\Theta^{12} \end{pmatrix} \tag{2.178}$$

The representation of the reverse $\tilde{M}$ of a multivector M is just the hermitean conjugate of its representation $M^\dagger$. It leaves the scalar and the basis vectors invariant, but it flips around the order of the matrix product. This is exactly what we want.

The matrix representation of a vector $\mathbf{v}$ would be:

$$\mathbf{v} = v^i \sigma_i = \begin{pmatrix} v^3 & v^1 - iv^2 \\ v^1 + iv^2 & -v^3 \end{pmatrix} \tag{2.179}$$

You can check for yourself that the matrix representation of the rotor law,

$$\mathbf{v}' = R^\dagger \mathbf{v} R \tag{2.180}$$

(reading all the symbols as matrices representing multivectors) yields the same result as the proper multivector equation

$$\mathbf{v}' = \tilde{R}\mathbf{v}R. \tag{2.181}$$

All of this might come as a bit of an epiphany - we have been doing three-dimensional geometric algebra with the Pauli matrices all the time! However, at the same time, we have left some important questions unanswered: The main application of Pauli matrices are spinors (aka spin states). How do they fit into all of this? What does it mean to write

$$s^i = \langle\psi|\, \sigma^i \,|\psi\rangle\,, \tag{2.182}$$

for instance? Stay tuned - we are going to answer these questions in the chapter about spinors.

Chapter 3

Tensor notation

> A tensor is something that transforms like a tensor
>
> Ancient theoretical physics proverb

Chapter summary

- Tensor notation is not the same as matrix representation.
- We can denote k-vectors as fully antisymmetric rank k tensors. Every fully antisymmetric rank k tensor is a bivector.
- Multivectors can't be denoted that way because the rank of a tensor is fixed.
- We can write the wedge product by using the antisymmetrization brackets $[\cdots]$ on tensors.
- We can write the interior product by forming sandwich contractions between tensors.
- The antisymmetric rank-2 tensors representing bivectors can be written as matrices. If we exponentiate such a matrix, we get the rotation matrix corresponding to the rotor generated by the bivector.

3.1 Overview

Geometric algebra is a very nice tool, but it has its limits. Sometimes, it can be preferable to use conventional tools like matrices or tensors to do a calculation. In addition, practically all physics literature right now is written in conventional maths. Therefore, it would be nice to have the tools to fluently switch between tensor notation and geometric algebra notation.

3.1.1 Tensor notation is not the same as matrix representation

First of all, a warning: Previously, we have introduced the Pauli matrix representation for Cl(3).

> However, matrix representation and tensor notation are two completely different things.

In particular:

- A matrix representation maps every multivector onto a matrix. However, tensor notation can only depict pure k-vectors.

- In matrix representation, we always work with matrices. In tensor representation, we work with tensors of arbitrary rank - and the rank-2 ones just happen to be matrices, too.
- In matrix representation, we can multiply two matrices to get the geometric product between the two multivectors. In tensor representation, we cannot denote the geometric product.
- In tensor notation, the indices of our tensors are vector indices - they run over basis vectors. In matrix notation, the indices of our tensors are spinor indices - they run over basis spinors. You will find out what that means in the spinor chapter.

3.1.2 Fully antisymmetric rank k tensors

The fundamental principle of converting geometric algebra experssions to tensors is the basis representation of k-vectors. We already know that we can write an arbitrary k-vector as a linear combination of basis k-vectors:

Tensor notation for a k-vector

$$X = \frac{1}{k!} X^{i_1 i_2 \cdots i_k} \quad e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k} \tag{3.1}$$

The collection of components $X^{i_1 i_2 \cdots i_k}$ is a **rank-k tensor**. It is **fully antisymmetric**. This means that it flips sign if we exchange any two indices:

$$X^{i_1 i_2 \cdots i_k} = -X^{i_2 i_1 \cdots i_k}. \tag{3.2}$$

This is because the basis bivectors also flip sign when we exchange two indices because of the wedge product antisymmetry. This antisymmetry implies that all indices $i_1, i_2, \cdots, i_k$ need to be distinct from each other - if there is a duplicate index, the component is zero:

$$X^{i_1 i_1 \cdots i_k} = -X^{i_1 i_1 \cdots i_k} = 0 \tag{3.3}$$

This equivalence always holds: Fully antisymmetric rank-k tensors correspond to k-vectors, and vice versa.

$$\text{Fully antisymmetric rank-}k\text{ tensors} \qquad \leftrightarrow \qquad k\text{-vectors}. \tag{3.4}$$

3.1.3 Vectors, bivectors and trivectors

The easiest example is the vector - we can write any arbitrary vector $\mathbf{v}$ as:

$$\mathbf{v} = v^i e_i. \tag{3.5}$$

Remember that vectors v^i are rank-1 tensors. Similarly, a bivector B can be written as

$$B = \frac{1}{2!} B^{ij} e_i \wedge e_j = \left(B^{12} e_{12} + B^{23} e_{23} + B^{31} e_{31}\right). \tag{3.6}$$

For bivectors only, we can also denote their components in matrix form:

$$B^{ij} = \begin{pmatrix} B^{11} & B^{12} & B^{13} \\ B^{21} & B^{22} & B^{23} \\ B^{31} & B^{32} & B^{33} \end{pmatrix} \tag{3.7}$$

$$= \begin{pmatrix} 0 & B^{12} & -B^{31} \\ -B^{12} & 0 & B^{23} \\ B^{31} & -B^{23} & 0 \end{pmatrix} \tag{3.8}$$

You are probably wondering why we used a factor of $1/k!$ to define the basis decomposition. The reason is simple - take the above bivector decomposition with its $1/2!$ factor as an example. The Einstein summation convention in this formula will sweep across both cyclic and anticyclic indices - for instance, somewhere in the sum, ij would assume the value 23, and then 32 a bit later. Because $e_i \wedge e_j = -e_j \wedge e_i$, we needed to define

$$B^{ij} = -B^{ji}. \tag{3.9}$$

Then, the 23 and 32 parts of the sum would resolve to:

$$B^{23}e_{23} + B^{32}e_{32} = 2B^{23}e_{23} \tag{3.10}$$

We see that by summing over all possible combinations of ij, we "overcount" the basis bivectors twice, because there are two possible permutations for every ij index pair. Similarly, there are $6 = 3!$ permutations for every index triple ijk, so in order to write down a trivector in terms of its components, we'd need to write

$$T = \frac{1}{3!}T^{ijk}e_i \wedge e_j \wedge e_k \tag{3.11}$$

in order to get the normalization right. The expression T^{ijk} might raise some eyebrows - after all, we are working in three dimensions, so all possible index triples ijk without duplicates are permutations of 123. The even permutations of 123 will all be equal to some $t \in \mathbb{R}$, and the odd permutations will be -1 times the even permutations:

$$T^{ijk} = \begin{cases} t & \text{if } ijk \text{ is an even permutation of 123} \\ -t & \text{if } ijk \text{ is an odd permutation of 123} \\ 0 & \text{if } ijk \text{ has duplicate indices.} \end{cases} \tag{3.12}$$

.

This should look somewhat familiar - the Levi-Civita symbol in three dimensions, ϵ^{ijk} is defined as:

$$\epsilon^{ijk} = \begin{cases} 1 & \text{if } ijk \text{ is an even permutation of 123} \\ -1 & \text{if } ijk \text{ is an odd permutation of 123} \\ 0 & \text{if } ijk \text{ is not a permutation of 123.} \end{cases} \tag{3.13}$$

. We can conclude that the tensor T^{ijk} denoting a trivector T in three dimensions must always be a multiple of ϵ^{ijk}:

$$T^{ijk} = t\epsilon^{ijk} \tag{3.14}$$

The Levi-Civita symbol itself denotes the unit pseudoscalar:

$$\frac{1}{3!}\epsilon^{ijk}e_i \wedge e_j \wedge e_k = \frac{1}{3!}3!\; e_{123} = e_{123}. \tag{3.15}$$

Therefore, it is absolutely valid to refer to the Levi-Civita symbol as the **pseudoscalar symbol**.

In geometric algebra, we always work with an orthonormal metric. In the three-dimensional case, our metric is just δ_{ij}, so we can pull indices up and down however we please. We could therefore also write B as

$$B = B_{ij}e^i \wedge e^j. \tag{3.16}$$

This will change though as soon as we treat the spacetime algebra, where the metric $\eta_{\mu\nu}$ is nontrivial.

3.1.4 Extracting components

Now, we should investigate how to extract the tensor components of a given k-vector algebraically. For vectors $\mathbf{v}$, this is simple: We just take the interior product between the vector and the basis vector in question:

$$v^i = \mathbf{v} \cdot e_i = \langle \mathbf{v} e_i \rangle . \tag{3.17}$$

Intuitively, we might expect that this extends to arbitrary k-vectors. For instance, for bivectors, we might try to write

$$B^{ij} = B \cdot (e^i \wedge e^j) = \left\langle B(e^i \wedge e^j) \right\rangle \qquad \text{(INCORRECT!)}. \tag{3.18}$$

But if we do this, we run into a problem: The bivectors e^{ij} square to -1, so for instance, for a bivector $B = e_{12}$, we'd obtain the components $B^{12} = -B^{21} = e_{12} \cdot e^{12} = -1$, the opposite of what we want.

The solution is to dot the bivector with the *inverse* of the basis bivectors:

$$(e_{ij})^{-1} = -e^{ij}. \tag{3.19}$$

Note that the inverse operation $(\cdots)^{-1}$ implicitly swaps upstairs and downstairs indices.

Hence, we obtain the correct tensor components by writing

$$B^{ij} = B \cdot (e_{ij})^{-1} = \left\langle B(e_i \wedge e_j)^{-1} \right\rangle \tag{3.20}$$

The generalized formula for the components of a k-vector X of arbitrary grade is:

Tensor components of a k-vector

$$X^{i_1 i_2 \cdot i_k} = \left\langle X \; (e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k})^{-1} \right\rangle . \tag{3.21}$$

For 3D vectors (3.17), the fact that we had to use the inverse $\cdots^{-1}$ does not matter - for every i, we have $e_i^2 = 1$ and thus $(e_i)^{-1} = e_i$. However, it will matter as soon as we treat special relativity and spacetime algebra.

3.1.5 Some examples

Whenever we see a fully antisymmetric tensor in conventional math, it is safe to translate it to a k-vector. Common examples include:

- T_1, T_2, T_3, the generator matrices of SO(3): Matrices are rank-2 tensors, so these are actually the bivectors e_{23}, e_{31} and e_{12}.
- As explained above, the Levi-Civita symbol ϵ actually is the unit pseudoscalar.
- The electromagnetic field tensor $F^{\mu\nu}$ actually is a bivector in the geometric algebra of spacetime - more on that in the Electrodynamics chapter.
- The Riemann tensor $R_{\mu\nu\rho\sigma}$ from general relativity is antisymmetric in its first two and in its last two indices. Thus, it is a linear map from bivectors to bivectors.

3.2 Operations in tensor notation

Now that we have seen how we can write k-vectors as tensors, the next logical step is to ask ourselves how to perform operations like the addition on them.

3.2.1 Linear operations

Linear operations on k-vectors in tensor notation are pretty simple and work as we'd intuitively expect. For instance, if we want to add two k-vectors, we just need to add their tensor notations together componentwise:

$$(B + C)^{ij} = B^{ij} + C^{ij}. \tag{3.22}$$

Scalar multiplication works similarly easy:

$$(\lambda B)^{ij} = \lambda B^{ij}. \tag{3.23}$$

However, you may have already noticed the important caveat here: In tensor notation, we can only add k-vectors of even grade. It is not possible to denote mixed-grade multivectors in tensor notation. Hence, it is not possible to denote the geometric product - it inherently produces mixed-grade multivectors. It is possible though to denote the interior and exterior product in tensor notation. But first, we will have to introduce some new notation.

3.2.2 Symmetric and antisymmetric parts

Let $Z^{i_1 i_2 \cdots i_k}$ be a rank-k tensor that is not necessarily symmetric or antisymmetric. First of all, we define the **symmetrization** of Z as:

Symmetrization of a tensor

$$Z^{(i_1 i_2 \cdots i_k)} = \frac{1}{k!} \left(Z^{i_1 i_2 \cdots i_k} + \text{all permutations of } i_1 i_2 \cdots i_k. \right) \tag{3.24}$$

We are averaging over all possible permutations. The resulting tensor is completely symmetric, as all antisymmetries have been "leveled out" by the sum over all permutations. If we were two swap two indices, the result would be exactly the same. For instance, the symmetrization of a rank-3 tensor Z^{ijk} would read

$$Z^{(ijk)} = \frac{1}{3!} \left(Z^{ijk} + Z^{kij} + Z^{jki} + Z^{kji} + Z^{jik} + Z^{ikj} \right). \tag{3.25}$$

In contrast, the **antisymmetrization** of a tensor $Z^{i_1 i_2 \cdots i_k}$ is defined as:

Antisymmetrization of a tensor

$$Z^{[i_1 i_2 \cdots i_k]} = \frac{1}{k!} \left(Z^{i_1 i_2 \cdots i_k} + \text{all other even permutations} - \text{all odd permutations} \right) \tag{3.26}$$

Now, we are not simply taking the average over all possible permutations - we are adding the even ones and subtracting the odd ones. This means that $Z^{[i_1 i_2 \cdots i_k]}$ is fully antisymmetric - whenever we exchange two indices, the perviously even permutations will become odd, and the previously odd permutations will become even. We can say that we have "averaged out" all symmetric parts of the tensor. For instance, the antisymmetrization of a rank-3 tensor would read

$$Z^{[ijk]} = \frac{1}{3!} \left(Z^{ijk} + Z^{kij} + Z^{jki} - Z^{kji} - Z^{ikj} - Z^{jik} \right) \tag{3.27}$$

If the tensor denotes a k-vector (i.e. if the tensor already is fully antisymmetric), the antisymmetrization operation will have no effect on the tensor. For instance, for a trivector T^{ijk}:

$$T^{[ijk]} = T^{ijk}. \tag{3.28}$$

3.2.3 Interior product

Now, we are equipped to understand the expressions for the interior and exterior product. We start by investigating the interior product. We already know the simplest case - the interior product between two vectors v, w:

$$v \cdot w = v^i w_i. \tag{3.29}$$

We simply form the **contraction** of their indices. Similarly, for bivectors and trivectors,

$$(B \cdot v)^i = B^{ij} v_j \tag{3.30}$$
$$(T \cdot v)^{ij} = T^{ijk} v_k, \tag{3.31}$$

or from the other side:

$$(v \cdot B)^i = v_j B^{ji} \tag{3.32}$$
$$(v \cdot T)^{ij} = v_k B^{kij} \tag{3.33}$$

The general formula involving a k-vector X and a vector v is:

Interior product between a k-vector and a vector

$$(X \cdot \mathbf{v})^{i_1 i_2 \cdots i_{k-1}} = X^{i_1 i_2 \cdots i_{k-1} j} v_j \tag{3.34}$$
$$(\mathbf{v} \cdot X)^{i_1 i_2 \cdots i_{k-1}} = v_j X^{j i_1 i_2 \cdots i_{k-1}}. \tag{3.35}$$

Now, we can ask ourselves how we can dot together k-vectors of arbitrary rank. We already know that if the k-vectors are in basis form, we can rewrite the inner product between them as a series of inner products of k-vectors with vectors. For instance, the inner product $e_{123} \cdot e_{23}$:

$$e_{123} \cdot e_{23} = (e_{123} \cdot e_2) \cdot e_3 \tag{3.36}$$

We are free to reorder the k-vector we are splitting up. For instance:

$$e_{123} \cdot e_{23} = e_{123} \cdot (-e_{32}) = -(e_{123} \cdot e_3) \cdot e_2. \tag{3.37}$$

Thus, to calculate such a interior product, we could naïvely try writing down a **sandwich contraction** like

$$(T \cdot B)^i = T^{ijk} B_{kj}. \qquad \text{(INCORRECT!)} \tag{3.38}$$

We run into an issue here. When we work in geometric algebra notation like in (3.36) and (3.37), we choose a *specific* ordering for the vectors we are dotting onto the trivector. However, in (3.38), we sum over *all* index pairs jk, so we include *all* possible orderings of the bivector. Hence, the correct formula has to divide by $l!$, where l is the number of dummy indices. The previous example therefore correctly translates to

$$(T \cdot B)^i = \frac{1}{2!} T^{ijk} B_{kj}. \tag{3.39}$$

Let's consider the general interior product between a k-vector X and a l-vector Y, where $k > l$. The resulting grade is $k - l$. The general sandwich contraction formula is:

Interior product between a k-vector and an l-vector

$$(X \cdot Y)^{i_1 \cdots i_{k-l}} = \frac{1}{l!} X^{i_1 \cdots i_{k-l} j_1 j_2 j_3 j_4 \cdots j_l} Y_{j_l \cdots j_4 j_3 j_2 j_1}. \tag{3.40}$$

Proof of double interior derivative identity

Now, we are ready to prove the identity

$$\boldsymbol{\partial} \cdot (\boldsymbol{\partial} \cdot M) = 0. \tag{3.41}$$

First of all, because the double interior derivative is linear, we can assume that M is a pure k-vector without loss of generality. We can thus denote M as a tensor $M^{i_1 i_2 \cdots i_k}$. The geometric derivative $\boldsymbol{\partial}$ is denoted by the rank-1 tensor ∂^i. The first inner product is a $k-1$ vector:

$$(\boldsymbol{\partial} \cdot M)^{i_2 i_3 \cdots i_k} = \partial_{i_1} M^{i_1 i_2 \cdots i_k}. \tag{3.42}$$

The second one can be sandwiched on top of the first one:

$$(\boldsymbol{\partial} \cdot (\boldsymbol{\partial} \cdot M))^{i_3 \cdots i_k} = \partial_{i_2} \partial_{i_1} M^{i_1 i_2 i_3 \cdots i_k}. \tag{3.43}$$

M is antisymmetric over all indices. However, the two partial derivatives commute with each other. Thus, the index contraction is equal to zero:

$$\partial_{i_2} \partial_{i_1} M^{i_1 i_2 \cdots} = \frac{1}{2} \partial_{i_2} \partial_{i_1} \left(M^{i_1 i_2} - M^{i_2 i_1} \right) \tag{3.44}$$

$$= \frac{1}{2} \left(\partial_{i_2} \partial_{i_1} M^{i_1 i_2} - \partial_{i_1} \partial_{i_2} M^{i_2 i_1} \right) \tag{3.45}$$

We now rename the indices in the second term - we can do that because they are dummy indices:

$$= \frac{1}{2} \left(\partial_{i_2} \partial_{i_1} M^{i_1 i_2} - \partial_{i_2} \partial_{i_1} M^{i_1 i_2} \right) \tag{3.46}$$

$$= 0. \tag{3.47}$$

Thus, we have proven the identity $\boldsymbol{\partial} \cdot (\boldsymbol{\partial} \cdot M) = 0$. Also, in general, we can remember that if we contract something symmetric with something antisymmetric, the result will always be zero.

3.2.4 Exterior product

Let's say we want to take the exterior product between the vectors $\mathbf{v} = e_1$ and $\mathbf{w} = e_2$, resulting in the bivector $\mathbf{v} \wedge \mathbf{w} = e_{12}$. The only nonzero components of the tensors v^i and w^i are $v^1 = 1$ and $w^2 = 1$, respectively, and the only two nonzero components of $(\mathbf{v} \wedge \mathbf{w})^{ij}$ are $(\mathbf{v} \wedge \mathbf{w})^{12} = -(\mathbf{v} \wedge \mathbf{w})^{21} = 1$. Therefore, our first impulse might be to write something like

$$(\mathbf{v} \wedge \mathbf{w})^{ij} = v^i w^j \qquad \text{(INCORRECT!)} \tag{3.48}$$

This tensor would have the component $v^1 w^2 = 1$, but the other one is missing, $v^2 w^1 = 0$. It is not antisymmetric. We could therefore try to antisymmetrize the tensor:

$$(\mathbf{v} \wedge \mathbf{w})^{ij} = v^{[i} w^{j]} \qquad \text{(INCORRECT!)} \tag{3.49}$$

However, the antisymmetrization brackets $[\cdots]$ *average* over the components, so the above tensor has the components

$$v^{[1} w^{2]} = v^{[2} w^{1]} = \frac{1}{2} \tag{3.50}$$

half of what we want. This is because the component $v^1 w^2$ has been "smeared out" over the $2! = 2$ index pairs 12 and 21. We therefore need to premultiply this expression with $2!$. We obtain the correct result,

$$(\mathbf{v} \wedge \mathbf{w})^{ij} = 2! \; v^{[i} w^{j]}. \tag{3.51}$$

This generalizes to arbitrarily long exterior products of k-vectors. For instance:

$$(\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c})^{ijk} = 3!\ a^{[i} b^j c^{k]}. \tag{3.52}$$

The general formula for the exterior product of l vectors $\mathbf{a}_{(1)}, \cdots, \mathbf{a}_{(l)}$ is:

$$(\mathbf{a}_{(1)} \wedge \cdots \wedge \mathbf{a}_{(l)})^{i_1 i_2 \cdots i_l} = l!\ a_{(1)}^{[i_1} a_{(2)}^{i_2} \cdots a_{(l)}^{i_l]}. \tag{3.53}$$

The next thing to ask ourselves is how to calculate the exterior product between a k-vector X and an l-vector Y. Looking at (3.53), we could naively try to write

$$(X \wedge Y)^{i_1 \cdots i_{k+l}} = (k+l)!\ X^{[i_1 \cdots i_k} Y^{i_{k+1} \cdots i_{k+l}]} \qquad \text{(INCORRECT!)} \tag{3.54}$$

But then, we run into a similar problem as we did in (3.38). The formula (3.53) refers to the exterior product of a *specific ordering* of vectors. However, a k-vector like $X^{i_1 \cdots i_k}$ contains *all* $k!$ *possible* orderings of the cross product. The formula (3.54) would therefore overcount the product by $k!$ times $l!$ times. We can thus see that the correct formula is:

Exterior product betweeen a k-vector and an l-vector

$$(X \wedge Y)^{i_1 \cdots i_{k+l}} = \frac{(k+l)!}{k!\ l!}\ X^{[i_1 \cdots i_k} Y^{i_{k+1} \cdots i_{k+l}]} \tag{3.55}$$

An important special case is the exterior product between a vector and a k-vector:

Exterior product between a vector and a k-vector

$$(\mathbf{a} \wedge X)^{i_1 i_2 \cdots i_{k+1}} = (k+1)\ a^{[i_1} X^{i_2 \cdots i_{k+1}]} \tag{3.56}$$

$$(X \wedge \mathbf{a})^{i_1 i_2 \cdots i_{k+1}} = (k+1)\ X^{[i_1 \cdots i_k} a^{i_{k+1}]} \tag{3.57}$$

3.3 Differential forms

Almost every one of us will already have encountered **differential forms** and **exterior algebra**. Geometric algebra and exterior algebra are very similar in principle, but there are some important differences. Broadly speaking, differential forms put a focus on mathematical naturalness and abstract beauty, while geometric algebra puts a focus on practicability and geometric intuition. In other words, exterior algebra is for physics gods, and geometric algebra is for mortals like us.

The most important difference is that there is no metric or interior product in exterior algebra - it is not possible to pull indices up and down as we please. Because we have a metric in GA, it does not matter whether our vectors are covariant or contravariant. In exterior algebra, however, we always work with **covariant vectors**. The basis covariant vectors, also called **one-forms**, are denoted

$$\mathrm{d}x^1, \ \ \mathrm{d}x^2, \ \ \cdots, \ \ \mathrm{d}x^n. \tag{3.58}$$

In our notation, they are just the basis vectors with raised indices:

$$e^1, \ \ e^2, \ \ \cdots, \ \ e^n. \tag{3.59}$$

In exterior algebra, k-vectors are called k-forms. For instance, the bivector

$$B = e_1 \wedge e_2 \tag{3.60}$$

is equivalent to the 2-form

$$B = \mathrm{d}x^1 \wedge \mathrm{d}x^2. \tag{3.61}$$

In differential geometry, pseudoscalars are called **top-forms**. The exterior derivative $\partial \wedge X$ is denoted

$$\mathrm{d}X \tag{3.62}$$

in exterior algebra. The commonly given definition for $\mathrm{d}X$,

$$\mathrm{d}X^{i_1 i_2 \cdots i_{k+1}} = (k+1)\, \partial^{[i_1} X^{i_2 \cdots i_{k+1}]}, \tag{3.63}$$

is just the tensor notation for the expression $\partial \wedge X$.

The greatest problem with differential forms is their **lack of an interior product**, and, as a consequence, a lack of an interior derivative. To remedy this, exterior algebra defines the so-called **Hodge dual** $*X$. It turns a k-vector X into an $n-k$-vector $*X$, where n is the number of dimensions. Its definiton is:

$$(*X)_{i_1 i_2 \cdot i_{n-k}} = \frac{1}{k!} \epsilon_{i_1 i_2 \cdots i_{n-k}}{}^{i_{n-k+1} i_{n-k+2} \cdots i_n} X_{i_{n-k+1} i_{n-k+2} \cdots i_n}. \tag{3.64}$$

This definition seems pretty intimidating, and the standard way of teaching it is completely devoid of any geometric intuition. This is part of the problem - there are good abstract mathematical reasons to define the Hodge dual the way we do, but they are only accessible to the initiates of expert differential geometry. Everyone else mostly just learns this definition by heart and hopes to never have to use it again.

Luckily, we have geometric algebra now. Let us take a step back and think about which parts of this definition we know already. The pseudoscalar symbol ϵ is the tensor notation for the pseudoscalar I, so it looks a bit like we're taking the interior product between the pseudoscalar I and the k-vector X. But that's not quite it yet - the index contractions in (3.64) are not in sandwich order like in (3.55).

But luckily, X is totally antisymmetric, so we can rearrange them into sandwich order. The first index of X we are pulling to the left will take $(k-1)$ permutations. The next one will take $(k-2)$ permutations. This goes on until we have performed

$$(k-1) + (k-2) + \cdots + 2 + 1 = \frac{k(k-1)}{2} \tag{3.65}$$

permutations, and the expression for the Hodge dual reads:

$$(*X)_{i_1 i_2 \cdot i_{n-k}} = (-1)^{\frac{k(k-1)}{2}} \frac{1}{k!} \epsilon_{i_1 i_2 \cdots i_{n-k}}{}^{i_{n-k+1} i_{n-k+2} \cdots i_n} X_{i_n \cdots i_{n-k+2} i_{n-k+1}}. \tag{3.66}$$

This looks exactly like the interior product in tensor notation (with lowered indices). We can now translate the definition of the Hodge dual of a k-vector into GA:

Hodge dual in GA

$$*X = (-1)^{\frac{k(k-1)}{2}} I X. \tag{3.67}$$

If we are careful with the signs, we can construct a kind of "interior product surrogate" with the Hodge dual, as we will see in the electrodynamics chapter when translating the differential form electrodynamic equations. However, in general, we have no reason to use the Hodge dual - everything gets a lot easier with a metric.

3.4 Transformation matrices

In the last chapter, we have seen how we can describe rotations with rotation bivectors Θ, rotors

$$R = \exp(\Theta/2), \tag{3.68}$$

and the rotor law

$$\mathbf{v}' = \tilde{R}\mathbf{v}R. \tag{3.69}$$

The normal formalism is somewhat different. Instead of rotors, we use rotations matrices $M = M^i{}_j$. They are linear maps mapping vectors to other vectors:

$$v'^i = M^i{}_j v^j. \tag{3.70}$$

Those matrices can also be written as matrix exponentials:

$$M = \exp\big(T_i\theta^i\big), \tag{3.71}$$

where

$$T_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \tag{3.72}$$

$$T_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{3.73}$$

$$T_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.74}$$

are the **generators** of SO(3) and θ^i are the components of the rotation "vector". For a specific θ^i, this results in the generator matrix

$$T_i\theta^i = \begin{pmatrix} 0 & \theta^3 & -\theta^2 \\ -\theta^3 & 0 & \theta^1 \\ \theta^2 & -\theta^1 & 0 \end{pmatrix} \tag{3.75}$$

Now, let's try to make sense of this with geometric algebra. The first thing we note is that the rotation axial vector is actually a bivector $\Theta = I\theta$:

$$\Theta = \theta^3 e_{12} + \theta^1 e_{23} + \theta^2 e_{31} \tag{3.76}$$

Its tensor components are:

$$\Theta^{12} = -\Theta^{21} = \theta^3 \tag{3.77}$$

$$\Theta^{23} = -\Theta^{32} = \theta^1 \tag{3.78}$$

$$\Theta^{31} = -\Theta^{13} = \theta^2 \tag{3.79}$$

So we would rewrite the above matrix as

$$T_i\theta^i = \begin{pmatrix} 0 & \Theta^{12} & -\Theta^{31} \\ -\Theta^{12} & 0 & \Theta^{23} \\ \Theta^{31} & -\Theta^{23} & 0 \end{pmatrix} = \Theta^{ij}. \tag{3.80}$$

This is just the tensor notation for the rotation bivector Θ! We can see that the SO(3) generator matrices T_i and the contraction $T_i\theta^i$ are just a really roundabout way to denote a bivector without calling it a bivector. Then, the rotation matrix is the matrix exponential of the tensor notation of the rotation bivector.

We can also do this the other way around: We can obtain the rotation matrix $M^i{}_j$ from a rotor by calculating

Rotation matrix corresponding to a rotor

$$M^i{}_j = e_i \cdot (\tilde{R} e_j R). \tag{3.81}$$

We will rarely need this formula in practice - after all, we want to stop using rotation matrices. It is still useful though to learn about how rotation matrices interlink with geometric algebra.

When we apply a rotation matrix M to a vector $\mathbf{v}$, it transforms like

$$\mathbf{v}' = M\mathbf{v}, \tag{3.82}$$

or in index notation:

$$v'^i = M^i{}_j v^j. \tag{3.83}$$

Vectors are rank-1 tensors. The above transformation law is a special case of the tensor transformation law. In general, rank (k, l) tensors transform like

Tensor transformation law

$$T'^{i_1 i_2 \cdots i_k}{}_{i_{k+1} \cdots i_{k+l}} = M^{i_1}{}_{j_1} M^{i_2}{}_{j_2} \cdots M^{i_k}{}_{j_k} \left(M^{-1}\right)_{i_{k+1}}{}^{j_{k+1}} \cdots \left(M^{-1}\right)_{i_{k+l}}{}^{j_{k+l}} \; T'^{j_1 j_2 \cdots j_k}{}_{j_{k+1} \cdots j_{k+l}}, \tag{3.84}$$

one rotation matrix for every index. Upstairs (contravariant) indices transform with the normal rotation matrix, and downstairs (contravariant) indices transform with the inverse rotation matrix. For instance, if we were to rotate a trivector T,

$$T' = \tilde{R} T R \tag{3.85}$$

would transform like

$$T'^{ijk} = M^i{}_l M^j{}_m M^k{}_n T^{lmn}, \tag{3.86}$$

the tensor transformation law for a rank (3, 0) tensor. However, if we pulled down one of the indices ($T^{ij}{}_k$), we'd have use the inverse for it:

$$T'^{ij}{}_k = M^i{}_l M^j{}_m \left(M^{-1}\right)_k{}^n T^{lm}{}_n. \tag{3.87}$$

Chapter 4

Geometric algebra of spacetime

> There is quite a large number of quotes falsely attributed to me. Sadly, people misattribute quotes to me almost as often as they misattribute quotes to Albert Einstein.
>
> Confucius

Chapter summary

- Special relativity unifies 1D time and 3D space into 4D spacetime.
- Four-vectors are constructed by combining one timelike quantity and one spacelike quantity.
- The geometric algebra of spacetime, or spacetime algebra (STA) consists of four basis vectors γ_μ. The timelike basis vector γ_0 squares to $(\gamma_0)^2 = +1$, and the spacelike basis vectors γ_i to $(\gamma_i)^2 = -1$.
- Lorentz boosts are just a special kind of rotation - hyperbolic rotations. In this context, we call normal rotations "circular rotations".
- The Dirac gamma matrices are a matrix representation of the spacetime algebra.
- The acceleration four-vector is always orthogonal to the velocity four-vector. We can use the bivector $\mathbf{a} = AU$ to handle the resulting mathematical difficulties.
- Relativistic angular momentum decomposes into circular momentum (what we're used to), and hyperbolic momentum (a new kind of angular momentum).
- The space-time split links together non-relativistic and relativistic physics.

4.1 Basics

4.1.1 Foundations of special relativity

Special relativity is often perceived as an arcane science only accessible to geniuses - and it doesn't exactly help that many physics curriculae prioritize other topics over relativity, for instance electromagnetism or quantum mechanics. We, however, want to show the reader that relativity doesn't need to be feared at all. Even though the historical development of relativity can be very complex and confusing, it becomes very simple to grasp its axiomatic foundations if one uses geometric algebra.

The basic axioms of special relativity can be summarized as follows:

- In addition to the three usual spatial dimensions x, y and z, we add in time t as the fourth (or zeroth) dimension. Vectors now do not consist of three, but of four elements.
- Normally, the three spatial unit vectors square to $+1$. We redefine them to square to -1. The unit vector pointing in the time direction squares to $+1$.

Based on these axioms, we are going to build the so-called spacetime algebra (STA). It is okay if you do not understand them yet - we are going to explain them in the next few subsections.

4.1.2 Four-vectors

Previously, the basis vectors of our geometric algebra space were

$$e_1, e_2, e_3.$$

Now, to construct the geometric algebra of spacetime, we add in a **zeroth basis vector** to represent **time**:

$$\gamma_0, \gamma_1, \gamma_2, \gamma_3.$$

An arbitrary four-vector x^μ can be written as

$$x = x^\mu \gamma_\mu \tag{4.1}$$

in this formalism, analogously to the three-dimensional space algebra.

However, this comes with a catch. In the space algebra, all the basis vectors square to one ($e_i^2 = 1$), but in spacetime algebra, we define

STA basis vector squares

$$(\gamma_0)^2 = 1 \tag{4.2}$$

$$(\gamma_1)^2 = (\gamma_2)^2 = (\gamma_3)^2 = -1. \tag{4.3}$$

This means that an arbitrary vector $a = a^\mu \gamma_\mu$ will square to:

$$a^2 = (a^0)^2(\gamma^0)^2 + (a^1)^2(\gamma_1)^2 + (a^2)^2(\gamma_2)^2 + (a^3)^2(\gamma_3)^2 \tag{4.4}$$

$$= (a^0)^2 - \sum_i (a^i)^2. \tag{4.5}$$

In space algebra, non-zero vectors always square to positive values ($a^2 > 0$). In STA, however, we have to distinguish between three cases:

- $a^2 > 0$: In this case, a is called a **timelike** vector. In this case, $T = \sqrt{a^2}$ is called the **proper time** of a[1]

[1] Proper time is normally called τ in the literature, but since we use $\tau = 2\pi$, we call it to T to avoid confusion.

- $a^2 < 0$: In this case, a is called a **spacelike** vector. In this case, $s = \sqrt{-a^2}$ is called **spacelike separation**.
- $a^2 = 0$: In this case, a is called a **lightlike** or **null** vector.

4.1.3 Natural units

The most important example of a four-vector is the time-position vector x. It describes the time and location of an event - its time component specifies at what time an even takes place, and its three space component specifies the location of the event:

$$x = t\gamma_0 + x^i\gamma_i \tag{4.6}$$
$$= x^\mu\gamma_\mu. \tag{4.7}$$

We will also write four-vectors as a combination of a time component and a spatial vector:

Time-position vector

$$x = (t, \mathbf{x}) \tag{4.8}$$

For instance, to describe the first nuclear explosion in history, we would write:

$$x = (\text{July 16, 1945}; \textbf{Trinity Site, New Mexico, Earth}). \tag{4.9}$$

However, please don't picture this as a combination of two quantities, but rather as a single point in a four-dimensional space. The way we humans have evolved has led us to perceive time and space fundamentally differently, but the core point of special relativity is that there is not that much of a difference - they are both dimensions of spacetime.

This should raise some eyebrows. The time component has the unit "second", while the space component has the unit "meters". How are we supposed to add meters onto seconds? The answer is simple: In special relativity, meters and seconds are just two different units for the same quantity - "time-length", which measures distances in spacetime. The difference between seconds and meters really just is like the difference between meter and feet. Just like there is a conversion factor to convert feet into meters,

$$\frac{1 \text{ foot}}{1 \text{ meter}} = 0.3048 \tag{4.10}$$

there is a conversion factor to convert seconds and meters into each other:

Natural units

$$\frac{1 \text{ second}}{1 \text{ meter}} = 299\ 792\ 458 \tag{4.11}$$

This number is better known as the speed of light. The reason for this is simple: In special relativity, the speed of light is 1 by definition. When we measure it with our backwards units of seconds and meters, we obtain:

$$c := 1 = 299\ 792\ 458 \text{ m/s} \tag{4.12}$$

In order for this to work out, the conversion ratio s/m needs to be 299 792 458.

It is somewhat inconvenient to have two separate units for the same thing. Just like the civilized world got rid of feet, special relativity got rid of seconds. Instead, we measure time in meters. For instance, instead of saying "The bus is going to arrive in three minutes", we would calculate $c \cdot 180\text{s} = 54\text{Gm}$ and say "The bus is going to arrive in 54 gigameters".

4.1.4 Examples of four-vectors

We are now going to take a look at various four-vectors to get a feeling for them. We start with the time-position four-vector.

Events

We have already given an example of a so-called **absolute** event four-vector:

$$x = (\text{July 16, 1945; } \textbf{Trinity Site, New Mexico, Earth}). \tag{4.13}$$

We call this vector "absolute" because it describes a specific point in spacetime. In contrast, relative event four-vectors describe space-time separations between two events. From everyday life, we are already familiar with some types of them:

Purely temporal relative four-vectors describe time separation. For instance, suppose we were standing at the bus stop with the bus arriving in three minutes. There are two relevant absolute four-vectors here: The current event

$$x_0 = (\text{now}, \textbf{here}), \tag{4.14}$$

and the event of the bus arriving

$$x_1 = (\text{now} + 3 \text{ minutes}, \textbf{here}). \tag{4.15}$$

The relative[2] four-vector describing the separation between these two events is

$$x = x_1 - x_0 = (3 \text{ minutes}, \mathbf{0}) = (54 \cdot 10^9, 0, 0, 0) \text{ m}. \tag{4.16}$$

If we square this vector, we find that we get a positive result:

$$x^2 = (54 \cdot 10^9)(\gamma_0)^2 = (54\text{e}9)^2 > 0. \tag{4.17}$$

This is the reason for why we call vectors x with $x^2 > 0$ "timelike" - they describe separations in time. The quantity $\sqrt{x^2}$ has a nice physical meaning: If an observer were to move along this vector through space-time, it would take a subjective time of $\sqrt{x^2}$. This quantity is called the proper time $T = \sqrt{x^2}$. It is called "proper" because its value is the same, no matter the coordinate system. This is also valid for two events that are separated spatially - in that case, the observer has to move through space.

In contrast, let's suppose that we are late to the bus stop, and the bus is still departing while we still are 100m away from the bus stop along the x-axis. This separation four-vector would look like

$$x = (0, 100\text{m}, 0, 0). \tag{4.18}$$

This vector would square to something negative:

$$x^2 = (100m)^2(\gamma_1)^2 = -(100m)^2 < 0 \tag{4.19}$$

This is why we call vectors x with $x^2 < 0$ "spacelike" - they describe separations in space. The quantity $\sqrt{-x^2}$ describes the normal, spatial distance between two points we are used to from classical physics. This is also valid for points that do not happen at the same time coordinate - in this case, $\sqrt{-x^2}$ describes why

The physical difference between timelike, lightlike and null vectors is perhaps best understood by looking at a spacetime diagram where the time axis is depicted geometrically alongside one or two[3] space axes - see Figure 4.1.

If a four-vector $U = U^\mu \gamma_\mu$ describes the velocity of an object, the three cases respectively have the following physical meanings:

[2] It should be noted that the vectors we call "absolute" are actually relative too - the time component is defined relative to an arbitrary moment in time, and the space component is defined relative to Earth. However, in curved spaces like in general relativity, this does not work out anymore - then, absolute points in spacetime and relative vectors are two distinct mathematical concepts.

[3] We'd have a hard time drawing full 1+3D spacetime diagrams, so we usually ignore one or two space axes.

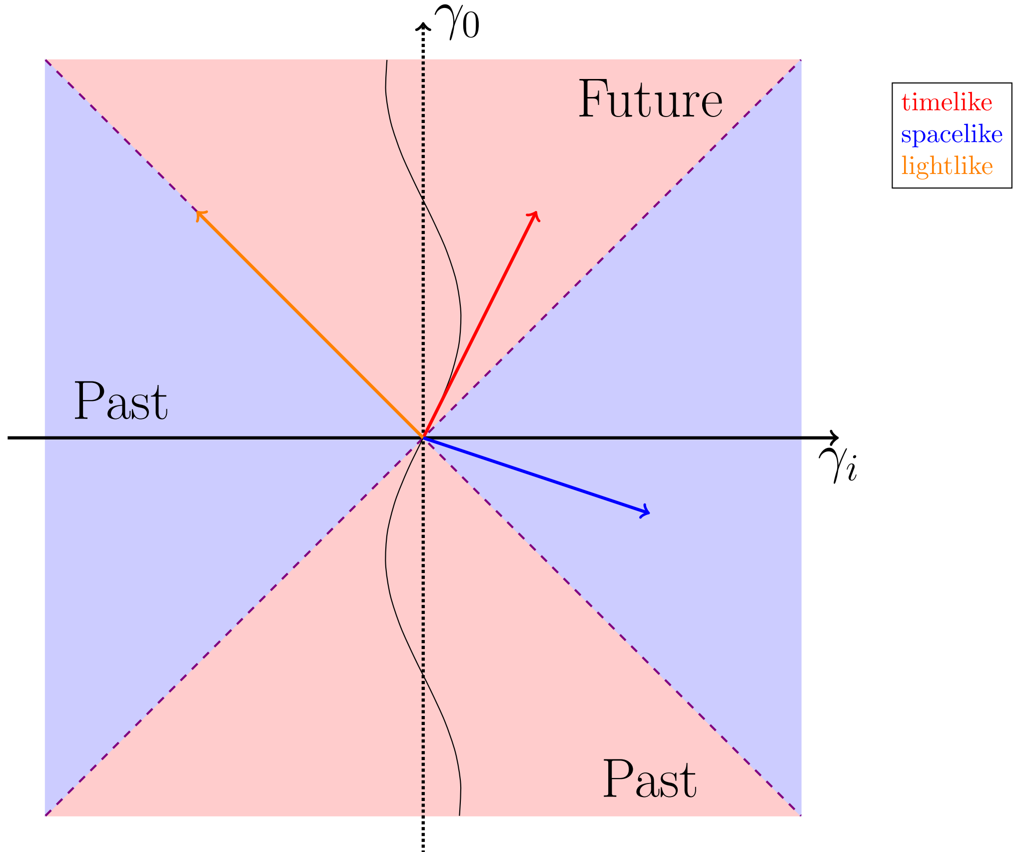


Figure 4.1: A simple spacetime diagram showing the difference between timelike, spacelike, and lightlike vectors. The dashed lines represent the path of a light pulse emitted or absorbed at the origin. The wavy line represents the path a moving object might take.

- timelike U: U describes a massive object moving slower than the speed of light.
- lightlike U: U describes a massless object moving at the speed of light.
- spacelike U: U describes an object moving faster than the speed of light. Currently, no such objects ("tachyons") are known.

Spacelike vectors are mostly used to describe spacelike separations between two events - i.e. the separation between two events occuring at the same time.

Four-velocity

The next type of four-vector we are going to examine is the velocity four-vector $U = U^\mu \gamma_\mu$. Its components tell us how much the time-position of an object changes per unit of proper time. This is related, but not identical to the concept of classical velocity $\mathbf{v}$:

The time component U^0 of describes how fast the time coordinate of the object changes per proper time of the object. For instance, if we watch a spaceship flying past Earth with near the speed of light, its passengers would experience time a lot more slowly than we do. The factor γ by which the astronauts experience time more slowly than we is given by

$$\gamma = \frac{1}{\sqrt{1 - \mathbf{v}^2}} \tag{4.20}$$

where $\mathbf{v}$ is the classical velocity of the spaceship. Remember that since meters and seconds are just different units for the same thing, $\mathbf{v}$ is dimensionless[4]. We can thus write that

$$U^0 = \gamma. \tag{4.21}$$

The classical velocity $\mathbf{v}$ measures how fast the space coordinates change per unit of time coordinate. Therefore, to obtain the spatial parts of the four-velocity U, we need to write:

$$U^i = \gamma v^i. \tag{4.22}$$

The full four-velocity is:

Four-velocity

$$U = (\gamma, \gamma \mathbf{v}). \tag{4.23}$$

A still-standing object with $\mathbf{v} = 0$ only moves through time:

$$U = (1, \mathbf{0}). \tag{4.24}$$

In general, U always squares to 1:

$$U^2 = 1. \tag{4.25}$$

For this reason, we call these U timelike. This relation has a physical interpretation: U describes spacetime separation per unit of proper time. $\sqrt{U^2}$ describes the proper time an observer moving along U experiences. But since U precisely describes separation per proper time, $\sqrt{U^2} = 1$. Therefore, $U^2 = 1$.

However, the story looks different for objects moving at the speed of light. Objects moving at the speed of light do not experience time (for $v \to 1$, $\gamma \to \infty$). Therefore, we cannot define a four-velocity of light.

Four-momentum

The four-momentum unites the classical concepts of energy and momentum. For massive objects, it is given by:

Four-momentum

$$p = mU \tag{4.26}$$

where m is the mass of the object. Its time component describe the energy of the object, while its space components describe the momentum:

$$p = (E, \mathbf{p}). \tag{4.27}$$

For a still-standing object, the four-momentum is given by:

$$p = (m, \mathbf{0}). \tag{4.28}$$

Classically, energy and momentum are two separate quantities, but in GR, they are two sides of the same coin. Similarly, they have the same unit:

$$[\mathbf{p}] = \mathrm{kg}\frac{\mathrm{m}}{\mathrm{s}} \sim \mathrm{kg}. [E] = J = \mathrm{kg}\frac{\mathrm{m}^2}{\mathrm{s}^2} \sim \mathrm{kg}. \tag{4.29}$$

[4]We can obtain the dimensionless $\mathbf{v}$ by dividing the dimensionful $\mathbf{v}$ by $c = 299\ 792\ 458 \mathrm{m/s}$.

We just need to insert the appropriate conversion factors. The famous equation $E = mc^2$ actually is just a mundane conversion formula from units of kilograms to units of joule. (And only valid for objects at rest.)

p squares to m^2:

$$p^2 = m^2 U^2 = m^2. \tag{4.30}$$

In fact, this is how mass is defined in special relativity - mass is the time-length of the momentum four-vector $m = \sqrt{p^2}$.

> Unfortunately, there is a bad habit among physicists to refer to the energy p^0 as the "mass", which is supposedly increasing while the object speeds up. This is not the case - the actual mass is $m = \sqrt{p^2}$. It does not change when we accelerate the object.

A photon with wavelength λ and wavenumber $k = \tau/\lambda$ has energy $E = \hbar k$ and 3-momentum $\mathbf{p} = \hbar k \mathbf{n}$, where $\mathbf{n}$ is some normal vector. The four-momentum of the photon then is

$$p = \hbar(k, \mathbf{n}k). \tag{4.31}$$

It squares to $p^2 = 0$, just as we'd expect - photons are massless.

4.1.5 Brief recap on time in special relativity

The word "time" has two meanings in special relativity: Coordinate time and proper time.

Coordinate time describes time as measured by one specific observer. In other words, coordinate time is the zeroth component x^0 of a four-vector x. Evidently, this quantity is not invariant under spacetime transformations - as we all know, observers moving at very high velocities measure time differently. Therefore, when we want to specify the trajectory a particle takes, we can't just write the position as a function of coordinate time, $x(t)$ - this would cause serious trouble if we were to change coordinate systems.

What we do will do instead is parametrize paths by **proper time**. Proper time T is defined as the time that has passed from the point of view of the object whose trajectory we are trying to describe. If we parametrize the trajectory by some other parameter $x(\lambda)$, the proper time passing for the object between the points λ_1, λ_2 is:

$$T = \int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda \left(\frac{\mathrm{d}x(\lambda)}{\mathrm{d}\lambda}\right)^2. \tag{4.32}$$

If we however parametrize the trajectory by T straightaway, i.e. $x(T)$, we get

$$\left(\frac{\mathrm{d}x(T)}{\mathrm{d}T}\right)^2 = 1 \tag{4.33}$$

we can define the four-velocity U as the tangent vector of the trajectory:

$$U = \frac{\mathrm{d}x(T)}{\mathrm{d}T}. \tag{4.34}$$

This way, we guarantee that U is a unit vector, i.e. $U^2 = 1$. We also see that the time component of the 4-velocity vector represents the amount of coordinate time that passes per unit proper time of the object, and the spatial components the spatial distance the object transverses per unit proper time.

1

γ_0 γ_1 γ_2 γ_3

γ_{10} γ_{20} γ_{30} γ_{23} γ_{31} γ_{12}

γ_{012} γ_{023} γ_{031} γ_{123}

γ_{0123}

Figure 4.2: The multivector pyramid of the spacetime algebra.

4.1.6 STA multivectors

In space algebra, there were four grades: one scalars, three vectors, three bivectors/pseudovectors and one trivector/pseudoscalar. Now, in STA, we have:

- 1 scalar
- 4 vectors
- 6 bivectors
- 4 trivectors/pseudovectors
- 1 tetravector/pseudoscalar.

Note that we have started calling trivectors and tetravectors pseudovectors and pseudoscalars respectively. This is because there are three trivectors and one tetravector in four dimensions, so the tetravector takes the role of the pseudoscalar and the trivector takes the role of the pseudovector. Elements of the individual grades can be represented by fully antisymmetric tensors as follows:

- scalar $a = a$
- vector $v = v^\mu \gamma_\mu$
- bivector $B = 1/2!\ B^{\mu\nu}\gamma_{\mu\nu}$
- trivector/pseudovector $T = 1/3!\ T^{\mu\nu\rho}\gamma_{\mu\nu\rho}$
- tetravector/pseudoscalar $P = 1/4!\ P^{\mu\nu\rho\sigma}\gamma_{\mu\nu\rho\sigma}$.

In general, tensor notation of multivectors works completely analogously to the Cl(3) case - we just have to exchange Latin space indices $(i, j, k, ...)$ for Greek spacetime indices $(\mu, \nu, \rho, ...)$ and remember to use the spacetime metric $\eta_{\mu\nu}$ whenever we need to raise or lower an index. For instance, the bivector B formed from the vectors a, b with the GA expression

$$B = a \wedge b \tag{4.35}$$

would have the tensor analogue

$$B^{\mu\nu} = 2!\ a^{[\mu}b^{\nu]} = a^\mu b^\nu - b^\mu a^\nu. \tag{4.36}$$

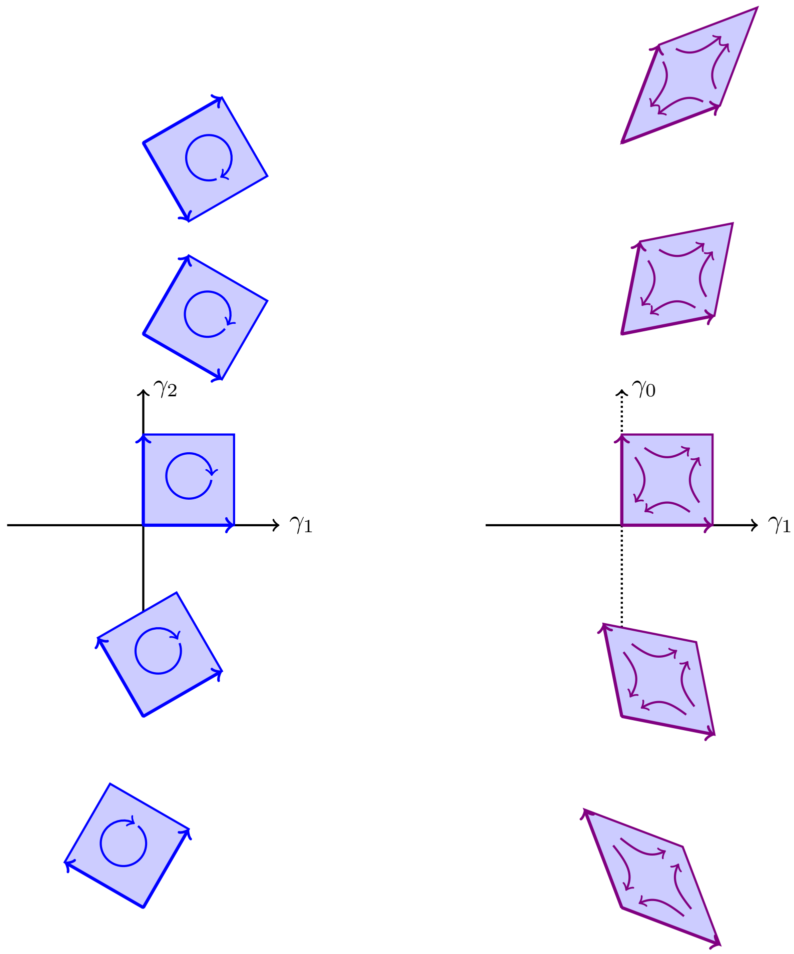


Figure 4.3: A circular bivector $\gamma_1 \wedge \gamma_2$ and a hyperbolic bivector $\gamma_0 \wedge \gamma_1$. Note that the bivectors on the respective sides are all equal to each other. The inscribed shapes describe what sort of transformation the bivectors generate (more on that in the next section).

If we wanted to interior-multiply a four-vector $U = U^\mu \gamma_\mu$ onto this bivector, the expression

$$C = B \cdot U \tag{4.37}$$

would be translated to

$$C^\mu = B^{\mu\nu} U_\nu = B^{\mu\nu} \eta_{\nu\lambda} B^\lambda. \tag{4.38}$$

In STA, there are six basis bivectors. They can be divided into the so-called **hyperbolic** bivectors

$$\gamma_{01}, \qquad \gamma_{02}, \qquad \gamma_{03},$$

and the **circular** bivectors

$$\gamma_{12}, \qquad \gamma_{23}, \qquad \gamma_{31}.$$

The circular bivectors represent purely spatial two-dimensional areas just like in ordinary space algebra. The hyperbolic bivectors, however, have no classical equivalent. They represent oriented areas with one axis in the timelike light cone. The "circular"/"hyperbolic" terminology will become clear soon.

Trivectors represent oriented three-dimensional volumes, but now, there is more than one possible orientation. For instance, the trivector could be purely spatial ($\gamma_1 \wedge \gamma_2 \wedge \gamma_3$) - this would represent a conventional spatial volume that exists for an instant only. We will call such a trivector a **spheroid trivector**. A trivector including a timelike axis, e.g. $\gamma_0 \wedge \gamma_2 \wedge \gamma_3$, would represent a spatial area existing for a finite amount of time. We will call such a trivector a **hyperboloid trivector**. In total, there are four basis trivectors in STA, so every possible trivector can be interpreted as the volume orthogonal to a vector in four dimensions. For this reason, we also call the STA trivectors **pseudovectors**. Specifically:

- spheroid trivectors $\Rightarrow$ **timelike pseudovectors** (for instance $T = I\gamma_0$)
- hyperboloid trivectors $\Rightarrow$ **spacelike pseudovectors** (for instance $T = I\gamma_1$)

Finally, there is only one tetravector in STA - for this reason, we call STA tetravectors **pseudoscalarars**. We conventionally define the unit pseudoscalar as

$$I = \gamma_0 \wedge \gamma_1 \wedge \gamma_2 \wedge \gamma_3. \tag{4.39}$$

Note that we could in principle also have choosen a different order - for instance $\gamma_{0132} = -I$. Choosing either $+I$ or $-I$ as the unit pseudoscalar amounts to choosing an orientation of the space we are working with (either right-handed or left-handed).

The STA pseudoscalar anticommutes with vectors, commutes with bivectors and anticommutes with trivectors:

$$I\gamma_\mu = -\gamma_\mu I \tag{4.40}$$

$$I\gamma_{\mu\nu} = \gamma_{\mu\nu} I \tag{4.41}$$

$$I\gamma_{\mu\nu\rho} = -\gamma_{\mu\nu\rho} I \tag{4.42}$$

4.1.7 A note on the Dirac gamma matrices

Those who have already studied the Dirac equation will have probably recognized the notation we have introduced in this section. The **Dirac gamma matrices** are conventionally defined in terms of the Pauli matrices σ_i as

$$\gamma_0 = \begin{pmatrix} 0 & 1_{2\times 2} \\ 1_{2\times 2} & 0 \end{pmatrix} \tag{4.43}$$

$$\gamma_i = \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix}. \tag{4.44}$$

They are just a matrix representation of the spacetime algebra, similar to how the Pauli matrices are a representation of the space algebra[5]. Any expression composed of sums and products of Dirac matrices can be interpreted as an STA multivector. Similarly to Wolfgang Pauli, Paul Dirac himself did not realize the geometric significance of these matrices - this connection was made by David Hestenes.

[5]To be precise, they represent the *complexified* spacetime algebra, in which the coefficients of k-vectors are complex number-valued. The reason for that is that Dirac fermions are naturally charged, i.e. exhibit $U(1)$ gauge invariance - and complex numbers are precisely what is needed to represent the $U(1)$ group.

4.2 Spacetime rotations

In this section, we are going to learn how to describe Lorentz transformations in geometric algebra. To emphasize the geometric nature of them, we are going to use a slightly different terminology. Conventionally, Lorentz transformations are divided into "rotations" and "Lorentz boosts". However, we are going to see that in spacetime algebra, Lorentz boosts are just a special type of rotation and can be described by rotors. Hence, we will call the "normal" rotations **circular rotations**, and Lorentz boosts **hyperbolic rotations**. Lorentz transformations in general will be referred to as **spacetime rotations**.

4.2.1 Circular rotations

The circular bivectors γ_{ij} generate rotations just like their Cl(3) analogues e_{ij}. There is just one small caveat: The spacelike basis vectors square to -1, so the basis bivectors in STA do not generate positive (counterclockwise), but negative (clockwise) rotations. Consider the following example: The Cl(3) rotor

$$R = \frac{1}{\sqrt{2}}\left(1 + e_{12}\right) \tag{4.45}$$

describes a positive 90° rotation along the xy plane:

$$\tilde{R} e_1 R = e_2. \tag{4.46}$$

To translate this rotor to STA, we need to use the circular bivector γ_{12}. A first guess for the translation of R into STA might be

$$L = \frac{1}{\sqrt{2}}\left(1 + \gamma_{12}\right), \tag{4.47}$$

but checking it reveals that

$$\tilde{L}\gamma_1 L = \frac{1}{2}\left(\gamma_1 + \gamma_{211} + \gamma_{112} + \gamma_{21112}\right) = -\gamma_2 \tag{4.48}$$

i.e. the rotor L performs a negative rotation along the xy plane. This is because γ_1 and γ_2 square to -1.[6] The correct translation of R into STA thus is:

$$L = \frac{1}{\sqrt{2}}\left(1 - \gamma_{12}\right). \tag{4.49}$$

This is an important point, so let us emphasize it:

> The STA circular bivectors γ_{ij} generate negative rotations. The equivalent of the spatial bivector e_{ij} is the STA bivector $-\gamma_{ij} = \gamma_{ji}$, which generates positive rotations.

Generally speaking, the STA equivalent of a Cl(3) rotation around some bivector e_{ij} by an angle of θ

$$R = \exp\left(\frac{e_{ij}}{2}\theta\right) = \cos(\theta/2) + e_{ji}\sin(\theta/2) \tag{4.50}$$

is

$$L = \exp\left(-\frac{\gamma_{ij}}{2}\theta\right) = \cos(\theta/2) - \gamma_{ij}\sin(\theta/2). \tag{4.51}$$

The bivector exponential works exactly like in the Cl(3) case, because the STA circular bivectors square to -1 as well. For instance,

$$(\gamma_{12})^2 = \gamma_{1212} = -\gamma_{1221} = \gamma_{11} = -1 \tag{4.52}$$

Apart from this subtlety, spatial rotations in STA work just as we'd expect based on our three-dimensional intuition. We will call the rotors and rotations generated by the circular bivectors γ_{ij} **circular rotations**.

[6] This is the reason for why we drew the swirls inside the circular bivector in Figure 4.3 in the clockwise sense.

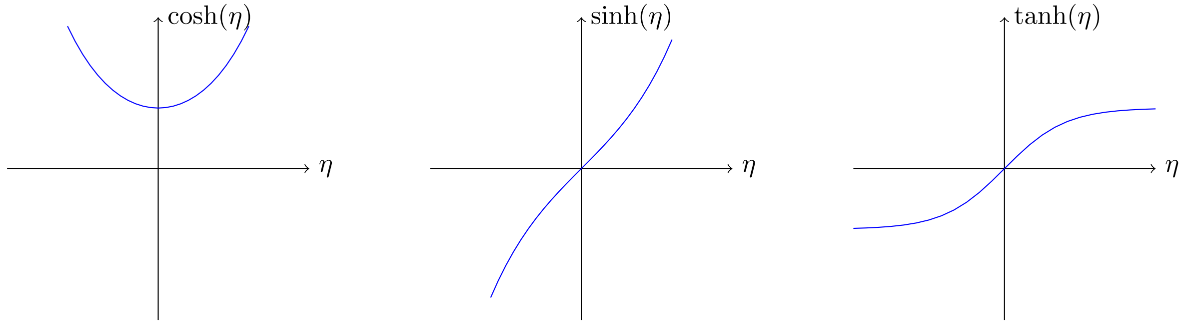


Figure 4.4: The hyperbolic trigonometric functions.

4.2.2 Hyperbolic rotations

We have seen that the circular bivectors γ_{ij} generate circular rotations. In contrast, the hyperbolic bivectors γ_{0i} generate **hyperbolic rotations** - conventionally called Lorentz boosts. We call them **hyperbolic rotations** to emphasize that boosts really are nothing but rotations of the time and space dimensions into each other. However, there are some important differences - for instance, we cannot expect a large enough boost in one direction to return the system to its original state. In the following, we are going to examine the differences between the circular rotations and hyperbolic rotations.

Contrary to the circular bivectors, the hyperbolic bivectors square to +1. For instance,

$$(\gamma_{01})^2 = \gamma_{0101} = -\gamma_{0110} = \gamma_{00} = +1 \tag{4.53}$$

This means that the power expansion of a bivector exponential now resolves to

$$L = \exp\left(\frac{\gamma_{0i}}{2}\eta\right) = \sum_{n=0}^{\infty} \frac{(\gamma_{0i}\ \eta/2)^n}{n!} \tag{4.54}$$

$$= \sum_{n=0}^{\infty} (\gamma_{0i})^n \frac{(\eta/2)^n}{n!} \tag{4.55}$$

$$= \left(\sum_{n=0}^{\infty} \frac{(\eta/2)^{2n}}{(2n)!}\right) + \gamma_{0i}\left(\sum_{n=0}^{\infty} \frac{(\eta/2)^{2n+1}}{(2n+1)!}\right) \tag{4.56}$$

$$= \cosh(\eta/2) + \gamma_{0i}\sinh(\eta/2), \tag{4.57}$$

where cosh and sinh are the hyperbolic trigonometric functions (see Figure 4.4). They describe hyperbolic rotations - Figure 4.5 shows how hyperbolic rotations describe bectors being moved along hyperbolae.

The parameter η is the hyperbolic analogue to the regular, circular angle. We call it the **hyperbolic angle**. It is also called the **hyperbola parameter**, **boost parameter**, or **rapidity**.

Let's consider the timelike unit vector $U = \gamma_0$ (i.e. a still-standing object) being boosted by the hyperbolic rotor

Hyperbolic rotation

$$L = \exp\left(\frac{\gamma_{03}}{2}\eta\right) = \cosh(\eta/2) + \gamma_{03}\sinh\eta/2. \tag{4.58}$$

This is a boost along the z axis, i.e. a hyperbolic rotation in the tz plane. We obtain:

$$\tilde{L}\gamma_0 L = \cosh(\eta)\gamma_0 + \sinh(\eta)\gamma_i \tag{4.59}$$

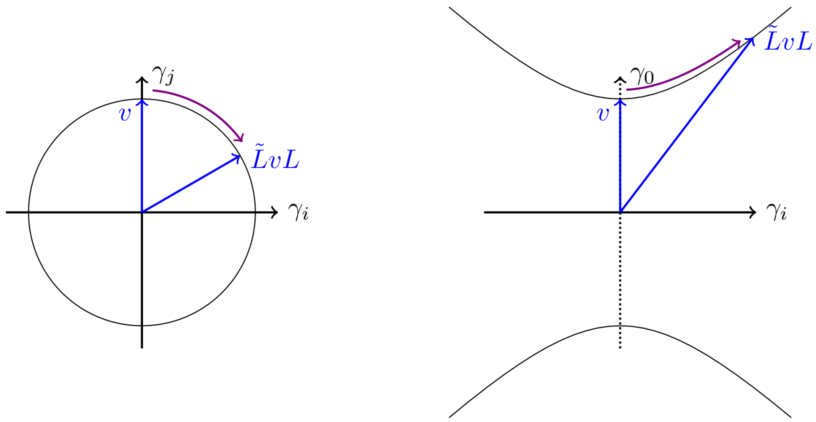


Figure 4.5: The difference between circular rotations and hyperbolic rotations. The vector v might for instance represent the 4-velocity of an object moving through spacetime. In the top diagram, the 4-velocity is being rotated along a purely spatial plane, and in the top diagram, it is being rotated from the time axis to the space axis. In both cases, the 4-magnitude $v^2 = v_\mu v^\mu$ stays invariant - rotations are magnitude-preserving per definition. The mathematical way to state this is that both circular and hyperbolic rotations are elements of the special orthogonal group $\mathrm{SO}(1,3)$.

Its time component $\cosh(\eta)\gamma_0$ represents the amount of coordinate time passing per unit subjective time of the object. The space component $\sinh(\eta)\gamma_3$ represents the amount of coordinate distance the object moves along the z axis per unit subjective time. Therefore, the classical 3-velocity magnitude of the boosted object is

$$v = \frac{\gamma_3 \cdot (L\gamma_0\tilde{L})}{\gamma_0 \cdot (L\gamma_0\tilde{L})} = \frac{\sinh\eta}{\cosh\eta} = \tanh\eta. \tag{4.60}$$

This relation describes how the hyperbolic angle relates to the velocity we boost our reference frame by.

Velocity and hyperbolic angle

$$v = \tanh(\eta) \tag{4.61}$$

For small velocities, $v \approx \eta$, but once η becomes very large, v asymptotically approaches the speed of light $c = 1$. In other words: The components U^0, U^3 of the four-velocity can get arbitrarily large, but the classical speed $v = U^3/U^0$ a bystander would measure can't exceed $c = 1$.

4.2.3 Rotations bivectors revisited

In three dimensions, rotations bivectors consisted of three components,

$$\Theta = \frac{1}{2!}\Theta^{ij} e_i \wedge e_j = \Theta^{12}e_{12} + \Theta^{23}e_{23} + \Theta^{31}e_{31}. \tag{4.62}$$

Now, in 3+1 dimensions, there are six possible bivectors - three hyperbolic and three circular - so our new spacetime rotation bivector Ξ (uppercase xi) has six components:

$$\Xi = \frac{1}{2!}\Xi^{\mu\nu}\gamma_\mu \wedge \gamma_\nu \tag{4.63}$$

We can relate this STA bivector to the hyperbolic angles and classical circular rotational angles:

$$\Xi = \eta^i \gamma_{0i} - \frac{1}{2!}\Theta^{ij}\gamma_i \wedge \gamma_j \tag{4.64}$$

Mind the minus in front of the circular part. This is because the STA circular bivectors generate negative rotations, as discussed previously.

Our treatment of rotation matrices from the last chapter carries over 1:1 to relativistic bivectors. Previously, we saw how the three generators T_i of SO(3) are just the matrix/tensor representations of the bivectors. Now, we have six generators and six matrices representing them (conventionally denoted $M_{\mu\nu}$). However, things get somewhat messy because we constantly need to raise and lower indices with the spacetime metric $\eta_{\mu\nu}$. Spacetime algebra clearly is the superior formalism here.

4.3 Relativistic dynamics

Now that we have seen how spacetime transformations work, we will learn how to do relativistic mechanics in geometric algebra. First, we are going to examine relativistic acceleration and velocity addition. Then, we are going to learn about the basics of relativistic kinematics and rotor mechanics. As a particularly good example of how spacetime algebra simplifies calculations, we are going to learn about relativistic orbital angular momentum.

4.3.1 Relative velocities

Classically, the relative velocity between two moving objects is calculated by subtracting their velocities, but this does not work in special relativity for a simple reason: The relative velocity between two objects is defined as the velocity of one of the objects as seen from the reference frame of the second one. Classically, "boosting" an object to another reference frame just involves subtracting velocities, but as we have seen, we need to perform a full hyperbolic rotation in special relativity for that.

Let U_1, U_2 be the 4-velocities of two observers. Now, we want to calculate the relative velocity between them. Previously, we have seen that the circular angle θ between two unit 3-vectors $\mathbf{a}, \mathbf{b}$ is given by

$$\mathbf{a} \cdot \mathbf{b} = \cos(\theta) \tag{4.65}$$
$$\mathbf{a} \wedge \mathbf{b} = \sin(\theta) B, \tag{4.66}$$

where B is the unit bivector describing the plane of rotation in which A and B lie:

$$B = \frac{\mathbf{a} \wedge \mathbf{b}}{\sqrt{-(\mathbf{a} \wedge \mathbf{b})^2}}. \tag{4.67}$$

The angle θ can be reconstructed by calculating:

$$\tan(\theta) B = \frac{\mathbf{a} \wedge \mathbf{b}}{\mathbf{a} \cdot \mathbf{b}}. \tag{4.68}$$

Hyperbolic angles between two vectors work exactly like this. Given two four-velocity vectors U_1, U_2 with a hyperbolic angle η between them, we can write:

Hyperbolic parallelogram equations

$$U_1 \cdot U_2 = \cosh(\eta) \tag{4.69}$$
$$U_1 \wedge U_2 = \sinh(\eta) C, \tag{4.70}$$

where

$$C = \frac{U_1 \wedge U_2}{\sqrt{(U_1 \wedge U_2)^2}} \tag{4.71}$$

is the hyperbolic bivector generating the boost that rotates U_1 towards U_2 (remember that $U_1 \wedge U_2$ squares to a positive number, so we had to flip the sign under the root). Therefore:

$$\frac{U_1 \wedge U_2}{U_1 \cdot U_2} = \frac{\sinh(\eta)}{\cosh(\eta)} C = \tanh(\eta) C = |v| C. \tag{4.72}$$

is a hyperbolic bivector whose magnitude is equal to the classical velocity. We can also write the bivector $|v|C$ as:

$$|v| C = v\gamma_0, \tag{4.73}$$

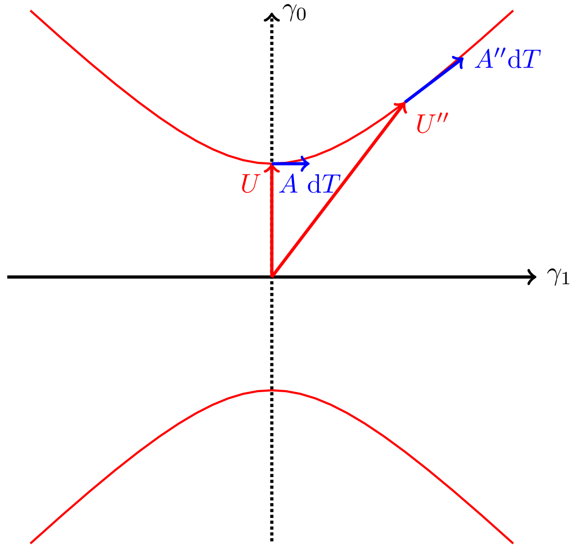


Figure 4.6: Infinitesimal relativistic acceleration. The red hyperbola depicts all possible values the 4-velocity vector of an object with mass can take (i.e. all vectors U with $U^2 = 1$. If we want U to remain on this hyperbola, we need to choose a vector A orthogonal to U. Note that in hyperbolic geometry, our Euclidean intuition of "orthogonal" does not work anymore - in this diagram, U'' is orthogonal to A'' just like U is orthogonal to A.

where v is the spatial classical velocity vector $v = v^i \gamma_i$. For simplicity, we are going to denote this hyperbolic bivector as

$$\mathbf{v} := v\gamma_0 = |v|C. \tag{4.74}$$

The formula for the relative velocity bivector now reads

Relative velocity bivector

$$\mathbf{v} = \frac{U_1 \wedge U_2}{U_1 \cdot U_2} \tag{4.75}$$

such that $v = |\mathbf{v}|$. Note that this v can never become larger than $c = 1$ - just like we'd expect. Don't be confused by the boldface $\mathbf{v}$ - it is indeed a hyperbolic bivector. We will explain why we chose this notation soon.

4.3.2 Relativistic acceleration

If we were to deal with non-accelerated objects only, the systems we are studying would look pretty dull. Therefore, the first problem we need to tackle to do relativistic dynamics is relativistic acceleration.

The four-acceleration A is defined as the derivation of U by the proper time of the object.

Relativistic acceleration

$$A = \dot{U}. \tag{4.76}$$

A direct consequence of this is that A and U have to be orthogonal to each other:

$$U^2 = 1 \tag{4.77}$$

$$\Rightarrow \frac{\mathrm{d}}{\mathrm{d}T} U^2 = 0 \tag{4.78}$$

$$= \dot{U}U + U\dot{U} = 0 \tag{4.79}$$

$$= AU + UA = 0 \tag{4.80}$$

$$= U \cdot A = 0. \tag{4.81}$$

This can be pictured as the requirement that A should only move U along the $U^2 = 1$ hyperboloid - see Figure 4.6 for a depiction.

We now accelerate a non-moving object $U = \gamma_0$ with an acceleration $A = A^i \gamma_i$ orthogonal to U. After a short proper time $\mathrm{d}T$, this results in:

$$U' = \gamma_0 + A\mathrm{d}T. \tag{4.82}$$

After this infinitesimal step, U still approximately squares to one and thus is a valid 4-velocity vector:

$$U'^2 = (U + A\mathrm{d}T)^2 \tag{4.83}$$

$$= U^2 + \underbrace{2U \cdot A\mathrm{d}T}_{0} + \underbrace{A^2 (\mathrm{d}T)^2}_{\mathcal{O}(\mathrm{d}T^2)} \tag{4.84}$$

$$= (\gamma_0)^2 = 1. \tag{4.85}$$

However, this scheme does not work out for faster $U'' \neq \gamma_0$ anymore - then, the new U'' and the original A are not orthogonal anymore, so we would need to boost A into the reference frame of $U'' = \gamma_0$ first (see Figure 4.6). Luckily, there is an easier way to achieve this.

Consider the **acceleration bivector**

Acceleration bivector

$$\mathbf{a} := A \wedge U = AU. \tag{4.86}$$

Again, don't be irritated yet by the boldface notation - $\mathbf{a}$ is indeed a hyperbolic bivector. If we take the interior product between this bivector and a velocity 4-vector, we always get the acceleration 4-vector we need to perform an infinitesimal acceleration. For instance, for $U = \gamma_0$ and $A = a\gamma_3$, we have $\mathbf{a} = a\gamma_{30}$ and:

$$\mathbf{a} \cdot U = a\gamma_{30} \cdot \gamma_0 = a\gamma_3 \tag{4.87}$$

If we instead want to accelerate an already moving object, for instance

$$U'' = \frac{1}{\sqrt{3}} \left(2\gamma_0 + \gamma_3\right), \tag{4.88}$$

we would need to use the updated acceleration four-vector

$$A'' = \mathbf{a} \cdot U'' = \frac{1}{\sqrt{3}} \left(\gamma_0 + 2\gamma_3\right). \tag{4.89}$$

It is easy to check that they are orthogonal to each other. The prescription $A = \mathbf{a} \cdot U$ works for all U.

We now remind ourselves that an expression like $\mathbf{a} \cdot U$ is basically just an infinitesimal hyperbolic rotation along the hyperbolic bivector $\mathbf{a} = a\gamma_{30}$:

$$\mathbf{a} \cdot U = -\left[U, \frac{\mathbf{a}}{2}\right], \tag{4.90}$$

Therefore, a 4-vector U infinitesimally accelerated in the direction of $\mathbf{a}$ for a time $\mathrm{d}T$ reads

$$U' = U - \mathrm{d}T \left[U, \frac{\mathbf{a}}{2}\right]. \tag{4.91}$$

This is exactly the infinitesimal rotor transformation law with a hyperbolic bivector.

Therefore, if we accelerate U for a non-infinitesimal time T with a constant acceleration a, the resulting velocity is given by the full rotor transformation law:

$$U(T) = \tilde{L}(T) U L(T) \tag{4.92}$$

with

$$L(T) = \exp\left(-\frac{\mathbf{a}}{2} T\right) \tag{4.93}$$

$$= \cosh\left(\frac{a}{2} T\right) + \gamma_{03} \sinh\left(\frac{a}{2} T\right). \tag{4.94}$$

The infinitesimal version of this rotor is

$$L(\mathrm{d}T) = 1 - \frac{\mathbf{a}}{2} \mathrm{d}T + \mathcal{O}(\mathrm{d}T^2). \tag{4.95}$$

$$= 1 + \frac{a \mathrm{d}T}{2} \gamma_{03}. \tag{4.96}$$

When we apply the full rotor $L(T)$ to the vector γ_0, we get:

$$U(T) = \tilde{L} \gamma_0 L = \cosh(aT) \gamma_0 + \sinh(aT) \gamma_3. \tag{4.97}$$

We can therefore conclude that the constant acceleration of an object by an acceleration A for a proper time T is the same as boosting the object with boost parameter $\eta = aT$. We see how the speed of light barrier plays out - both a and T can be arbitrarily large, such that the hyperbolic angle η can grow arbitrarily large too. However, $v = \tanh(\eta)$ is bounded to 1 from above and -1 from below.

4.3.3 Relativistic kinematics

The relativistic 4-momentum of an object is given by its 4-velocity multiplied by its mass:

$$p = mU \tag{4.98}$$

The time component of p represents the energy of the object, while the space components represent the classical 3-momentum. The derivation of 4-momentum by proper time is called the 4-force, denoted by f:

4-force

$$f = \dot{p}. \tag{4.99}$$

Therefore, thinking about 4-forces as "currents" of 4-momentum flowing from one object to another is perfectly valid. The timelike component of F represents the kinetic energy flow between them, while the spacelike components are the classical Newtonian force (and therefore the flow of classical 3-momentum).

With these definitions, it is easy to see that the relativistic version of Newton's third law holds:

Relativistic Newton's third law

$$f = mA. \tag{4.100}$$

Relativistic angular momentum

In three dimensions, the expression for the angular momentum bivector of a system is

$$\mathbb{L} = \sum_a \mathbf{x}_{(a)} \wedge \mathbf{p}_{(a)}. \tag{4.101}$$

where $\mathbf{x}$ and $\mathbf{p}$ are the position and momentum 3-vectors, respectively. With STA, we can effortlessly generalize this expression to four dimensions:

Relativistic angular momentum

$$L = \sum_a p_{(a)} \wedge x_{(a)} = -\sum x_{(a)} \wedge p_{(a)} \tag{4.102}$$

where x and p are the position and momentum 4-vectors. We can split up this L into a purely hyperbolic bivector $\mathbf{N} = N^i \gamma_{i0}$ and a purely circular bivector $\mathbb{L}$:

$$L = \mathbf{N} + \mathbb{L} \tag{4.103}$$

Here, the circular components $\mathbb{L}$ are the conventional 3D angular momentum bivector we are used to from classical physics. We will call this type of relativistic angular momentum **circular momentum**. We had to flip the sign of $x \wedge p$ because the circular bivectors of the STA describe clockwise and not counterclockwise rotations.

The hyperbolic components $\mathbf{N}$ have no direct analogue in classical physics. We will call them **hyperbolic momentum**[7]. They represent the position of the center of mass of the system at the time $t = 0$. To see why, we remind ourselves that the hyperbolic components of L are formed from the wedge of the time components of x and the space components of p, and vice versa:

$$N^i = L^{i0} = \sum_a x^i_{(a)} p^0_{(a)} - x^0_{(a)} p^i_{(a)} \tag{4.104}$$

For a slow (non-relativistic) object, $x^0 = t$ and $p^0 = m$. We therefore obtain

$$N^i = L^{i0} = \sum_a m_{(a)} x^i_{(a)} - t p^i_{(a)} \tag{4.105}$$

We see that at $t = 0$, this is just the expression for the center of mass of the system multiplied by the total mass[8]:

$$N^i = M(x_{\text{CoM}})^i = M \frac{1}{M} \sum_a m_{(a)} x^i_{(a)} = \sum_a m_{(a)} x^i_{(a)} \tag{4.106}$$

where $M = \sum_a m_{(a)}$ is the total mass of the system.

However, when the total momentum of the system is not equal to zero, the center of mass will move over time. In fact, the time derivative of $M\mathbf{x}_{\text{CoM}}$ is just the total momentum of the system:

$$\dot{N}^i = M \dot{x}^i_{\text{CoM}} = \sum_a m_{(a)} \dot{x}^i_{(a)} = \sum_a m_{(a)} v^i_{(a)} = \sum_a \mathbf{p}_{(a)}. \tag{4.107}$$

So in order to calculate the "original" ($t = 0$) center of mass, we subtract the total momentum times the time from the current center of mass:

$$N^i = M x^i_{\text{CoM},t=0} = M x^i_{\text{CoM}} - t \sum_a p^i_{(a)} \tag{4.108}$$

$$= \sum_a m_{(a)} x^i_{(a)} - t p^i_{(a)}, \tag{4.109}$$

[7] Sometimes, they are also referred to as the "dynamic mass moment"

[8] The unit of the hyperbolic momentum therefore is position times mass. In natural units, mass is energy, and energy $E = p^0$ is momentum in the direction of time. Thus, hyperbolic and circular momentum have the same units.

which is just the expression for the hyperbolic momentum components $N^i = L^{[i0]}$. Note that this is a conserved quantity - it doesn't change over time. In fact, this makes perfect sense in light of Noether's theorem too - just like the regular circular momentum conservation is linked to the symmetry of circular rotations, hyperbolic momentum conservation is linked to the symmetry of hyperbolic rotations.

It is important to note that when we only apply circular rotations, the hyperbolic and circular components of L (and any other STA bivector) stay separate. However, once we apply a hyperbolic rotation, they get mixed. For instance, consider a system of two point masses orbiting each other in the xy plane with zero net 3-momentum and their center of mass at $\mathbf{x} = 0$. The angular momentum of such a system would be

$$L = |L|\gamma_{21}. \tag{4.110}$$

Once we apply a boost in the y direction with the hyperbolic angle η, this becomes

$$L = |L|(\cosh(\eta)\gamma_{21} - \sinh(\eta)\gamma_{01}). \tag{4.111}$$

We see that a part of the circular momentum has been hyperbolically rotated into the time axis. The resulting hyperbolic momentum describes the original location of the center of mass. You are encouraged to draw a diagram with the γ_0, γ_1 and γ_2 axes to depict this hyperbolic rotation geometrically.[9]

[9] In traditional math, the subject of relativistic angular momentum is notorious for being extremely complicated. It is often not taught until advanced general relativity.

4.4 Space-time split

The conventional four-vector formalism for SR and the classical matrix-vector formalism are two wholly separate algebras. This means that if we want to find the classical equivalent of some given physical law, we need to manually convert the vector expressions between these two formalisms in addition to taking the low-energy limit.

This is not the case for geometric algebra - there is a very elegant device called the **space-time split**, which links the geometric algebras of space and spacetime together seamlessly. We will derive it based on a couple of observations.

4.4.1 The even subalgebra of Cl(1, 3)

The **STA even subalgebra** $\mathrm{Cl}^+(1,3)$ is defined as the subset of $\mathrm{Cl}(1,3)$ that only contains multivectors of even grade. This means that while the full STA contains scalars, vectors, bivectors, pseudovectors and pseudoscalars, the even subalgebra of the STA only contains scalars, bivectors and pseudoscalars. We call this subset a "subalgebra" because it is closed - adding or multiplying even multivectors with each other will always result in even multivectors.

As a "basis" for the even subalgebra, we define the STA bivectors[10]

Even subalgebra basis

$$\sigma_i = \gamma_i \gamma_0. \tag{4.112}$$

such that:

$$\sigma_i = \gamma_{i0} \tag{4.113}$$

$$\sigma_{ij} := \sigma_i \sigma_j = -\gamma_{ij} \tag{4.114}$$

$$\sigma_{123} := \sigma_1 \sigma_2 \sigma_3 = \gamma_{0123} = I. \tag{4.115}$$

We see that by multiplying the σ_μ bivectors with each other, we can span the whole even subalgebra $\mathrm{Cl}^+(1,3)$. In total, we get:

- 1 scalar
- 3 hyperbolic bivectors σ_i
- 3 circular bivectors σ_{ij}
- 1 pseudoscalar σ_{123}

Furthermore, the following anticommutation relation holds:

$$\sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij} \tag{4.116}$$

This looks familiar to us - it is exactly the structure of Cl(3). We can reinterpret the σ_i as the vectors of Cl(3) and the σ_{ij} as the bivectors of Cl(3), and calculate with them as if they were e_i and e_{ij}. In mathematical language, we say that $\mathrm{Cl}^+(1,3)$ and Cl(3) are isomorphic.

Geometrically, we can imagine this as follows: We first single out the spacelike hyperplane of spacetime that is orthogonal to our current γ_0. Hyperbolic bivectors will intersect with that hyperplane along a line - the vector in question. Circular bivectors, on the other hand, will completely lie inside the hyperplane. See Figure 4.7 for an intuitive visualization. This way of reinterpreting the 4D hyperbolic bivectors as the 3D vectors and the 4D circular bivectors as the 3D bivectors is the core of the space-time split.

[10]They have nothing to do with the Pauli matrices - the notation just happens to be the same because of a series of historic accidents.

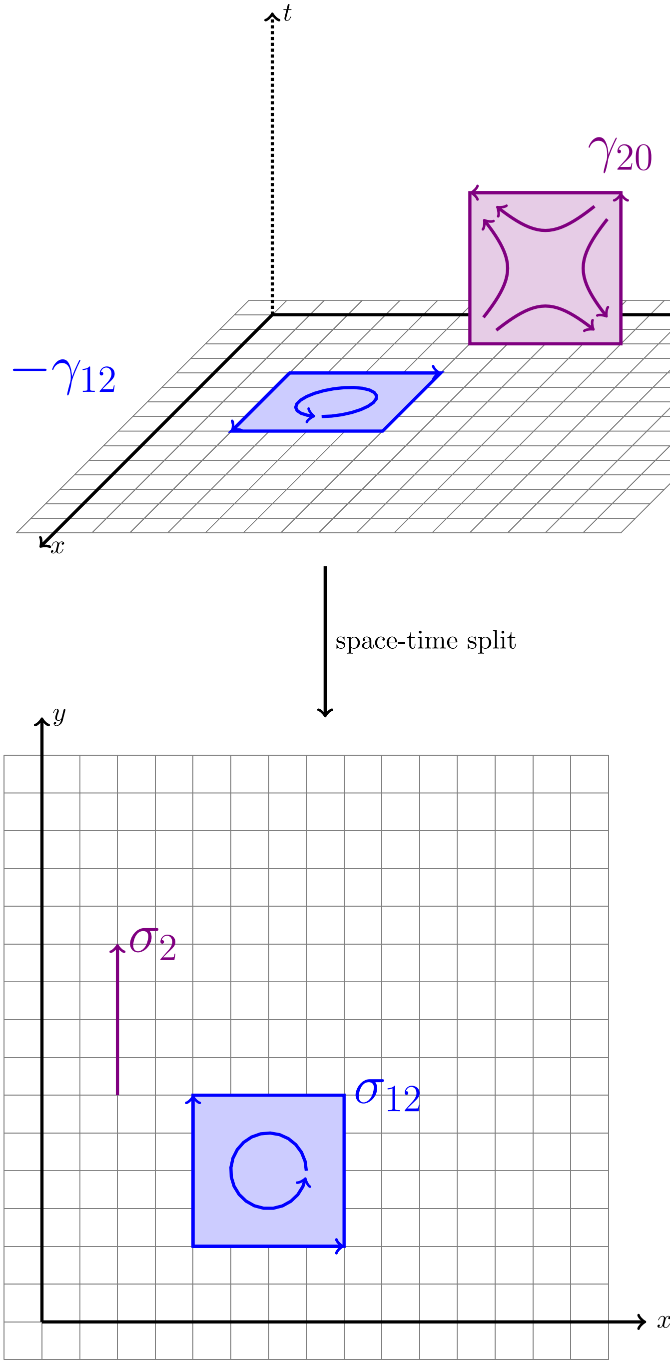


Figure 4.7: The space-time split visualized. We can see how the 4D hyperbolic bivectors are reinterpreted as 3D vectors, and how the 4D circular bivectors are reinterpreted as 3D bivectors. Note that the equivalent of σ_{12} is **not** γ_{12}, but $-\gamma_{12}$. Intuitively speaking, this is because σ_{12} generates positive rotations (equivalently to e_{12}), but γ_{12} generates negative rotations.

4.4.2 Splitting bivectors

In fact, interpreting hyperbolic bivectors of the form $B = b^i \gamma_{i0}$ as classical 3-vectors is not new to us at all:

- When we treated relativistic accelerations, we found that the boosts were generated by the acceleration bivector $\mathbf{a} = aU$, where a is a purely spatial 4-vector $a = a^i \gamma_i$. This means that we can interpret $\mathbf{a} = a\gamma_0 = a^i \sigma_i$ as a classical 3-vector with the space-time split. This makes perfect sense - the classical acceleration is a 3-vector.

- When we calculated the relative velocity between two objects, we derived the relative velocity hyperbolic bivector

$$\mathbf{v} = \frac{U_1 \wedge U_1}{U_1 \cdot U_2}.$$

  Now, we can see that we have actually derived the formula for the classical velocity three-vector.

- In our treatment of the relativistic angular momentum, we saw that while the circular components of L behave like the "normal" 3D angular momentum bivector, the hyperbolic components $L^{[i0]}$ behave like a classical vector indicating the center of mass at $t = 0$. Using the space-time split, we can decompose L into a 3-vector and a 3-bivector:

$$L = N^i \sigma_i + \frac{1}{2!} \mathbb{L}^{ij} \sigma_{ij}$$

Note again that the hyperbolic and circular components of an STA bivector mix as soon as we perform a hyperbolic rotation. Hence, the space-time split is only useful if we commit ourselves to our current inertial frame - otherwise, we are forced to treat STA bivectors as STA bivectors instead of as the sum of a classical 3D vector and bivector.

4.4.3 Splitting vectors

Now that we have seen how we can easily reinterpret STA bivectors as classical vectors, it is somewhat tempting to try to do that for STA vectors too. For instance, a relativistic 4-momentum vector $p = p^\mu \gamma_\mu$ decomposes two parts when we go back to classical physics: the energy $E = p^0$, and the 3-momentum vector $\mathbf{p}$. To obtain them, we calculate:

Splitting a four-vector p

$$pU = E + \mathbf{p} \tag{4.117}$$

where U is the 4-velocity of the observer observing the 4-vector p. Here, the interior part of the geometric product pU represents the energy of the object in question, while the exterior part represents the classical 3-momentum. Conventionally, we choose U to be equal to γ_0 (i.e. the current observer). Then, the split is just

$$p\gamma_0 = p^0 + p^i \sigma_i. \tag{4.118}$$

This is a combination of a scalar and a hyperbolic STA bivector, with the latter being reinterpreted as a classical 3-vector. In general, such quantities are called **split vectors**.

This vector space-time split perfectly showcases the connection between relativistic and classical physics: In special relativity, we use 4-vectors, which unify one timelike and one spacelike quantity. Examples of such pairs include:

- 4-position $x = (t, x^i)$: Time and position

- 4-velocity $U = (\gamma, U^i)$: Time velocity/time dilation and spatial velocity, where γ is the Lorentz factor
- 4-momentum $p = (E, p^i)$: Energy and momentum
- 4-force $f = (P, F^i)$: Power and force (energy current and momentum current)
- electromagnetic 4-potential $A = (\phi, A^i)$: Electric potential and magnetic vector potential
- 4-current $j = (\rho, j^i)$: Charge density and current density
- etc.

The core of special relativity is to treat those quantities as components of unified 4-vectors. Classical physics, on the other hand, splits them up into two separate quantities. The space-time split for vectors does exactly that: It splits up a single 4-vector, for instance the 4-momentum vector p, into a sum of a scalar and a classical 3-vector $p\gamma_0 = E + \mathbf{p}$.

To convert back from the classical $E + \mathbf{p}$ to a 4-vector, we calculate:

Converting the split vector back into a 4-vector

$$p = (E + \mathbf{p}) \cdot U \tag{4.119}$$

where again, U is the 4-velocity of the coordinate system in which the classical quantities E and $\mathbf{p}$ were measured. For instance, if we performed the measurement in the reference frame $U = \gamma_0$, we convert the split vector $E + \mathbf{p}$ to a four-vector p like this:

$$(E + \mathbf{p}) \cdot \gamma_0 = E\gamma_0 + \mathbf{p} \cdot \gamma_0 \tag{4.120}$$

$$= E\gamma_0 + p^i \sigma_i \cdot \gamma_0 \tag{4.121}$$

$$= E\gamma_0 + p^i \frac{1}{2}(\sigma_i\gamma_0 - \gamma_0\sigma_i) \tag{4.122}$$

$$= E\gamma_0 + p^i \frac{1}{2}(\gamma_{i00} - \gamma_{0i0}) \tag{4.123}$$

$$= E\gamma_0 + p^i\gamma_{i00} = E\gamma_0 + p^i\gamma_i = p^\mu\gamma_\mu = p \tag{4.124}$$

This procedure allows us to seamlessly link relativistic and nonrelativistic physics in general - in practice, we can convert a relativistic law into a nonrelativistic one by inserting γ_0's in the right places.

4.4.4 Relativistic acceleration revisited

Now, it is easy to see why we previously defined relativistic 4-acceleration the way we did. By "constant acceleration", we meant to say that the classical 3-acceleration $\mathbf{a} = a^i\sigma_i$ as measured by the accelerated object should be constant thorough the whole process. If the current 4-velocity is U, this classical 3-acceleration $\mathbf{a}$ is equivalent to the 4-acceleration

$$A = \mathbf{a} \cdot \gamma_0 = a^i\gamma_i. \tag{4.125}$$

Hence, the rotor describing an acceleration $\mathbf{a}$ which lasts for a proper time T is just

Relativistic acceleration rotor

$$L(T) = \exp\left(-\frac{\mathbf{a}T}{2}\right). \tag{4.126}$$

In other words: The classical change in 3-velocity $\Delta\mathbf{v} = \mathbf{a}T$ is a 3-vector, but with the space-time split, we can reinterpret it as an STA hyperbolic bivector. This bivector gives us the plane of hyperbolic rotation that describes the relativistic acceleration process, and its magnitude is the hyperbolic angle of this hyperbolic rotation.

4.5 STA geometric derivatives

In STA, we define the geometric derivative analogously to the 3D case:

STA geometric derivative

$$\partial = \gamma^\mu \partial_\mu \tag{4.127}$$

Is is sometimes also known as the **Dirac operator** in quantum field theory. It behaves exactly like the 3D geometric derivative, except for some specialities related to the hyperbolic structure of spacetime: Most importantly, we had to pull up the index of the γ^μ in (4.127) to contract it with ∂_μ. This means that the spacelike basis vectors changed sign:

$$\gamma^0 = \gamma_0 \tag{4.128}$$

$$\gamma^i = -\gamma_i. \tag{4.129}$$

Therefore, our geoemtric derivative is written:

$$\partial = \gamma^\mu \partial_\mu = \gamma_0 \partial_0 - \sum_{i=1}^{3} \gamma_i \partial_i. \tag{4.130}$$

This might seem like a bug at first, but it is a feature - the STA geometric derivative perfectly captures the hyperbolic structure of spacetime. For instance, if we square the geometric derivative:

$$\partial^2 = \partial^\mu \partial_\mu \tag{4.131}$$

This operator is known as the **d'Alembertian** $= \partial^2$ in conventional maths.

The interior derivative of a vector field $A = A^\mu \gamma_\mu$ is:

$$\partial \cdot A = \partial_\mu A^\mu \tag{4.132}$$

It is somewhat comparable to the divergence in 3D, except that it also contains changes in time. Its main use are conservation equations. To understand this better, we first consider the space-time split version of the geometric derivative:

Geometric derivative with spacetime split applied

$$\partial \gamma_0 = \partial_0 - \boldsymbol{\partial} \tag{4.133}$$

Now, in classical physics, the statement that a charge density ρ together with a current density $\mathbf{j} = j^i \sigma_i$ is conserved is written as

$$\partial_0 \rho = -\boldsymbol{\partial} \cdot \mathbf{j} \tag{4.134}$$

The change of the charge density in time should be equal to the influx (negative divergence) of current density. We can rearrange this and substitute $\rho = j^0$:

$$
\begin{aligned}
0 &= \partial_0 \rho + \boldsymbol{\partial} \cdot \mathbf{j} && (4.135)\\
&= \partial_0 j^0 + \frac{1}{2}(\boldsymbol{\partial}\mathbf{j} + \dot{\mathbf{j}}\dot{\boldsymbol{\partial}}) && (4.136)\\
&= \partial_0 j^0 + \partial_i j^k \frac{1}{2}(\sigma^i \sigma_k + \sigma_k \sigma^i) && (4.137)\\
&= \partial_0 j^0 + \partial_i j^k \delta^i_k && (4.138)\\
&= \partial_0 j^0 + \partial_i j^i && (4.139)\\
&= \partial_\mu j^\mu && (4.140)\\
&= \partial \cdot j. && (4.141)
\end{aligned}
$$

Therefore, in STA, the law of current conservation states that the interior derivative of the current field is zero:

Current conservation law

$$\partial \cdot j = 0 \quad (4.142)$$

Any divergence in the spatial directions must be cancelled out by a divergence in the time direction. The exterior derivative $\partial \wedge$ exists too in STA. We shall see how it works in the next chapter.

Chapter 5

Electrodynamics

Now, to make a mistake is easy and natural to man. But that is not enough. The next thing is to correct it: When a mistake has once been started, it is not necessary to go on repeating it for ever and ever with cumulative inconvenience.

Oliver Heaviside on natural units,
Electromagnetic Theory (1893)

Chapter summary

- The electromagnetic field bivector F is the bivector denoted by the antisymmetric rank 2 electromagnetic tensor $F^{\mu\nu}$.
- Electric fields are the hyperbolic components of F. Magnetic fields are the circular components of F.
- With the space-time split, we can reinterpret the E-fields as 3D vectors. However, regardless of the space-time split, B-fields are bivectors.
- In GA, the electromagnetic equation reads $\partial F = j$.
- We can formulate right-chiral and left-chiral plane-wave solutions with the pseudoscalar exponential $\exp(\mp I k \cdot x)$, respectively.
- In electrostatics, we can find solutions by using the Green's function approach.
- Magnetic fields generated by currents and electric fields generated by charges are two very similar phenomena - but this is obscured by traditional maths.
- The electromagnetic force $f = qF \cdot U$ is the force exerted on a charged particle by an electromagnetic field. Hyperbolic EM fields hyperbolically rotate the particle's velocity, while circular EM fields circularly rotate the it.

5.1 Overview

Electromagnetism is one of the most beautiful theories ever discovered - that is, in geometric algebra. In this chapter, we will show how the jumble of vectors, pseudovectors, tensors, minus signs and Lorentz transformations neatly falls into place once we translate the maths to geometric algebra.

5.1.1 Traditional formalism for electromagnetism

Traditionally, all physics students start their studies of electrodynamics by learning about the electric and magnetic "vector" fields $\mathbf{E}, \mathbf{B}$ and the famous Maxwell equations. The Maxwell equations can be divided into the **homogenous** Maxwell equations, in which no charges or currents appear:

$$\nabla \cdot \mathbf{B} = 0 \tag{5.1}$$

$$\nabla \times \mathbf{E} = -\partial_t \mathbf{B} \tag{5.2}$$

and the **inhomogenous** Maxwell equations, which contain the charge density ρ and the current density $\mathbf{j}$:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \tag{5.3}$$

$$\nabla \times \mathbf{B} = \mu_0(\mathbf{j} + \epsilon_0 \partial_t \mathbf{E}). \tag{5.4}$$

Later, when studying relativistic electrodynamics, we conventionally introduce natural units ($c = \epsilon_0 = \mu_0 = 0$). Furthermore, we learn that $\mathbf{E}$ and $\mathbf{B}$ are not vectors anymore, but components of the antisymmetric rank-2 **electromagnetic field tensor**

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^x & -E^y & -E^z \\ E^x & 0 & -B^z & B^y \\ E^y & B^z & 0 & -B^x \\ E^z & -B^y & B^x & 0 \end{pmatrix} \tag{5.5}$$

which then obeys the two relativistic electrodynamic equations

$$\partial_\mu F^{\mu\nu} = j^\nu \tag{5.6}$$

$$3\ \partial^{[\mu} F^{\nu\rho]} = 0 \tag{5.7}$$

Here, (5.6) is equivalent to the two inhomogenous Maxwell equations (5.3, 5.4). It is called the **inhomogenous electrodynamic equation**. Similarly, (5.7) is equivalent to (5.1, 5.2), and hence called the **homogenous electrodynamic equation**. Normally, the factor of 3 is left out - we are using it to foreshadow an important conclusion.

In the conventional approach, there is justification for the form of $F^{\mu\nu}$ whatsoever. Instead, the expression for $F^{\mu\nu}$ is just postulated, with the post-hoc justification that (5.6, 5.7) agree with the four Maxwell equations when all components are written out. The final tensor formalism is completely devoid of any geometric intuition.

5.1.2 Geometric formalism for electromagnetism

In the classical EM formalism, the magnetic field $\mathbf{B}$ is an axial 3-vector. This means that the magnetic field normally looks like a vector, but does an additional sign flip under a parity flip, similar to the angular momentum axial vector. We can easily see this when considering how a current loop generating a B-field would behave under a reflection (Figure 5.1). However, in geometric algebra, we have sworn a sacred oath to replace the concept of axial vectors with the concept of bivectors. Hence, we will stop talking about "magnetic field vectors $\mathbf{B}$" and instead describe magnetic fields with the bivector

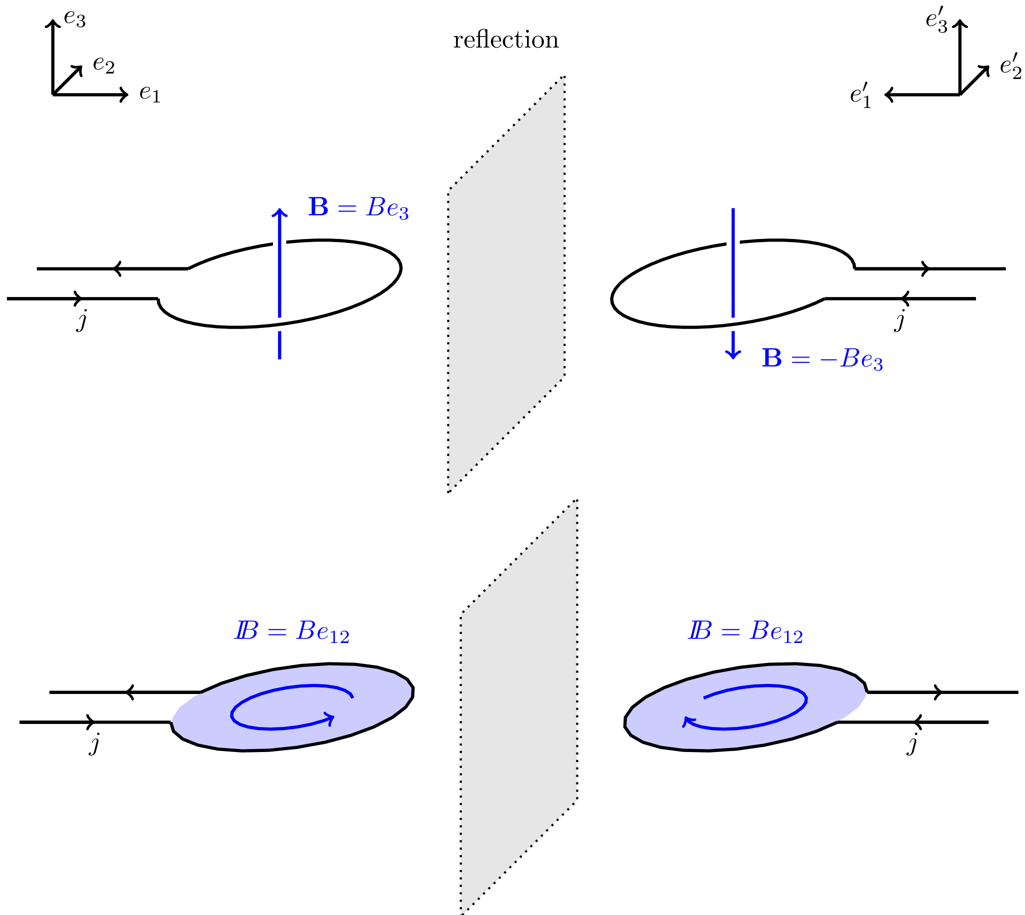


Figure 5.1: The behaviour of B-fields under reflections. We see that it makes much more sense to describe B-fields as bivectors orthogonal to the traditional magnetic "vector" B. To avoid confusion, we will denote the B-field bivector as $I\!B := I\mathbf{B}$.

Magnetic field bivector

$$I\!B = I\mathbf{B} \tag{5.8}$$

i.e. the bivector orthogonal to the original B-field vector.

In contrast, E-fields still are vectors. It is somewhat tempting to add $\mathbf{E}$ and $I\!B$ to obtain the mixed multivector

$$F = \mathbf{E} + I\!B \tag{5.9}$$

$$= E^i e_i + \frac{1}{2!} I\!B^{ij} e_{ij}. \tag{5.10}$$

Here, the vector grade of F is the E-field, and the bivector grade is the B-field. $I\!B^{ij}$ are the components of the bivector $I\!B$.

It might puzzle the reader why we can simply add together the electric and magnetic field into a single quantity, even though they have two different units. It turns out that when we use natural units in which m and s are two units for the same quantity such that $c = 1$, their units (V/m and T) are the same. If we weren't using natural units, we'd have to write:

$$F = \mathbf{E} + c I\!\!B. \tag{5.11}$$

Previously, we have learned that we can use the space-time split to reinterpret STA hyperbolic bivectors as classical vectors, and STA circular bivectors as classical bivectors. We can also do this the other way around - we reinterpret our newly defined F as an STA bivector:

Electromagnetic field bivector

$$\begin{aligned} F &= \mathbf{E} + I\!\!B && (5.12)\\ &= E^i\sigma_i + \frac{1}{2!}B^{ij}\sigma_{ij} && (5.13)\\ &= E^i\gamma_{i0} - \frac{1}{2!}B^{ij}\gamma_{ij} && (5.14) \end{aligned}$$

Now, electric fields are hyperbolic STA bivectors, and magnetic fields are circular STA bivectors. We can see that the components $F^{\mu\nu}$ of this bivector are:

$$\begin{aligned} F^{i0} &= -F^{0i} = E^i && (5.15)\\ F^{ij} &= -I\!\!B^{ij}, && (5.16) \end{aligned}$$

or, more succinctly in matrix notation,

$$\begin{aligned} F^{\mu\nu} &= \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -I\!\!B^{12} & I\!\!B^{31} \\ E^2 & I\!\!B^{12} & 0 & -I\!\!B^{23} \\ E^3 & -I\!\!B^{31} & I\!\!B^{23} & 0 \end{pmatrix} && (5.17)\\ &= \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. && (5.18) \end{aligned}$$

This is where the electromagnetic field "tensor" $F^{\mu\nu}$ comes from. It is just the components of the STA **electromagnetic field bivector** F. Its hyperbolic components are the electric fields, and its circular components are the magnetic fields. With the space-time split, we can interpret the hyperbolic part as a classical vector, and the circular part as a classical bivector/pseudovector.

5.1.3 The electrodynamic equation

Next, we are going to translate the electrodynamic equations into geometric algebra.

The index contraction in the inhomogenous electrodynamic equation

$$\partial_\mu F^{\mu\nu} = j^\nu \tag{5.19}$$

is the component representation of the interior product between the vector derivative ∂ and F, or in other words, the interior derivative of F:

$$\partial \cdot F = j \tag{5.20}$$

with $j = j^\mu\gamma_\mu$. On the other hand, the index diatraction in $3\,\partial^{[\mu}F^{\nu\rho]} = 0$ is the exterior derivative of F:

$$\partial \wedge F = 0. \tag{5.21}$$

The interior derivative $\partial \cdot F$ yields a vector, while the exterior derivative $\partial \wedge F$ yields a trivector. We can therefore simply combine them into a single multivector equation without losing any information:

Electrodynamic equation

$$\partial F = j, \tag{5.22}$$

the GA electrodynamic equation. This simplification of all four Maxwell equations into such a short and beautiful statement perhaps is the main triumph of geometric algebra. And it's not shorthand notation - this is the form we are actually going to use for practical computations in electromagnetism.

5.1.4 A note on differential forms

In exterior calculus, the homogenous and inhomogenous electrodynamic equations read

$$\mathrm{d}F = 0 \tag{5.23}$$
$$*\mathrm{d} * F = -j. \tag{5.24}$$

The first equation is just the exterior derivative, $\partial \wedge F = 0$. The second one is a bit more difficult to translate, though: We remember the GA translation of the Hodge dual of a k-vector,

$$(*X) = (-1)^{\frac{k(k-1)}{2}} IX \tag{5.25}$$

Specifically, in geometric algebra, this means that for a vector V, a bivector B, and a trivector T:

$$*V = IV \tag{5.26}$$
$$*B = -IB \tag{5.27}$$
$$*T = -IT \tag{5.28}$$

In GA notations, these expressions translate to

$$\partial \wedge F = 0 \tag{5.29}$$
$$I\partial \wedge (-IF) = -j \tag{5.30}$$

Rewriting the latter equation reveals that:

$$-I\partial \wedge IF = -I\frac{1}{2}(\dot{\partial} I \dot{F} + I\dot{F}\dot{\partial}) \tag{5.31}$$
$$= -\frac{1}{2}(I\dot{\partial} I \dot{F} + I^2\dot{F}\dot{\partial}) \tag{5.32}$$
$$= \frac{1}{2}(\dot{\partial} I^2 \dot{F} + I^2\dot{F}\dot{\partial}) \tag{5.33}$$
$$= -\frac{1}{2}(\dot{\partial}\dot{F} - \dot{F}\dot{\partial}) \tag{5.34}$$
$$= -\partial \cdot F. \tag{5.35}$$

Therefore, (5.24) translates to

$$\partial \cdot F = j, \tag{5.36}$$

in GA, such that the two equations are equivalent to $\partial F = j$. However, we can directly see the problem with them - exterior algebra does not have the concept of an interior derivative, so we have to use two successive Hodge duals to emulate it. Also, in practical terms, the Hodge dual is very difficult to evaluate by hand because of the abundance of minus signs in its derivation.

<table>
<tr><td>$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$</td><td rowspan="2">$\partial_\mu F^{\mu\nu} = j^\nu$</td><td rowspan="2">$*\mathrm{d} * F = j$</td><td rowspan="2">$\partial \cdot F = j$</td><td rowspan="4">$\partial F = j$</td></tr>
<tr><td>$\nabla \times \mathbf{B} = \mu_0(\mathbf{j} + \epsilon_0 \partial_t \mathbf{E})$</td></tr>
<tr><td>$\nabla \cdot \mathbf{B} = 0$</td><td rowspan="2">$3\ \partial^{[\mu} F^{\nu\rho]} = 0$</td><td rowspan="2">$\mathrm{d}F = 0$</td><td rowspan="2">$\partial \wedge F = 0$</td></tr>
<tr><td>$\nabla \times \mathbf{E} = -\partial_t \mathbf{B}$</td></tr>
</table>

Figure 5.2: The electromagnetic equation in its various formulations.

5.1.5 Gauge fields

Electromagnetism is also commonly formulated in terms of a **gauge field** $A = A^\mu \gamma_\mu$. In literature, this is also called the "electromagnetic four-potential" or "gauge potential". It is the four-vector made up of the electric potential ϕ and the magnetic vector potential $\mathbf{A} = A^i \boldsymbol{\sigma}_i$:

$$A = (\phi + \mathbf{A})\gamma_0 \tag{5.37}$$
$$= \phi\gamma_0 + A^i \gamma_i \tag{5.38}$$
$$= A^\mu \gamma_\mu. \tag{5.39}$$

The electromagnetic field bivector F is derived from A by taking the exterior derivative:

Electromagnetic field bivector in terms of the gauge field

$$F = \partial \wedge A. \tag{5.40}$$

In traditional maths, this relation is commonly known in its component form

$$F^{\mu\nu} = 2!\ \partial^{[\mu} A^{\nu]} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{5.41}$$

The field strenght F is an unambiguous physical parameter we can measure in a lab - for instance, we can measure is hyperbolic components with $F^{i0} = E^i$ with voltmeters and its circular components $F^{ij} = -I\!B^{ij}$ with Hall sensors. The situation is a bit more complicated for A, though. We cannot measure e.g. the electric potential $A^0 = \phi$ - we can just measure differences of ϕ from place to place (these differences are called electric fields). For instance, the two electric fields $\phi(x)$ and $\phi'(x) = \phi(x) + \phi_0$ are physically identical.

This principle is called **gauge symmetry**. In its most general form, it states that if α is some scalar field, the two gauge fields A and $A' = A + \partial\alpha$ are physically identical. This can be seen by calculating the corresponding field strengths:

$$F = \partial \wedge A \tag{5.42}$$
$$F' = \partial \wedge A' \tag{5.43}$$
$$= \partial \wedge (A + \partial\alpha) \tag{5.44}$$
$$= \partial \wedge A + \partial \wedge \partial \wedge \alpha \tag{5.45}$$
$$= \partial \wedge A = F \tag{5.46}$$

because the double exterior derivative $\partial \wedge \partial \wedge$ is always zero (remember that for a scalar, $\partial\alpha = \partial \wedge \alpha$).

Because of the huge variety of gauge fields A representing the same physical field F, we commonly use a procedure called **gauge fixing** to simplify the resulting equations. The most commoly used gauge fixing is the so-called **covariant gauge** (sometimes also called **Loren(t)z gauge**). It requires that our gauge field A fulfills the condition

$$\partial \cdot A = 0, \tag{5.47}$$

or $\partial_\mu A^\mu = 0$ in component notation. It turns out that for every gauge field A, we can find an α such that $A' = A + \partial\alpha$ fulfills the gauge condition. Such a field A' is referred to as **gauge-fixed**.

In terms of a gauge field A, the electrodynamic equation $\partial F = j$ reads

Electrodynamic equation in terms of the gauge field

$$\partial(\partial \wedge A) = j \tag{5.48}$$

Using $\partial = \partial \cdot + \partial\wedge$, we can rewrite this slightly as

$$\partial(\partial - \partial\cdot)A = j \tag{5.49}$$

$$= \partial^2 A - \partial(\partial \cdot A) = j \tag{5.50}$$

In tensor notation, the first term reads

$$(\partial^2 A)^\mu = \partial_\lambda \partial^\lambda A^\mu \tag{5.51}$$

$$= \partial_\lambda \partial^\lambda \eta^{\mu\nu} A_\nu. \tag{5.52}$$

The second term is the gradient of the scalar field $\partial \cdot A$:

$$(\partial(\partial \cdot A))^\mu = (\partial(\partial^\nu A_\nu))^\mu \tag{5.53}$$

$$= \partial^\mu(\partial^\nu A_\nu). \tag{5.54}$$

Now, we are ready to write (5.48) in tensor form:

$$(\partial_\lambda \partial^\lambda \eta^{\mu\nu} - \partial^\mu \partial^\nu) A_\nu = j^\mu. \tag{5.55}$$

The reason for why this equation looks so complicated is that there is no easy way to express the geometric derivative in (5.48) in tensor notation.

If we gauge-fix A to the covariant gauge $\partial \cdot A = 0$, we know that $\partial \wedge A = \partial A$ and hence

Gauge-fixed electrodynamic equation in terms of the gauge field

$$\partial^2 A = j. \tag{5.56}$$

Note that ∂^2 is a fully scalar operator. It is also called the d'Alembertian operator - traditionally denoted $\Box = \partial_\lambda \partial^\lambda$.

This is a very important result - it means that we can treat the different index components of A and j as fully separate:

$$\partial^2 A^\mu = j^\mu. \tag{5.57}$$

This is very useful in the Green's function approach to solving the differential equations. We shall soon see why.

It should also be noted that the homogenous electrodynamic equation $\partial \wedge F = 0$ is automatically fulfilled if we define $F = \partial \wedge A$, because two successive exterior derivatives always yield zero. Viewed slightly differently, the homogenous electrodynamic equation $\partial \wedge F = 0$ *is* the statement that there is a gauge field A corresponding to F.

5.2 Solutions to the electromagnetic equation

Now, we are going to find solutions to the electrodynamic equation. We are first going to take a look at wave solutions in a vacuum, and then present the Green's function method for solving the electrodynamic equation for an arbitrary current density.

5.2.1 Plane-wave solutions

The simplest case of the electrodynamic equation $\partial F = j$ is the vacuum case - the case where $j = 0$ everywhere. The resulting equation $\partial F = 0$ is very easy to solve with a plane-wave approach. Normally, imaginary numbers are used for this purpose. We are going to show that the introduction of imaginary numbers is not necessary with GA.

But first, we have to introduce the pseudoscalar exponential $\exp(I\theta)$. Because $I^2 = -1$, this evaluates to:

$$\exp(I\theta) = \cos(\theta) + I\sin(\theta). \tag{5.58}$$

We are going to use this exponential to write our plane-wave solutions. We make the approach

Left-chiral plane-wave

$$F(x) = F_0 \exp(Ik \cdot x), \tag{5.59}$$

where F_0 is some STA bivector to be determined and k is the 4-wavevector specifying the direction of propagation of the light:

$$k = \omega\gamma_0 + k^i\gamma_i. \tag{5.60}$$

To check whether this is a solution, we take the geometric derivative:

$$\begin{aligned} \partial F(x) &= \partial F_0 \exp(Ik \cdot x) && (5.61)\\ &= \gamma^\mu \partial_\mu F_0 \exp(Ik \cdot x) && (5.62)\\ & && (5.63) \end{aligned}$$

The basis bivector F_0 is constant. The derivative of the pseudoscalar exponential is

$$\partial_\mu \exp(Ik \cdot x) = \frac{\mathrm{d}(k \cdot x)}{\mathrm{d}x^\mu} \frac{\mathrm{d}\exp(I\theta)}{\mathrm{d}\theta}|_{\theta = k\cdot x} = k_\mu I \exp(Ik \cdot x). \tag{5.64}$$

Therefore, the geometric derivative of $F(x)$ evaluates to:

$$\begin{aligned} \partial F(x) &= \gamma^\mu F_0 \partial_\mu \exp(Ik \cdot x) && (5.65)\\ &= \gamma^\mu F_0 k_\mu I \exp(Ik \cdot x) && (5.66)\\ &= kF_0 I \exp(Ik \cdot x) && (5.67)\\ &= kIF_0 \exp(Ik \cdot x) && (5.68)\\ &= kIF(x) && (5.69) \end{aligned}$$

Note that we could shift around the ∂_μ just like we're used to - however, the γ^μ had to stay in place, as vectors do not commute with bivectors.

We now demand that the vacuum electrodynamic equation holds:

$$0 \overset{!}{=} \partial F(x) = kIF_0 \exp(Ik \cdot x) \tag{5.70}$$

We now postmultiply both sides with $\exp(-Ik \cdot x)$ to obtain:

$$0 = kIF_0 \tag{5.71}$$
$$= -IkF_0 \tag{5.72}$$

We premultiply with I:

$$0 = kF_0. \tag{5.73}$$

This already is a very tangible requirement for the basis bivector F_0. But first of all, a brief detour: If we premultiply this equation with k, we find out that:

$$0 = k^2 F_0. \tag{5.74}$$

k^2 is a scalar quantity multiplied to F_0. If we want the vacuum electrodynamic equation to be fulfilled, either k^2 or F_0 should be zero. We want to describe nonzero waves, so F_0 shouldn't be zero. Therefore, we know that $k^2 = 0$, i.e. the wave vector has to be lightlike.

Now, we are ready to solve $kF_0 = 0$. This equation implies that somehow both $k \cdot F_0 = 0$ and $k \wedge F_0 = 0$. Doesn't that mean that k has to be both orthogonal and parallel to F?

Not quite! Normally, if we want to check whether two vectors a, b are orthogonal, we'd take the dot product between them. If it is zero, $a \cdot b = 0$, we'd normally say that the two vectors are orthogonal to each other.

> This scheme to determine orthogonality does not work out anymore for lightlike vectors. If k is a lightlike vector, then $k^2 = k \cdot k = 0$ - even though k most certainly is parallel to k.

Therefore, the only way we can resolve the apparent contradiction $k \cdot F_0 = 0$ and $k \wedge F_0 = 0$ is that F_0 needs to be of the form

$$F_0 = kA_0 \tag{5.75}$$

where A_0 is a vector orthogonal to k, such that F_0 is a pure bivector. Then, we know that:

$$kF_0 = k^2 A_0 = 0. \tag{5.76}$$

General plane-wave solutions

We conclude that our plane-wave approach

$$F(x) = F_0 \exp(Ik \cdot x) \tag{5.77}$$

is a solution to the vacuum electrodynamic equation $\partial F = 0$ if

- k is lightlike: $k^2 = 0$
- F_0 contains k: $F_0 = k \wedge A_0$, with A_0 another spacetime vector.

All of this might be a little bit too abstract, so let's derive an explicit example: We want to derive the solution of an electromagnetic wave propagating in the z direction. The corresponding wavevector is given by:

$$k = \omega(\gamma_0 + \gamma_3), \tag{5.78}$$

where ω is the frequency of the wave. An example vector A_0 orthogonal to this k is

$$V = \frac{1}{\omega}\gamma_1. \tag{5.79}$$

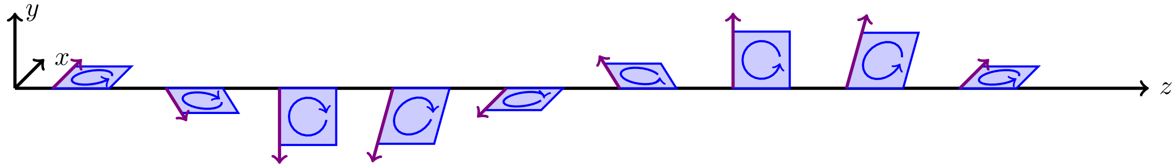


Figure 5.3: The structure of the derived plane-wave solution. We can see how the field starts out as $\sigma_1 + \sigma_{13}$, and then starts to rotate negatively in the xy plane.

We obtain the basis bivector

$$F_0 = k \wedge A_0 \tag{5.80}$$
$$= \gamma_{10} + \gamma_{13} \tag{5.81}$$
$$= \sigma_1 + \sigma_{31}. \tag{5.82}$$

As we can see from the space-time split, this F_0 represents an electric field in the x direction, plus a magnetic bivector field in the zx direction (i.e. a magnetic vector field in the y direction). We now take a look at what happens if we proceed by one quarter-wave (90°), such that $k \cdot x = \tau/4$:

$$F(x') = F_0 \exp(Ik \cdot x) \tag{5.83}$$
$$= F_0 \exp(I\tau/4) \tag{5.84}$$
$$= F_0 I. \tag{5.85}$$

In our case, this just means that

$$F(x') = (\sigma_1 + \sigma_{31})I \tag{5.86}$$
$$= \sigma_1 I + \sigma_{31} I \tag{5.87}$$
$$= \sigma_{23} - \sigma_2. \tag{5.88}$$

We see that the multiplication by the pseudoscalar turns the E-field in the x direction into a B-field in the yz direction, and the zx B-field into a $-y$ E-field[1]. In other words, after one quarter-wave turn, F_0 has been rotated by -90° in the xy plane. The rotation continues as we'd expect - see Figure 5.3 for a depiction. The electric and magnetic field follow an upwards spiral in the direction of the propagation. If we stick out the thumb of our left hand into the direction of the wavevector and form a fist with the remaining four fingers, they will point along the direction of the spiral which leads us into the direction of propagation. It is for this reason that we refer to this type of electromagnetic radiation as **left-hand circularly polarized** or **left-chiral** light.

In contrast, the **right-chiral** wave solutions are given by

Right-chiral plane wave

$$F(x) = F_0 \exp(-Ik \cdot x). \tag{5.89}$$

If we want to describe linear instead of circular polarization, we use a superposition of the two chiral solutions. Namely,

$$F(x) = F_0 \frac{\exp(Ik \cdot x) + \exp(-Ik \cdot x)}{2} \tag{5.90}$$

[1] This multiplication of F by I is the same as the electric-magnetic duality transform $F \to *F$ in differential forms language.

describes linear polarization along the axis of the initial electric field, and

$$F(x) = F_0 \frac{\exp(Ik \cdot x) - \exp(-Ik \cdot x)}{2} \tag{5.91}$$

describes linear polarization perpendicular to the initial electric field.

It is important to note that we did not have to introduce complex numbers at all. Normally, they are ubiquitous, with several disadvantages - for instance, when calculating the energy density and current of the electromagnetic field, we need to explicitly remove the imaginary part.
There is a compelling reason for why we were able to remove complex numbers with such ease: The SO(2) duality symmetry of vacuum electromagnetism. In simple terms, SO(2) duality symmetry states that if F is a vacuum solution,

$$\partial F = 0, \tag{5.92}$$

then $F' = Fe^{I\alpha}$ is a vacuum solution too,

$$\partial(Fe^{I\alpha}) = 0. \tag{5.93}$$

In our derivation, we repurposed this symmetry to rotate around the initial EM field bivector of our plane wave. In traditional maths, this symmetry is expressed via the Hodge dual - we'd say that if the 2-form F is a solution, then $*F$ is a solution too. But since $*F = IF = FI = e^{I\tau/4}$, we can see that this way only allows us to make quarter-turns in the SO(2) circle. Thus, this way of using SO(2) duality to describe plane waves is unique to geometric algebra.

5.2.2 Electrostatics

Electrostatics is the study of electromagnetic fields and currents that are constant in time. We are first going to solve the electrodynamic equation for constant fields in gauge field form using Green's functions, and then derive the corresponding electromagnetic fields.

The electrodynamic equation in gauge field form reads

$$\partial^2 A = j, \tag{5.94}$$

assuming the covariant gauge condition $\partial \cdot A = 0$. In the static case, the operator

$$\partial^2 = \partial_0^2 - \boldsymbol{\partial}^2 \tag{5.95}$$

reduces to $\partial^2 = -\boldsymbol{\partial}^2$, such that the electrodynamic equation becomes the **electrostatic** equation:

Gauge-fixed electrostatic equation

$$-\boldsymbol{\partial}^2 A = j. \tag{5.96}$$

As we have assumed that all fields are time-invariant, we can write

$$A(x) \to A(\mathbf{x}) \tag{5.97}$$
$$j(x) \to j(\mathbf{x}) \tag{5.98}$$

such that we are dealing with a purely spatial PDE

$$-\boldsymbol{\partial}^2 A(\mathbf{x}) = j(\mathbf{x}). \tag{5.99}$$

To solve this equation, we are going to take the Green's function approach. First of all, we note that this equation is linear - if A_1 is a solution for a current j_1 and A_2 a solution for j_2, then $A_1 + A_2$ will be a solution for $j_1 + j_2$. Therefore, in order to solve the electrostatic equation for an arbitrary current density j, we are first going to find solutions $G_{\mathbf{y}}(\mathbf{x})$ to the equation

$$-\boldsymbol{\partial}^2 G_{\mathbf{y}}(\mathbf{x}) = \delta^{(3)}(\mathbf{x} - \mathbf{y}). \tag{5.100}$$

Then, the solution to j will be given by:

$$A(\mathbf{x}) = \int \mathrm{d}^3\mathbf{y}\ j(\mathbf{y}) G_{\mathbf{y}}(\mathbf{x}) \tag{5.101}$$

We can show that this is a solution by inserting it into the electrostatic equation:

$$-\boldsymbol{\partial}^2 A(\mathbf{x}) = \int \mathrm{d}^3\mathbf{y}\ j(\mathbf{y})(-\boldsymbol{\partial}^2 G_{\mathbf{y}}(\mathbf{x})) \tag{5.102}$$

$$= \int \mathrm{d}^3\mathbf{y}\ j(\mathbf{y})\delta^{(3)}(\mathbf{x} - \mathbf{y}) \tag{5.103}$$

$$= j(\mathbf{x}) \tag{5.104}$$

Intuitively speaking: It would be somewhat tedious to find the solutions for the electrostatic equation for some arbitrary j. Hence, we decompose j into lots of infinitesimal and similar-looking fragments. It is very easy to find the solution $G_{\mathbf{y}}(\mathbf{x})$ for these fragments and then join them together with an integral to find the total solution for j.

Now, we just need the expression for $G_{\mathbf{y}}(\mathbf{x})$. We already know that the Green's function for a Laplacian equation is given by:

Electrostatic Greens's function

$$G_{\mathbf{y}}(\mathbf{x}) = \frac{1}{2\tau}\frac{1}{|\mathbf{x} - \mathbf{y}|}. \tag{5.105}$$

Therefore, the explicit expression for the solution of the electrostatic equation is given by

Solution of the electrostatic equation

$$A(\mathbf{x}) = \frac{1}{2\tau}\int \mathrm{d}^3\mathbf{y} \frac{j(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} \tag{5.106}$$

If we want to find the corresponding electromagnetic field, we take the exterior derivative:

$$F(\mathbf{x}) = \frac{1}{2\tau}\int \mathrm{d}^3\mathbf{y}\ \partial \wedge \frac{j(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|}. \tag{5.107}$$

We now simplify the integrand. First, we evaluate:

$$\frac{\mathrm{d}\left(\frac{1}{|\mathbf{r}|}\right)}{\mathrm{d}r^i} = \frac{\mathrm{d}\left(\frac{1}{|\mathbf{r}|}\right)}{\mathrm{d}|\mathbf{r}|}\frac{\mathrm{d}|\mathbf{r}|}{\mathrm{d}r^i} \tag{5.108}$$

$$= \frac{\mathrm{d}\left(\frac{1}{|\mathbf{r}|}\right)}{\mathrm{d}|\mathbf{r}|}\frac{\mathrm{d}\sqrt{\mathbf{r}^2}}{\mathrm{d}\mathbf{r}^2}\frac{\mathrm{d}\mathbf{r}^2}{\mathrm{d}r^i} \tag{5.109}$$

$$= -\frac{1}{\mathbf{r}^2}\frac{1}{2\sqrt{\mathbf{r}^2}}2r_i \tag{5.110}$$

$$= -\frac{r_i}{|\mathbf{r}|^3} \tag{5.111}$$

Therefore, defining $\mathbf{r} = \mathbf{x} - \mathbf{y}$ and $r = \mathbf{r}\gamma_0$ such that $r_0 = 0$:

$$\partial \wedge \frac{j(\mathbf{y})}{|\mathbf{r}|} = \gamma^\mu \wedge \partial_\mu \frac{j(\mathbf{y})}{|\mathbf{r}|} \tag{5.112}$$

$$= -\gamma^\mu \wedge r_\mu \frac{j(\mathbf{y})}{|\mathbf{r}|^3} \tag{5.113}$$

$$= -\frac{r \wedge j(\mathbf{y})}{|\mathbf{r}|^3}. \tag{5.114}$$

We therefore obtain the law

Solution of the electrostatic equation in EM bivector form

$$F(\mathbf{x}) = -\frac{1}{2\tau} \int \mathrm{d}^3\mathbf{x}\ \frac{r \wedge j(\mathbf{y})}{|\mathbf{r}|^3}. \tag{5.115}$$

This formula unites both the Biot-Savart law for magnetic fields generated by spacelike currents and the corresponding law for electric fields generated by timelike currents (charges). This completely unified description of electric and magnetic fields allows us to easily see that they are the same phenomenon. If you're confused by this, you may wish to ponder Figure 5.4 and 5.5.

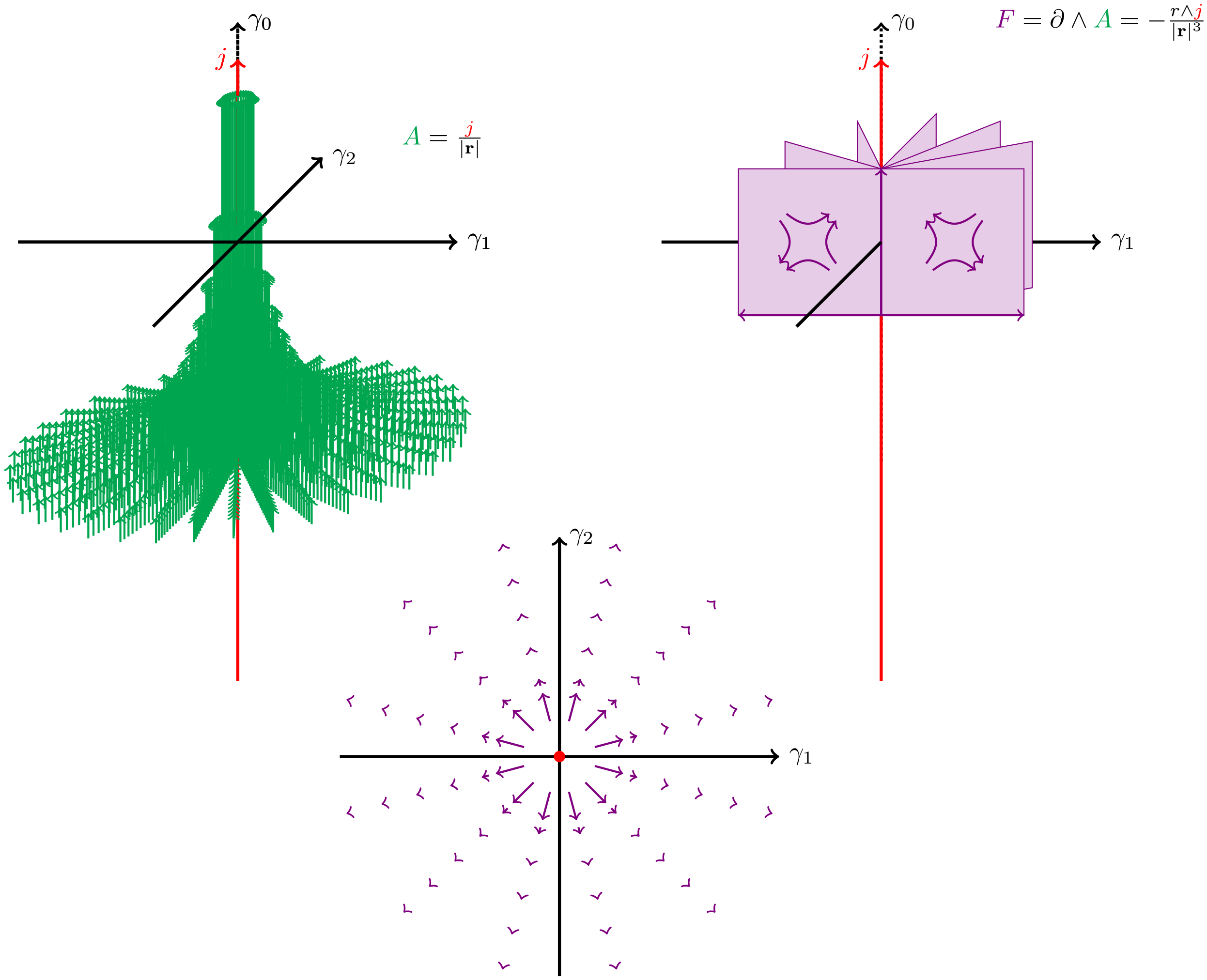


Figure 5.4: The electric field generated by a timelike current (aka charge). The gauge field vectors point in the same spacetime direction as the current and falls off with $1/r$. The electromagnetic field bivectors point in the direction of the current with one axis, and away from the current with the other axis. They fall off with $1/r^2$. Classically, this hyperbolic F-field is interpreted as a vector field (see lower diagram).

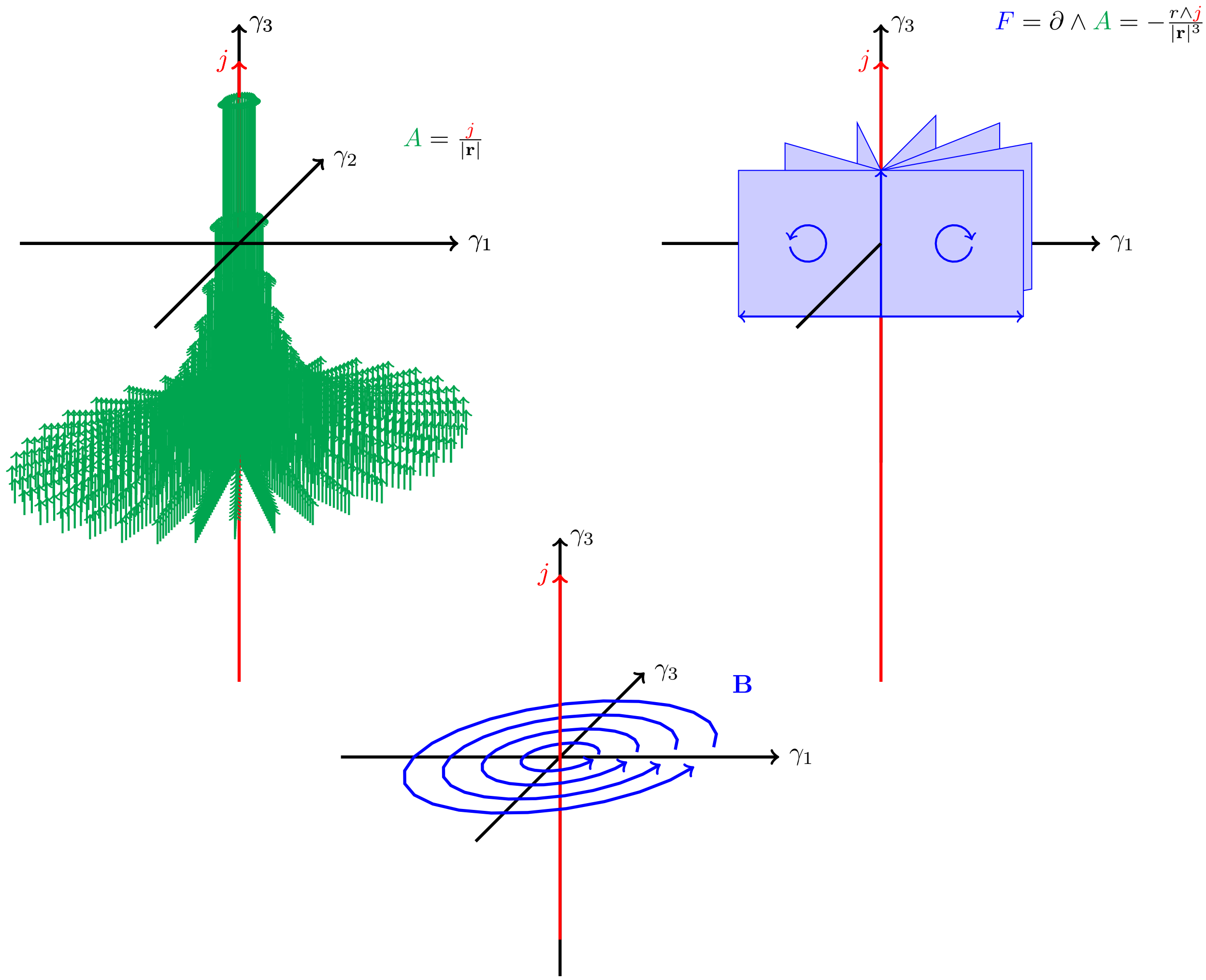


Figure 5.5: The magnetic field generated by a spacelike current. We can see that magnetic fields and electric fields are two sides of the same coin - namely, circular and hyperbolic electromagnetic field vectors, respectively. This relationship gets obscured by the fact that B-fields are conventionally mistakenly interpreted as vectors instead of bivectors, and that the time axis is projected out in the E-field case.

5.3 The electromagnetic force

Electromagnetic fields exert forces on charged particles. Conventionally, this force is divided in two parts:

- The electric or Coulomb force, exerted by electric fields on charged particles: $\mathbf{F} = q\mathbf{E}$
- The magnetic or Lorentz force, exerted by magnetic fields on moving particles: $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$.

In traditional math, these two expressions look like they have little in common. However, in this section, we are going to see that the electric and magnetic forces are just two sides of the same coin.

5.3.1 Reformulation of the electric and magnetic force in STA

Conventionally, the electric force is given by

$$\mathbf{F}_e = q\mathbf{E}. \tag{5.116}$$

This is a classical 3-vector expression. To build a relativistically covariant expression, we apply a reverse space-time split and post-interiormultiply with γ_0:

$$\mathbf{F}_e \cdot \gamma_0 = q\mathbf{E} \cdot \gamma_0 \tag{5.117}$$

We leave the right-hand side untouched, because the electric field is supposed to be the hyperbolic bivector

$$\mathbf{E} = E^i \sigma_i = E^i \gamma_{i0}. \tag{5.118}$$

When we resolve the product on the left-hand side, we obtain the 4-force $f_e = \mathbf{F}_e \cdot \gamma_0 = (F_e)^i \gamma_i$. We therefore obtain the following STA expression for the electric force:

$$f_e = q\mathbf{E} \cdot \gamma_0 \tag{5.119}$$

Now, to the magnetic force. In traditional vector algebra, it is given by

$$\mathbf{F}_m = q\mathbf{v} \times \mathbf{B}. \tag{5.120}$$

First of all, we need to untangle the cross product. We know that

$$\mathbf{v} \times \mathbf{B} = -I(\mathbf{v} \wedge \mathbf{B}) \tag{5.121}$$

$$\tag{5.122}$$

The magnetic field "vector" $\mathbf{B}$ is given by $\mathbf{B} = -I\mathbb{B}$. Therefore,

$$\begin{aligned}
\mathbf{v} \times \mathbf{B} &= -I(\mathbf{v} \wedge (-I\mathbb{B})) &(5.123)\\
&= I(\mathbf{v}I\mathbb{B} - I\mathbb{B}\mathbf{v}) &(5.124)\\
&= I\mathbf{v}I\mathbb{B} - I^2\mathbb{B}\mathbf{v} &(5.125)\\
&= -\mathbf{v}\mathbb{B} + \mathbb{B}\mathbf{v} &(5.126)\\
&= \mathbb{B} \cdot \mathbf{v} &(5.127)
\end{aligned}$$

We therefore arrive at the nonrelativistic magnetic force law

$$\mathbf{F}_m = q\mathbb{B} \cdot \mathbf{v}. \tag{5.128}$$

This already allows for a nice geometric interpretation: The magnetic field bivector $\mathbb{B}$ rotates the velocity along its own plane.

To make this law relativistic, we again post-interiormultiply with γ_0. Note that because $\mathbf{F}_m \wedge \gamma_0 = 0$, we can just postmultiply with the geometric product:

$$f_m = \mathbf{F}_m\gamma_0 = q(\mathbb{B} \cdot \mathbf{v})\gamma_0 \tag{5.129}$$
$$= q(\mathbb{B}\mathbf{v} - \mathbf{v}\mathbb{B})\gamma_0 \tag{5.130}$$
$$= q(\mathbb{B}\mathbf{v}\gamma_0 - \mathbf{v}\mathbb{B}\gamma_0) \tag{5.131}$$

$\mathbb{B}$ does not have any time components, so it commutes with γ_0:

$$f_m = q(\mathbb{B}\mathbf{v}\gamma_0 - \mathbf{v}\gamma_0\mathbb{B}) \tag{5.132}$$
$$= q(\mathbb{B}v - v\mathbb{B}) \tag{5.133}$$
$$= q\mathbb{B} \cdot v \tag{5.134}$$

where $v = v^i\gamma_i$ is the spacelike part of the non-relativistic 4-velocity. Remember that for small $\mathbf{v} = v^i\sigma_i$, the corresponding 4-velocity is given by:

$$U = \gamma_0 + v = \gamma_0 + v^i\gamma_i + \mathcal{O}(v^2) \tag{5.135}$$

We can therefore conclude that we'd need to add an extra γ_0 into (5.134) to obtain the four-velocity. Luckily, we can just get this if we combine it with the electric force law to a single electromagnetic force law:

$$f_{em} = f_e + f_m \tag{5.136}$$
$$= q\mathbf{E} \cdot \gamma_0 \tag{5.137}$$
$$= q(\mathbf{E} \cdot \gamma_0 + \mathbb{B} \cdot v) \tag{5.138}$$

Now, we can use that $\mathbf{E} \cdot v$ is on the order of $1/c^2$ and $\mathbb{B} \cdot \gamma_0 = 0$ to rewrite this as:

$$f_{em} = q(\mathbf{E} + \mathbb{B}) \cdot (\gamma_0 + v) \tag{5.139}$$

Inserting the relations $F = \mathbf{E} + \mathbb{B}$ and $U \approx \gamma_0 + v$ yields:

Electromagnetic force law

$$f_{em} = qF \cdot U \tag{5.140}$$

This is the **electromagnetic force law**. Some of you might recognize it from special relativity - the conventional tensor notation is

$$f^{\mu}_{em} = qF^{\mu\nu}U_{\nu}. \tag{5.141}$$

Again, we see how geometric algebra gives this abstract tensor equation a geometric meaning.

5.3.2 Geometric interpretation

We can reformulate the electromagnetic force law to look like the infinitesimal rotor transformation law:

$$\dot{U} = \frac{f}{m} = \frac{q}{m}F \cdot U \tag{5.142}$$
$$= \left[\frac{q}{2m}F, U\right] \tag{5.143}$$
$$= -\left[U, \frac{q}{2m}F\right]. \tag{5.144}$$

Viewed like this, electric fields (hyperbolic EM fields) generate hyperbolic rotations of U, and magnetic fields (circular EM fields) generate circular rotations of U. This is exactly what we are used to:

- electric fields accelerate, i.e. hyperbolically rotate particles,
- magnetic fields force them on circular cyclotron motion trajectories, i.e. circularly rotate them.

There's just one small caveat: the electromagnetic force law (5.140) acts on U from the *left* side, while the rotor transformation law acts on U from the *right* side. This is the reason for the minus sign in (). We have to pay attention to it: If the particle is positively charged, the rotation generated by the electromagnetic force is opposite to the rotation the bivectors F consists of generate. See Figure 5.6 for an illustration.

Integrating (5.3.2) yields the rotor

$$L(T) = \exp\Big(-\frac{q}{2m}FT\Big) \tag{5.145}$$

such that the velocity of the particle after the proper time T is

$$U(T) = \tilde{L}(T)U_0L(T). \tag{5.146}$$

If F only has a hyperbolic part ($F = \mathbf{E} = E^i\sigma_i = E^i\gamma_{i0}$), this will lead to uniform acceleration:

$$\begin{aligned} L(T) &= \exp\Big(-\frac{q}{2m}FT\Big) && (5.147)\\ &= \exp\Big(-\frac{q}{2m}\mathbf{E}T\Big) && (5.148)\\ &= \cosh\Big(\frac{q}{2m}|\mathbf{E}|T\Big) - \frac{\mathbf{E}}{|\mathbf{E}|}\sinh\Big(\frac{q}{2m}|\mathbf{E}|T\Big), && (5.149)\end{aligned}$$

such that, if we apply this rotor to an initially still-standing ($U = \gamma_0$) object:

$$\begin{aligned} U(T) &= \tilde{L}(T)UL(T) && (5.150)\\ &= \cosh\Big(\frac{q}{2m}|\mathbf{E}|T\Big)\gamma_0 + \sinh\Big(\frac{q}{2m}|\mathbf{E}|T\Big)\frac{E^i}{|E|}\gamma_i. && (5.151)\end{aligned}$$

On the other hand, if F is purely circular ($F = I\!B^{ij}\sigma_{ij}$), the particle will be forced onto a cyclotron motion trajectory:

$$\begin{aligned} L(T) &= \exp\Big(-\frac{q}{2m}FT\Big) && (5.152)\\ &= \exp\Big(-\frac{q}{2m}I\!BT\Big) && (5.153)\\ &= \cos\Big(\frac{q}{2m}|I\!B|T\Big) - \sin\Big(\frac{q}{2m}|I\!B|T\Big)\frac{I\!B}{|I\!B|}. && (5.154)\end{aligned}$$

For instance, if $I\!B = B\sigma_{12}$ and the initial object is moving along the z axis:

$$\begin{aligned} U &= \frac{1}{3}\left(5\gamma_0 + 4\gamma_3\right) && (5.155)\\ \Rightarrow U(T) &= \frac{1}{3}\left(5\gamma_0 + 4\left(\cos\left(\frac{q|I\!B|T}{m}\right)\gamma_1 - \sin\left(\frac{q|I\!B|T}{m}\right)\gamma_2\right)\right), && (5.156)\end{aligned}$$

a spiral trajectory through spacetime. Mind the minus sign in front of the sine.

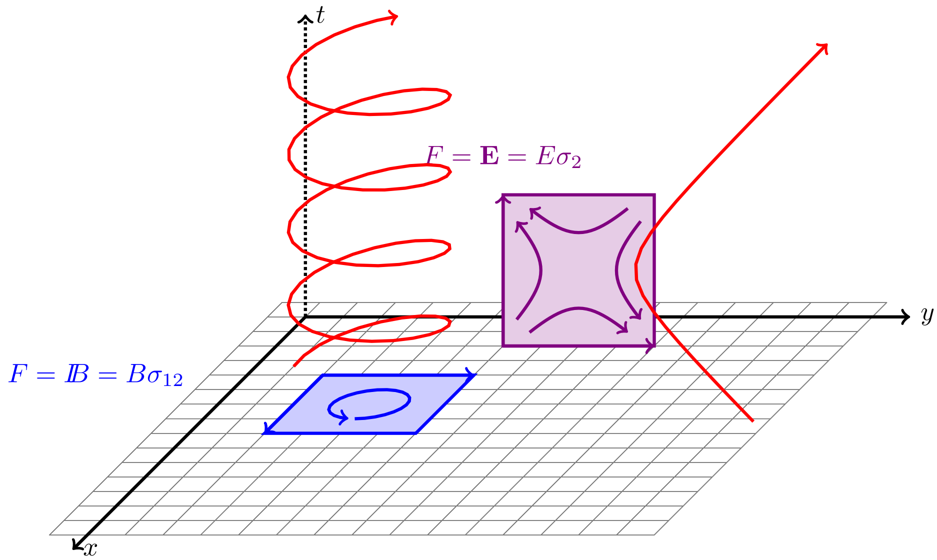


Figure 5.6: The effects of electric and magnetic fields on a positively charged particle. Electric fields are hyperbolic bivectors and hyperbolically rotate the four-velocity of the particle. Magnetic fields are circular bivectors and thus circularly rotate the four-velocity of the particle.

Chapter 6

Geometric calculus

Note: This is a preliminary version of the chapter and needs revisions in content and style.

6.1 Curves and Surfaces

Let us take a step back and consider manifolds, one and two dimensional examples for those are already familiar, as they are curves and surfaces in space. In physics, they also occur in the treatment of mechanics, as a trajectory $(x(t), y(t), z(t))$ of some object describes a curve in three dimensional space, which is parameterized using the time coordinate t. We can write

$$f : \mathbb{R} \to \mathrm{Cl}(3) \tag{6.1}$$

$$t \mapsto f^i(t)e_i. \tag{6.2}$$

So far, our geometric algebra treatment does not tell us much more about this function, just that it sends some scalar parameter t into the vectors in Cl(3), or, if we generalize, Cl(n).
Going further, we can also use multiple parameters, giving us a higher dimensional manifold. By defining a manifold $(x(\lambda, \mu), y(\lambda, \mu), z(\lambda, \mu))$ "living" in three dimensions, which itself has the two parameters (λ, μ), we get a surface plane. We can write

6.1.1 Tangent Vectors

Continuing with $m = 2$, so f defines a surface, we take the derivative with respect to λ or μ, which enables us to investigate the change of the surface vector along the coordinate. So, we have

$$\frac{\partial}{\partial \lambda} f^i(\lambda, \mu)e_i. \tag{6.3}$$

The result is a tangent vector pointing along the λ coordinate.

Going further, the space spanned by all such tangent vectors is called the tangent space, which is a linear space. So, in turn, this linear space, generates a tangent geometric algebra, which we will just call the tangent algebra.
Goign back to our example of a surface plane $m = 2$ in 3D space $n = 3$. This tangent algebra is spanned by two generating vectors, which are the two tangent vectors to the surface. Thus, it is isomorphic to Cl(2), since the tangent space to a surface is a plane, the only new elements to be added (apart from

scalar) is the tangent bivector

$$\frac{\partial f}{\partial \lambda} \wedge \frac{\partial f}{\partial \mu}. \tag{6.4}$$

Not how the tangent bivector represents the surface at a certain point just as well as the pair of tangent vectors does.

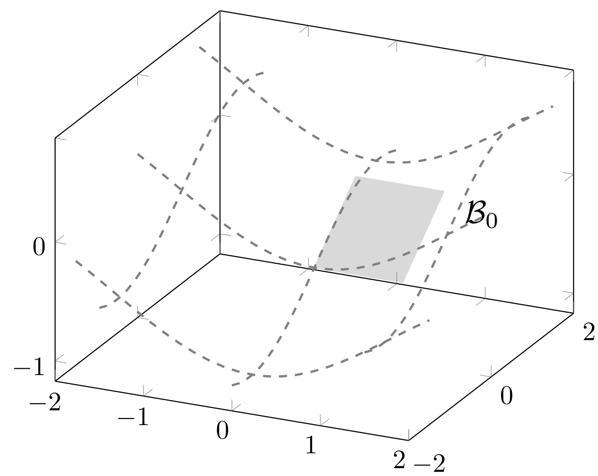


Figure 6.1

6.2 Geometric Calculus

We have already seen that we can understand the elements (λ, μ), coming from $\mathbb{R}^2$, to define the surface $f(\lambda, \mu)$.
Much in the same way as we would like to study the coordinate free geometric algebra, instead of classical vector algebra using coordinates, or row vectors, which are just a basis-dependant representation of these vectors, we should much rather think of surfaces

6.2.1 Nabla Operator

Next, we introduce the familiar nabla operator, in geometric algebra we can treat it as a vector valued operator, which is defined by

$$\partial = e^i \frac{\partial}{\partial x^i} = e^i \partial_i. \tag{6.5}$$

Applying it to some multivector valued expression F, we can take the geometric product

$$\partial F = e^i \partial_i F. \tag{6.6}$$

This is also called the gradient of F, it generalizes the gradient of classical vector algebra, which we would get back if F only has a scalar grade. Remembering that the geoemtric product is not necessarily commutative, note that we could have taken the right gradient just as well

$$\dot{F}\dot{\partial} = \partial_i F(\mathbf{x}) e^i, \tag{6.7}$$

where the overdot was introduces to denote the quantity that the ∂ operator acts on. Without overdots we will just assume the default, that it acts towards its right. Also note, that since ∂_i is a scalar operator, we could pull it to the left, since scalars commute with every element of the geometric algebra. So, we may encounter more complicated expressions, for example

$$\acute{\partial}\dot{\mathbf{u}}\dot{\partial}\acute{\mathbf{v}}, \tag{6.8}$$

where we used overdots and acute accents to denote which part the ∂ operators act on. This can be read as

$$e_i(\partial_k \mathbf{u})e_k(\partial_i \mathbf{v}). \tag{6.9}$$

The e_i and e_k vectors remain where $\acute{\partial}$ and $\dot{\partial}$ were in expression 6.8. The partial derivates are placed where they act on, so ∂_i and ∂_k are in front of $\acute{\mathbf{v}}$ and $\dot{}$

u.

The simplest example for a gradient of an n dimensional vector function would be

$$\begin{aligned} \partial \mathbf{x} &= e^i \partial_i x_j e^j && (6.10)\\ &= e^i e^j \partial_i x_j && (6.11)\\ &= e^i e^j \delta_{ij} && (6.12)\\ &= e^i e^i = n. && (6.13) \end{aligned}$$

Considering a simple scalar function $u(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x}$ on the other hand, we get

$$\begin{aligned} \partial u &= \partial(\mathbf{a} \cdot \mathbf{x}) && (6.14)\\ &= e^i \partial_i a_j x^j && (6.15)\\ &= e^i a_j \partial_i x^j && (6.16)\\ &= e^i a_j \delta_i^j && (6.17)\\ &= e^i a_i = \mathbf{a}. && (6.18) \end{aligned}$$

Now, instead of using components, the overdot notation enables us to quickly calculate expressions in practice, for example the gradient of $r(\mathbf{x}) = \mathbf{x}^2$, using the product rule

$$\begin{aligned} \partial r &= \partial \mathbf{x}^2 && (6.19)\\ &= \dot{\partial}(\dot{\mathbf{x}} \cdot \mathbf{x}) + \dot{\partial}(x \cdot \dot{\mathbf{x}}) && (6.20)\\ &= 2\dot{\partial}(\dot{\mathbf{x}} \cdot \mathbf{x}) = 2\mathbf{x}. && (6.21) \end{aligned}$$

6.2.2 Gradient, Divergence and Curl

Taking the gradient of a general vector valued function $\mathbf{f}$, i.e. a vector field, gives us the geometric product between two vector valued quantities, so the result has scalar and bivectorial components

$$\partial \mathbf{f} = \partial \cdot \mathbf{f} + \partial \wedge \mathbf{f}, \tag{6.22}$$

which are called the divergence and curls respectively.

Conversely, the divergence and curl can be written in terms of the geometric product as

$$\begin{aligned} \partial \cdot \mathbf{f} &= \frac{1}{2}\left(\partial \mathbf{f} + \dot{\mathbf{f}}\dot{\partial}\right) && (6.23)\\ \partial \wedge \mathbf{f} &= \frac{1}{2}\left(\partial \mathbf{f} - \dot{\mathbf{f}}\dot{\partial}\right). && (6.24) \end{aligned}$$

In a way, the curl tells us how much the vector field $\mathbf{f}$ commutes with the derivative operator ∂ and the divergence tells us how much it anticommutes with the derivative operator ∂.
Looking at the curl, we can expand it into coordinates, which yields

$$\partial \wedge \mathbf{f} = e^i \wedge \partial_i \mathbf{f} = e^i \wedge e^j \partial_i J_j. \tag{6.25}$$

In three dimensions, this gives us an expression much similar to the curl of classical vector calculus, but it is a bivectorial expression. In fact, taking the dual returns us the axial vector

$$\partial \times \mathbf{f} = -I \left(\partial \wedge \mathbf{f}\right), \tag{6.26}$$

which is the curl we are familiar with, defined using the cross product. This definition, in contrast to the definition using the wedge product does not carry over to an arbitrary number of dimensions. So, looking at the curl as the wedge product $\partial \wedge \mathbf{f}$ makes it much easier to generalize, while fitting into our geometric algebra treatment.

6.2.3 The Vector Differential

In vector calculus, instead of thinking of derivatives as differentiating with respect to some coordinate, we may want to go further by asking about coordinate free differentiation. Before we can come to that, let us see how the partial derivative is, in a way, the derivative along the coordinate frame vectors e_i. So, the cartesian coordinate chart x^i also defines the coordinate frame vectors e_i, using the frame vectors we know we can just project out the partial derivatives from the ∂ operator by taking the inner product

$$\frac{\partial}{\partial x^i} = \partial_i = e_i \cdot \partial. \tag{6.27}$$

Now, what if we had a different coordinate frame? Defining this vector $\mathbf{h}$ to be some arbitrary vector, we can generalize the expression to yield the directonal derivative

$$\partial_{\mathbf{h}} = \partial_{h^i e_i} = h^i \partial_{e_i} \tag{6.28}$$

and recognize that the directional derivative is the e_i direction is exactly the partial derivative

$$\partial_{e_i} = \partial_i. \tag{6.29}$$

Therefore the directional derivative is simply given by

$$\partial_{\mathbf{h}} = h^i \partial_i = \mathbf{h} \cdot \partial. \tag{6.30}$$

We also see that this expression is linear in $\mathbf{h}$, since it fulfills the properties

$$\partial_{\mathbf{h}+\mathbf{k}} = (\mathbf{h} + \mathbf{k}) \cdot \partial = \partial_{\mathbf{h}} + \partial_{\mathbf{k}} \tag{6.31}$$

and also

$$\partial_{\alpha\mathbf{h}} = \alpha \mathbf{h} \cdot \partial = \alpha \partial_{\mathbf{h}}. \tag{6.32}$$

Looking at this expression as defining a tranformation, that is sending some vector h to the directional derivative of $\mathbf{f}(x)$ along $\mathbf{h}$, we call it the differential of $\mathbf{f}$. It is denoted by

$$\mathbf{f}'_{\mathbf{x}} : \mathbf{h} \mapsto (\mathbf{h} \cdot \partial)\mathbf{f}(\mathbf{x}). \tag{6.33}$$

The directional derivative oeprator $\mathbf{h} \cdot \partial$ is scalar valued, so the differential $\mathbf{f}'_{\mathbf{x}}$, has the same domain and image as the original vector field $\mathbf{f}$,

$$\mathrm{Dom}(\mathbf{f}) = \langle \mathrm{Cl}(m) \rangle_1 \to \langle \mathrm{Cl}(n) \rangle_1 = \mathrm{Im}(\mathbf{f}) \tag{6.34}$$

and the differential is thus also a vector valued function of a vector. But in contrast to $\mathbf{f}$, which is not necessarily linear, $\mathbf{f}'_x$ is a linear transformation. Such linear transformations, we already know how to represent by using matrices. Therefore we try to expand the differential $\mathbf{f}'_x$ as a matrix by using the frame vectors

$$[\mathbf{f}'_x]_{ij} = e_i \cdot \mathbf{f}'(e_j). \tag{6.35}$$

The matrix entries tell us how to tranform the basis vectors, plugging in e_j, which is equivalent to $[0, ... \overbrace{1,}^{j\text{-th entry}} , ..., 0]$, we get the j-th matrix column, wehre the i-th entry is projected out by finally taking the inner product with e_i. Plugging in our definition of the differential gives us

$$e_i \cdot \langle((e_j \cdot \partial)\mathbf{f}(\mathbf{x})\rangle) \tag{6.36}$$
$$= e_i \cdot \partial_j \mathbf{f}(\mathbf{x}) \tag{6.37}$$
$$= \partial_j f_i(\mathbf{x}) \tag{6.38}$$

and so the matrix representation of $\mathbf{f}'_{\mathbf{x}}$ is given by

$$[\mathbf{f}'_x]_{ij} = \partial_j f_i(\mathbf{x}). \tag{6.39}$$

Which is the Jacobian matrix of $\mathbf{f}$.

Chain Rule

Having the differential of vector functions at hand, we will often encounter compositions of such functions. Just as in scalar calculus, we need the chain rule to tell us how to differentiate these expressions. So, say we have the composition of vector functions f, g and h, where

$$h = g \circ f. \tag{6.40}$$

This indicates that the function $z = h(x)$ is defined by evaluating f at x and g at $f(x)$, therefore $z = g(y)$ and $y = f(x)$. For the differential h'_x, the chain rule then gives

$$h'_x = g'_y \circ f'_x. \tag{6.41}$$

Essentially, the differential of the compostion is given by the composition of the differentials, but we have to be careful in defining them with respect to the right variable.
In matrix form, this composition turns into the product of the two jacobian matrices, given by

$$[h'_x]_{ik} = [g'_y]_{ij}[f'_x]_{jk}. \tag{6.42}$$

Total differential

6.2.4 Coordinate Transformations

Instead of the standard cartesian coordinate chart, we might use a differt chart, that more accurately reflects the symmetries underlying our physical system for example. Let us look at a simple case in two dimensions at first, we can write the vector coordinates $\mathbf{x} = xe_1 + ye_2$ in terms of a polar coordinate system as

$$x = r\cos\theta \tag{6.43}$$
$$y = r\sin\theta. \tag{6.44}$$

By this, we rewrite any vector $\mathbf{x}$ in terms of a new basis

$$\mathbf{x} = re_r + \theta e_\theta, \tag{6.45}$$

the main difference to the cartesian basis being, that neither e_r nor e_θ are constant across the two dimensional linear space.
The first question is now how we do calculations, particularly vector derivatives, in this new curvilinear coordinate system. We remember that the coordinate frame belonging to the new chart is given by

$$\frac{\partial \mathbf{x}}{\partial r} = \cos\theta e_1 + \sin\theta e_2 = e_r \tag{6.46}$$
$$\frac{\partial \mathbf{x}}{\partial \theta} = r(-\sin\theta e_1 + \cos\theta e_2) = e_\theta. \tag{6.47}$$

As expected, both of them vary across r and θ, also, while we have

$$e_r \cdot e_r = \cos^2\theta + \sin^2\theta = 1 \tag{6.48}$$

and so e_r is already normalized, $\hat{e}_r = e_r$. The second basis vector e_θ is not

$$e_\theta \cdot e_\theta = r^2(\sin^2\theta + \cos^2\theta) = r^2. \tag{6.49}$$

Therefore, we normalize it by defining

$$\hat{e}_\theta = \frac{e_\theta}{r}. \tag{6.50}$$

The contravariant reciprocal of e_θ, which is denoted e^θ and the covariant e_θ now have to following relationship to the normalized basis vector

$$e^\theta = \frac{1}{r}\hat{e}_\theta \qquad\qquad e_\theta = r\hat{e}_\theta \tag{6.51}$$

This is important to keep in mind, because we can now decompose the vector derivative by taking

$$\partial = \mathrm{Proj}_{e_\theta}(\partial) + \mathrm{Proj}_{e_r}(\partial), \tag{6.52}$$

this results in

$$\partial = e^\theta(e_\theta \cdot \partial) + e^r(e_r \cdot \partial) \tag{6.53}$$

the terms in parenthesis are the partial derivatives with respect to θ and r and the contravariant vectors we can replace by their normalized counterpart

$$\partial = \frac{1}{r}\hat{e}_\theta\partial_\theta + \hat{e}_r\partial_r. \tag{6.54}$$

This is the expresson for the derivative operator in polar coordinates. The form of the Laplacian operator can also be obtained by squaring the expression

$$\partial^2 = \frac{1}{r^2}\partial_\theta^2 + \partial_r^2. \tag{6.55}$$

In general, this means we can calculate the form of the vector derivative in some curvilinear coordinate system, by simply calculating the associated contravariant basis, which is

$$f^i = \frac{1}{f_i} = \frac{\hat{f}_i}{|f_i|} = \left|\frac{\partial \mathbf{x}}{\partial w_i}\right|^{-1}\hat{f}_\theta \tag{6.56}$$

Therefore expanding in the new basis,

$$\partial = \sum_i \mathrm{Proj}_{f_i}\partial = \sum_i f^i\frac{\partial}{\partial w_i} \tag{6.57}$$

and the vector derivative turns into

$$\partial = \sum_i \left|\frac{\partial \mathbf{x}}{\partial w_i}\right|^{-1}\frac{\partial}{\partial w_i}. \tag{6.58}$$

multivector case vectors -¿ scalar + bivector representation

derivative on a manifold tangent maps = Jacobian jacobian determinant

6.2.5 Fundamental Theorem of Geometric Calculus

To be able to calculate multidimensional integrals of multivector valued functions we need an integral theorem, which will be the generalization of the Fundamental Theorem of Calculus, but as it turns out, a large number of unrelated theorems will also emerge as special cases. Because of that, we will start by stating the Fundamental Theorem of Geometric Calculus in its general form. Let M be an m dimensional manifold and F a multivector valued function, then we have

$$\int_M \dot{F}\dot{\partial} d^{\wedge m}x = \oint_{\partial M} F d^{\wedge m-1}x. \tag{6.59}$$

Where $d^{\wedge m}x$ is an m vector valued measure, which is not the same as taking multiple scalar integrals, because that measure would be $(dx)^m$. The integrand term $\dot{F}\dot{\partial}$ is the right gradient of a multivector valued function. This expression is but a certain form of the FT, because we have to consider that the measures and F do not necessarily commute, on the other hand, there is no reason why they should be written towards the right. So, more generally, the FT can be given as

$$\int_m \dot{L}(d^{\wedge m}x\dot{\partial}) = \oint_{\partial M} L(d^{\wedge m-1}x). \tag{6.60}$$

Here, L denotes a function that is linear in its argument, on the left hand side its dependance on x is differentiated, but not its argument. The case with the measures to the left, we can get back from this expression by considering $L(Y) = FY$ and so the integrands turn into

$$L(d^{\wedge m-1}x) = F d^{\wedge m-1}x \tag{6.61}$$

$$\dot{L}(\dot{\partial} d^{\wedge m}x) = \dot{F}\dot{\partial} d^{\wedge m}x. \tag{6.62}$$

Meaning of the Multivector Measure

We already know the notation for multiple integrals, where, for example $d^m x$ the measure for m-fold integration. So this means, we have to take m times the scalar integral each with measure dx. Often, this is also denoted as $dA = dxdy$ for the two dimensional case, where we integrate over a surface or $dV = dxdydz$ for a three dimensional volume.

The multivectorvector measure $d^{\wedge m}x$ on the other hand, is a pseudoscalar over the tangent algebra of the manifold. This means, that it can also encode the orientation of the tangent space. We will define it as

$$d^{\wedge m}x = I_m d^m x \tag{6.63}$$

using the grade m unit pseudoscalar I_m, so $|I_m| = 1$. By this, we can simply convert multivector measures into scalar measures for calculaton. In order to take into account the orientation of the tangent algebra of the manifold M and its $m-1$-dimensional boundary ∂M, we can use the following relationship between their unit pseudoscalars:

$$I_m = \hat{n} I_{m-1} \tag{6.64}$$

Where $\hat{n}$ is the unit normal vector, pointing out of the manifold M. Consider for example a circular area A with positive orientation $I_2 = e_{12}$, since it is not curved this pseudoscalar is constant along the surface.

Next, we see that the boudary curve ∂A has the outward pointing unit normal $\hat{n} = \cos\varphi e_1 + \sin\varphi e_2$. The unit pseudoscalar along the boundary is going to be a vector, since the tangent algebra along a curve is $Cl(1)$, so the pseudoscalar is equal to the vector grade. Inverting the expression for the unit pseudoscalar gives us

$$I_1 = \hat{n} I_2 = \cos\varphi e_2 - \sin\varphi e_1, \tag{6.65}$$

which is the tangent vector to a circle, as expected.

Line Integrals

As a first step in investigating the FT, we will apply it to the line integral case, which is the case where the itegration manifold turns into an $m = 1$ dimensional curve in some space.
Let C be a curve and F a multivector field, then the FT turns into

$$\int_C \dot{F}\dot{\partial}dx = \oint_{\partial C} F d^0 x. \tag{6.66}$$

The right hand side turns into a degenerate case with a scalar valued d^0x. By definition, we say that this means we integrate zero times and this terms has to turn into summation. Since the boundary of a curve ∂C consists of just two points and is in essence zero dimensional, the pseudoscalar is equal to the scalar grade and we can only have $I_0 = \pm 1$. Regardless of the orientation of the curve, if we integrate along a to b, this means the tangent vector points along the outward normal at b and opposite to is along a. So the right hand side becomes

$$I_0(b)F(b) + I_0(a)F(a) = F(b) - F(a) \tag{6.67}$$

and the FT applied to curves becomes

$$\int_C \dot{F}\dot{\partial}dx = F(b) - F(a). \tag{6.68}$$

In the case that we integrate over a scalar valued field ϕ, we clearly see that the right hand side is simply scalar valued, so we have that

$$\int_C (\partial\phi) \cdot dx = \phi(b) - \phi(a). \tag{6.69}$$

The inner product results from the left hand side containing only a non-vanishing scalar grade, the trivector grade necessarily being zero to fulfull the expression. For this reason, we may also call this integral theorem the Gradient Theorem.
Futhermore, in the simplest case, we are integrating over a curve thats embedded in $\mathbb{R}$, so $C = [a, b]$ and we get the standard Fundamental Theorem of Calculus, which tells us how to calculate the integral of $\phi' = \frac{d\phi}{dx}$ using its antiderivative ϕ

$$\int_a^b \phi' dx = \phi(b) - \phi(a). \tag{6.70}$$

Divergence and Curl Theorem

Next, following the Gradient Theorem, we will derive integral theorems involving the divergence and curl of a multivector field. They are also widely known as Gauss' Theorem and Stokes' Theorem.
Let F be a multivetor field of grade r, then we can take

$$\int_M \dot{F}\dot{\partial} \underbrace{d^{\wedge m}x}_{I_m d^m x} = \oint_{\partial M} F \underbrace{d^{\wedge m-1}x}_{\hat{n} I_m d^{m-1}x} \tag{6.71}$$

and cancel the pseudoscalars I_m, giving us the integration theorem with scalar measures

$$\int_M \dot{F} \cdot \dot{\partial} d^m x + \int_M \dot{F} \wedge \dot{\partial} d^m x = \oint_{\partial M} F \hat{n} d^{m-1} x, \tag{6.72}$$

where we split the multivector gradient into divergence and curl, which shows us that the left hand side consists of an $r-1$ grade and $r+1$ grade respectively. Investigating the former component only, we also extract the inner in the integral over the boundary, which results in the relation

$$\int_M \dot{F} \cdot \dot{\partial} d^m x = \oint_{\partial M} F \cdot \hat{n} d^{m-1} x. \tag{6.73}$$

This is Gauss' Theorem, as formulated for a multivector field F.
In the three dimensional case, we are integrating over a volume V, let E be a vector valued field and $d\sigma$ the area measure over the boundary ∂V, then Gauss' Theorem is given by

$$\int_V \partial \cdot E d^3x = \oint_{\partial V} E \cdot d\sigma. \tag{6.74}$$

To derive the second integration theorem, we consider an $m-1$ grade multivector field F on our m-dimensional manifold M, thus the scalar part of the FT reads as

$$\int_M \left(\dot{F} \wedge \dot{\partial}\right) * d^{\wedge m}x = \oint_{\partial M} F * d^{\wedge m-1}x, \tag{6.75}$$

where the asterisk symbol denotes the scalar part of the geometric product between two multivectors $A * B = \langle AB \rangle_0$, which is equal to the inner product if one of the multivectors is of grade one. This theorem is now a generalized form of Stokes Theorem, in the three dimensional vector case, where we integrate over an area A, so $m = 2$. The left hand sides then presents us the scalar part of the product between two bivectors

$$\left(\dot{F} \wedge \dot{\partial}\right) * d^{\wedge 2}x. \tag{6.76}$$

This expression is simply calculated from the duals vectors of the bivectors by taking

$$-\left(\dot{F} \times \dot{\partial}\right) \cdot d^2x, \tag{6.77}$$

where we take note of the minus sign, also, we see that we get the classical definition of the curl using the cross product.
Thus, the curl theorem is given by

$$\int_A (\partial \times F) \cdot d^2x = \oint_{\partial A} F \cdot dx. \tag{6.78}$$

Green's Integration Theorem

For this we consider the general curl theorem in the two dimensional case, so we take the integral over an area A in two dimensions, thus

$$\int_A \left(\dot{F} \wedge \dot{\partial}\right) * d^{\wedge 2}x = \oint_{\partial A} F * dx. \tag{6.79}$$

If we write the 2D vector field as $F(x) = f(x,y)e_1 + g(x,y)e_2$, the left hand side becomes

$$\int_A \left(\frac{\partial f}{\partial y} - \frac{\partial g}{\partial x}\right) e_{12} * dxdye_{12} = \int_A \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dxdy. \tag{6.80}$$

For the left hand side of the divergence theorem, we considers that the vector measure dx can be decomposed $dx = e_1dx + e_2dy$, so the full theorem can now be written as

$$\int_A \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dxdy = \oint_{\partial A} fdx + gdy. \tag{6.81}$$

This is Green's Integration Theorem, which is essentially just the two dimensional case of the general divergence theorem.

Green's Functions

The FT, in fact, does not only generalize vector calculus theorems, but also integration theorems of complex analysis. Before we come to this, we will introduce Green's functions.
Let the differential operator D be given, Green's functions arise when we are asked to solve the inhomogenous dífferential equation

$$Du(x) = f(x) \tag{6.82}$$

where f is a given vector valued function.
In order to solve this differential equation, we can use the Green's function of the operator D, that satisfies the property

$$DG_a(x) = \delta(x - a). \tag{6.83}$$

The function value $f(x)$ can be rewritten by taking an integral over the variable a

$$f(x) = \int \delta(x - a) f(a) da \tag{6.84}$$

therefore we can use the Green's function, replacing the Dirac delta

$$f(x) = \int DG_a(x) f(a) da. \tag{6.85}$$

The differential operator D only acts on x by definition, not on a, so we can simply pull it out of the integral now

$$f(x) = D\left(\int G_a(x) f(a) da\right) \tag{6.86}$$

and, by doing this, we obtain a solution $u(x)$ to the differential equation

$$u(x) = \int G_a(x) f(a) da. \tag{6.87}$$

6.2.6 Analytic Functions

Geometric Algebra enables us to generalize much of vector calculus, much in the same way this is also possible for complex analysis. For this we need the notion of an analytic function. First, we take the GA of the two dimensional plane Cl(2), a vector is thus given by

$$\mathbf{r} = xe_1 + ye_2, \tag{6.88}$$

the e_1-axis can now be singled out as corresponding to the real axis of the complex plane by simply taking the geoemtric product $z = e_1\mathbf{r}$ such that now

$$z = x + ye_{12}. \tag{6.89}$$

This is essentially equivalent to a complex number, since the even subalgebra $\mathrm{Cl}^+(2)$ is isomorphic to the complex plane. In the following we will denote the pseudoscalar simply as $I = e_{12}$, representing that the pseudoscalar takes onthe role of the imaginary unit by fulfilling $I^2 = -1$.

$$z = x + Iy \tag{6.90}$$

The reverse $z^\dagger$, then takes the role of the complex conjugate, which is thus defined by $z^\dagger = \mathbf{r}e_1$.

$$z^\dagger = x - Iy \tag{6.91}$$

On this even subalgebra we can now investigate certain types of functions, let $\psi = \psi(z(x, y))$ be such a complex field. We take its vector derivative

$$\partial\psi = \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) e_1 + \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) e_2. \tag{6.92}$$

The bracketed terms vanish, if ψ fulfills the Cauchy-Rieman equations, so ψ being holomorphic is equivalent to

$$\partial\psi = 0. \tag{6.93}$$

We can also introduce the complex partial derivatives, which have the following properties

$$\frac{\partial z}{\partial z} = 1, \qquad \frac{\partial z^\dagger}{\partial z} = 0. \tag{6.94}$$

Therefore, we see that the holomorphic function ψ is independent of $z^\dagger$, by taking

$$\frac{\partial \psi}{\partial z^\dagger} = \frac{\partial \psi}{\partial z}\frac{\partial z}{\partial z^\dagger} = 0 \tag{6.95}$$

Also, in terms of the vector derivate, the complex derivatives can be expressed by

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - I\frac{\partial}{\partial y}\right) = \frac{1}{2}\widetilde{e_1\partial} \tag{6.96}$$

$$\frac{\partial}{\partial z^\dagger} = \frac{1}{2}\left(\frac{\partial}{\partial x} + I\frac{\partial}{\partial y}\right) = \frac{1}{2}e_1\partial \tag{6.97}$$

Let us look at some examples for analytic functions, the easiest function we can try is simply the identity $f(z) = z$. We calculate its gradient

$$\partial z = 2e_1\frac{\partial}{\partial z^\dagger}z = 0, \tag{6.98}$$

which vanishes, so this functions is indeed analytic. In fact we can therefore conclude that the expression $\partial(z-a)^n$ vanishes as well, by considering the product rule

$$\partial(z-a)^n = n(z-a)^{n-1}\partial(z-a) = 0. \tag{6.99}$$

Cauchy's Integral Formula

Finally, from what we have gathered, we will derive Cauchy's Integral formula. The Cauchy kernel in complex form is given by

$$\frac{1}{z-a} = \frac{1}{e_1(\mathbf{z} - e_1\mathbf{a})} \tag{6.100}$$

$$= \frac{1}{\mathbf{r}-\mathbf{a}}\frac{1}{e_1} \qquad = \frac{\mathbf{r}-\mathbf{a}}{(\mathbf{r}-\mathbf{a})^2}e_1, \tag{6.101}$$

which enables us to switch between the two. Take the integral

$$\oint_{\partial A} \frac{f(z)}{z-a}dz, \tag{6.102}$$

which we now know how to convert into an integral along a vector curve ∂A

$$\oint_{\partial A} f(\mathbf{r})\frac{\mathbf{r}-\mathbf{a}}{(\mathbf{r}-\mathbf{a})^2}e_1 e_1 dr. \tag{6.103}$$

This expression we convert to an integral along the area A by using the Fundamental Theorem of Calculus

$$\int_A f(\mathbf{r})\dot{\frac{\mathbf{r}-\mathbf{a}}{(\mathbf{r}-\mathbf{a})^2}}\dot{\partial}Idxdy \tag{6.104}$$

$$=\int_A f(\mathbf{r})\tau\delta(\mathbf{r}-\mathbf{a})Idxdy \tag{6.105}$$

$$=\tau If(\mathbf{a}) \tag{6.106}$$

Evaluating the delta function integral gives us the solution to the complex path integral, so we can finally write out Cauchy's Integral formula

$$f(a)=\frac{1}{\tau I}\oint_{\partial A}\frac{f(z)}{z-a}dz. \tag{6.107}$$

In the n-dimensional case, the derivation can be followed analogously, the Green's function for the vector derivative is then given by

$$G_y(x)\frac{1}{\sigma_n}\frac{x-y}{|x-y|^n}. \tag{6.108}$$

The factor σ_n in front being the surface area of the n-hypersphere and the Green's function satisfies the property

$$\partial G_y(x)=\delta(x-y). \tag{6.109}$$

The fundamental theorem is given by

$$\oint_{\partial V} GdS\psi=\int_V\left(\dot{G}\dot{\partial}\psi+G\dot{\partial}\dot{\psi}\right)dV, \tag{6.110}$$

here the second part vanishes since ψ is analytic, ans we are left with

$$\int_V \dot{G}\dot{\partial}\psi dV, \tag{6.111}$$

which turns into the integral over the delta function again

$$\int_V \sigma_n\delta(x-y)\psi(x)dV=\sigma_n\psi(y)I. \tag{6.112}$$

Therefore, Cauchy's Integral formula in its n-diemnsional form can written as

$$\psi(y)=\frac{1}{\sigma_n I}\oint_{\partial V}\frac{x-y}{|x-y|^n}dS\psi(x). \tag{6.113}$$

Chapter 7

Spinors

> No one fully understands spinors. Their algebra is formally understood but their general significance is mysterious. In some sense they describe the "square root" of geometry and, just as understanding the square root of 1 took centuries, the same might be true of spinors.
>
> Michael Atiyah (1929-2019), expert on spinors

7.1 Matrix representations

In the last couple of chapters, we have encountered two examples of **matrix representations**: The Pauli matrix representation, which represents the geometric algebra of space Cl(3), and the Dirac matrix representation, which represents the geometric algebra of spacetime $\mathrm{Cl}(1,3)$.

At first, it looks like the concept of matrix representation is just a neat trick to "emulate" geometric algebra with matrix algebra. However, as we will see in this chapter, the concept of matrix representations is central to describing quantum-mechanical spin and spinors.

First of all, we should become a little more concrete about what we mean by "matrix representation". Representation theory is a rather complicated branch of mathematics that deserves a lecture of its own. Luckily, we don't have to do all of it - we only need **complex matrix representations**.

Definition of a complex matrix representation

Mathematically, an n-dimensional complex matrix representation is defined as a bijection that assigns a complex $n \times n$ matrix to every multivector:

$$\rho : \mathrm{Cl}(p, q) \mapsto \mathrm{Mat}(n \times n, \mathbb{C}) \tag{7.1}$$

This bijection should be an **isomorphism** - this means that adding and multiplying these matrices should be equivalent to adding and multiplying their underlying multivectors:

$$\rho(A + B) = \rho(A) + \rho(B) \tag{7.2}$$
$$\rho(AB) = \rho(A)\rho(B) \tag{7.3}$$

The unit scalar 1 in the geometric algebra is always mapped to the identity matrix:

$$\rho(1) = I_{n\times n}. \tag{7.4}$$

Generally, we can specify the rest of the representation by specifying what matrices the basis vectors of the geometric algebra map to.

7.1.1 Pauli representation

For instance, the Pauli representation is an isomorphism ρ between Cl(3) and Mat($2\times 2, \mathbb{C}$). The basis vectors e_1, e_2, e_3 are mapped to the Pauli matrices:

Pauli matrix representation

$$\rho : \mathrm{Cl}(3) \to \mathrm{Mat}(2\times 2, \mathbb{C}) \tag{7.5}$$

$$\rho(e_1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{7.6}$$

$$\rho(e_2) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{7.7}$$

$$\rho(e_3) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.8}$$

Using the isomorphism property, we can determine the matrices belonging to the other basis k-vectors:

$$\rho(e_{12}) = \rho(e_1)\rho(e_2) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \tag{7.9}$$

$$\rho(e_{23}) = \rho(e_2)\rho(e_3) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \tag{7.10}$$

$$\rho(e_{31}) = \rho(e_3)\rho(e_1) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{7.11}$$

$$\rho(I) = \rho(e_1)\rho(e_2)\rho(e_3) = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \tag{7.12}$$

The multivector reverse $\tilde{M}$ can be performed by taking the hermitean conjugate of the corresponding matrix:

$$\rho(\tilde{M}) = \rho(M)^\dagger. \tag{7.13}$$

This makes perfect sense - the hermitean conjugate of a matrix product reverses its order, and the matrices $\rho(e_i)$ representing the basis vectors are all hermitean. Therefore, for instance:

$$(\rho(e_1)\rho(e_2))^\dagger = \rho(e_2)^\dagger\rho(e_1)^\dagger \tag{7.14}$$

$$= \rho(e_2)\rho(e_1) \tag{7.15}$$

There is another important operation to derive: The scalar grade projection $\langle\cdots\rangle$. First, we notice that the trace of the matrices representing multivectors is almost always zero:

$$\mathrm{tr}(\rho(1)) = 2 \tag{7.16}$$

$$\mathrm{tr}(\rho(e_i)) = 0 \tag{7.17}$$

$$\mathrm{tr}(\rho(e_{ij})) = 0 \tag{7.18}$$

$$\mathrm{tr}(\rho(I)) = 2i \tag{7.19}$$

This means that we can extract the scalar grade of a multivector by:

Scalar grade projection

$$\langle M \rangle = \frac{1}{2} \operatorname{Re} \operatorname{tr}(\rho(M)) \tag{7.20}$$

We can now do calculations with the complex 2x2 matrices as if they were 3D multivectors. In fact, you shouldn't think of e.g. the bivector e_{12} and the matrix $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ as two separate things - the matrix is simply a different way to write the multivector e_{12}. Therefore, once we have chosen a specific matrix representation, we will treat a multivector M and its matrix representation $\rho(M)$ as the same thing and write e.g.

$$e_3 = \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix} \tag{7.21}$$

$$e_{12} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{7.22}$$

We see that the Pauli matrices are just a way of doing calculations with 3D multivectors using matrices. There is, however, an open question: In quantum mechanics, Pauli matrices are mainly used to operate on **spin states** - also called **spinors** - like

$$|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix}, \qquad a, b \in \mathbb{C}. \tag{7.23}$$

How do these spin states fit in the picture? They can't be multivectors, because they're not 2×2 matrices. We shall find out soon - but first, we will take a brief look at the STA equivalent of the Pauli representation.

7.1.2 STA representation

The geometric algebra of spacetime $\mathrm{Cl}(1, 3)$ has much more elements than the geometric algebra of space. Thus, we need larger matrices to represent it. This is what the Dirac gamma matrices are for. They don't just represent the normal STA, but the **complexified STA** - the spacetime algebra where multivectors can have complex instead of real coefficients.

Gamma matrix representation of the STA

$$\rho : \mathrm{Cl}(1, 3, \mathbb{C}) \to \mathrm{Mat}(4 \times 4, \mathbb{C}) \tag{7.24}$$

$$\rho(\gamma_0) = \begin{pmatrix} & & 1 & 0 \\ & & 0 & 1 \\ 1 & 0 & & \\ 0 & 1 & & \end{pmatrix} \qquad \rho(\gamma_1) = \begin{pmatrix} & & 0 & -1 \\ & & -1 & 0 \\ 0 & 1 & & \\ 1 & 0 & & \end{pmatrix} \tag{7.25}$$

$$\rho(\gamma_2) = \begin{pmatrix} & & 0 & i \\ & & -i & 0 \\ 0 & -i & & \\ i & 0 & & \end{pmatrix} \qquad \rho(\gamma_3) = \begin{pmatrix} & & -1 & 0 \\ & & 0 & 1 \\ 1 & 0 & & \\ 0 & -1 & & \end{pmatrix} \tag{7.26}$$

Some authors also use a slightly different representation in which γ_0 is represented by the matrix

$$\begin{pmatrix} I_{2\times 2} & 0 \\ 0 & -I_{2\times 2} \end{pmatrix} \tag{7.27}$$

Our representation is called the **Weyl** or **chiral** representation, while the representation using the above matrix for γ_0 is called the **Dirac** representation for historical reasons. (Yes, the terminology is confusing and messy. It's because most physicists can't be bothered to do their jobs and find good names for the stuff they invent.)

Again, the other basis k-vectors can be constructed by e.g.

$$\gamma_{12} = \rho(\gamma_1)\rho(\gamma_2) = \begin{pmatrix} -i & 0 & & \\ 0 & i & & \\ & & i & 0 \\ & & 0 & -i \end{pmatrix} \tag{7.28}$$

Just as in the 3D case, we will say that the multivectors like γ_{12} and matrices like $\rho(\gamma_1)\rho(\gamma_2)$ are just two ways to depict the same object and use the notation γ_{12} for both. We will thus drop the ρ's from our notation again.

In the Pauli matrix representation, we could form the multivector reverse by taking the hermitean conjugate $M^\dagger$ of the matrix representing the multivector. This does not work as easily anymore for the Dirac matrices. If we conjugate them, we find that:

$$(\gamma_0)^\dagger = \gamma_0 \tag{7.29}$$

$$(\gamma_i)^\dagger = -\gamma_i. \tag{7.30}$$

This should ring a bell. It is a parity flip with respect to the time axis γ_0. The formula for conducting such a parity flip on a multivector M is:

$$P_{\gamma_0}(M) = \gamma_0 M \gamma_0. \tag{7.31}$$

Note that in four dimensions, there is more than one possible parity flip. The statement "We leave the time coordinate invariant and flip the sign of all spatial coordinates" is dependent on how exactly we define the time coordinate. Observers with different velocities would therefore define parity flips differently. This is why we added the γ_0 index to P here - it indicates that the parity flip is done w.r.t γ_0. In other words, the parity flip P_{γ_0} is the parity flip an observer moving with the four-velocity $U = \gamma_0$ would naturally dois the parity flip an observer moving with the four-velocity $U = \gamma_0$ would naturally do.

We have gotten ourselves into quite some trouble right now! We just wanted to reverse the order of multiplication and ended up introducing an extra parity flip with the hermitean conjugate $M^\dagger$. Even worse, this parity flip is not covariant! If we conduct the hermitean conjugate $M^\dagger$ in two different coordinate systems, we will get two different results, because the parity flip is .

The solution is surprisingly simple: $M^\dagger$ reverses the order multiplication $\tilde{M}$ and does an additional parity flip w.r.t. the current time axis:

$$M^\dagger = \gamma_0 \tilde{M} \gamma_0 \tag{7.32}$$

We only want the former, so we just conduct another parity flip:

$$\tilde{M} = \gamma_0 M^\dagger \gamma_0 = \gamma_0(\gamma_0 \tilde{M} \gamma_0)\gamma_0. \tag{7.33}$$

This is how we evaluate the multivector reverse when we are working with a matrix representation. But remember - the actual definition of $\tilde{M}$ does not involve γ_0. What we are actually doing here is cleaning up the mess of the hermitean conjugate.

Because the hermitean conjugate behaves so weirdly for STA representations, we will avoid it from now on as best as we can. Just remember - when you are given a matrix representing an STA multivector and want to take its reverse, first take the hermitean conjugate, and then do another parity flip with the current γ_0 to cancel it out.

The spinors associated with the Dirac gamma matrices have four complex components:

$$|\psi\rangle = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}, \qquad a, b, c, d \in \mathbb{C} \tag{7.34}$$

They are not STA multivectors, because they aren't 4×4 matrices. Again, the same question arises - what are they? How do they fit into the picture of the Dirac matrices representing multivectors? We can interpret multivectors geometrically, but what about spinors? How can we picture spinors, geometrically?

7.2 Fundamenentals of spinors

To answer the central question of this chapter "what is a spinor?" - we will first have to clear up some confusion about matrices and vectors.

7.2.1 What is a vector?

In physics, there are two types of objects that are commonly called "vectors".

- "Vectors" in the mathematical sense: An element of a vector space, i.e. a space that has an addition and a distributive scalar multiplication. This type of "vector" can be anything that behaves linearly in some way - quantum-mechanical states, functions, spinors, tensors, etc. These "vectors" can be represented by a column of numbers - for instance, spinors

$$|\psi\rangle = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}. \tag{7.35}$$

- Vectors in the physical sense: Grade-1 elements of a geometric algebra, aka a length with an orientation. Roughly speaking, arrows pointing through space (spacetime in STA).

The second definition is a lot more specific than the first one. So far, we have only used the second definition in this script. To avoid confusion, we are going to rename "vectors" in the mathematical sense to **lineals** and vector spaces to **linear spaces**. It is important to strictly distinguish between these concepts.

If we make a change in basis U (aka a transformation), our lineals l and matrices M transform like

$$l \to U^{-1} l \tag{7.36}$$

$$M \to U^{-1} l U \tag{7.37}$$

The sandwich transformation law for matrices ensures that the product of a lineal and a matrix Ml transforms like a lineal:

$$Ml \to U^{-1} M U U^{-1} l = U^{-1}(Ml). \tag{7.38}$$

7.2.2 Going beyond vector matrices and vector tensors

Matrices are linear maps acting on lineals. The first lineals we physicists encounter in our studies are vectors, and the first matrices we encounter are **vector matrices** - matrices that map vectors onto vectors. Such vector matrices describe geometric transformations like rotations, reflections, scalings, stretchings, or projections. Similarly, all tensors we have handled so far are **vector tensors** - tensors over the space of vectors. They transform under the action of vector matrices, like in (3.84).

But of course, we can easily imagine matrices acting on different kinds of lineals! We could imagine **bivector matrices** acting on bivectors. Bivectors are lineals too, and the inertia tensor actually is such a bivector matrix - it maps the angular velocity bivector onto the angular momentum bivector.

But most importantly, the **spinors** are lineals, too. We act on Pauli spinors with the matrices representing Cl(3) multivectors, and we act on Dirac spinors with the matrices representing $\mathrm{Cl}(1,3)$ multivectors. Thus, we can say that multivectors are **spinor matrices**.

The rotor transformation law for a multivector M is:

$$M \to \tilde{R}MR. \tag{7.39}$$

Now, we remember that $R\tilde{R} = 1$ for all rotors. This means that

$$\tilde{R} = R^{-1}. \tag{7.40}$$

We can thus reformulate the rotor law as

$$M \to R^{-1}MR. \tag{7.41}$$

This looks just like the matrix transformation law (7.37). On the one hand, this shouldn't surprise us - we've seen that multivectors are (can be represented as) spinor matrices, so it's natural that they should follow the sandwich matrix transformation law. On the other hand, this is deeply surprising - we derived the rotor law from purely geometric considerations! Clearly, we're onto something big here. So let's extend this analogy - by extension, spinors $|\psi\rangle$ transform like

Spinor transformation law

$$|\psi\rangle \to R^{-1}|\psi\rangle. \tag{7.42}$$

Acknowledgements

The following people have contributed to the Heidelberg Geometric Algebra lecture so far:

- **Dr. Maarten DeKieviet** `<maarten@physi.uni-heidelberg.de>`: Leader of the GA-HD working group
- **Vanessa Alexandra Hollmeier** `<vhollmeier@physi.uni-heidelberg.de>`: Chapters "Introduction", "Basics of geometric algebra", "Geometric algebra of space", "Tensor notation", "Electrodynamics", "Spinors", technical infrastructure for the project
- **Amir Leidel** `<amir.leidel@stud.uni-heidelberg.de>`: First draft of section on Cl(2), chapter "Geometric calculus"