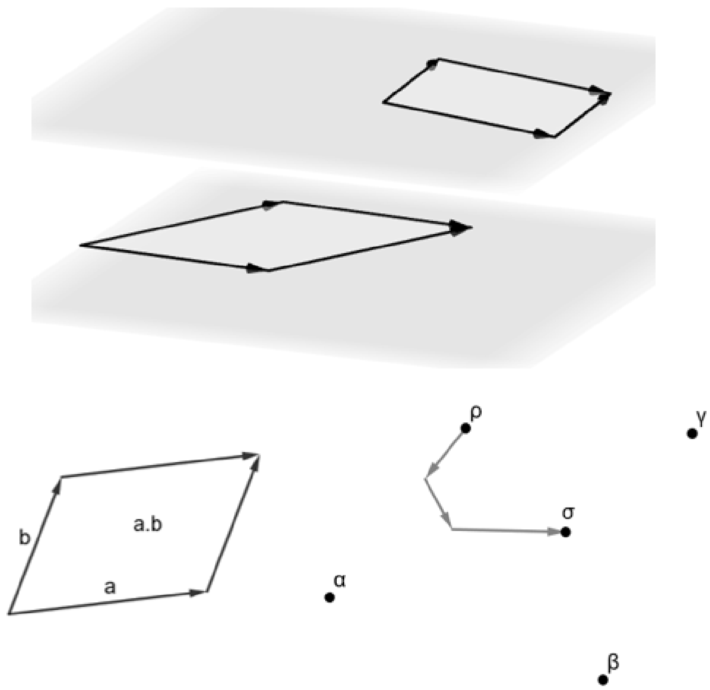
# History of affine spaces

A case study in the formation of modern mathematics

**Nim Steen**




Supervised by Viktor Blåsjö

Department of Mathematics
Utrecht University

# Contents

# Chapter 1

# Introduction

Over the course of the 19th and 20th century, mathematics moved towards an increasingly abstract approach, away from tangible and intuitive methods, and towards the general and axiomatic. In some areas of mathematics, this journey has been well documented, we know who contributed and what their motivations were to do so. In others this history is not so clear. One such area where its development is rather unclear is that of affine spaces as a generalization of geometry and a distinct topic within linear algebra.

The mathematical term affine was initially introduced by Euler in 1748 [7]. Almost a century later, Mobiüs uses the term in his 1827 'Barycentrische Calcul' [14], although he merely uses it to describes a geometric relation in line with congruent and similar figures. For almost a century, it is only this aspect of affinity, along with the transformations that preserve affinity, that gets studied and further developed. The first person to work out affine geometry as a whole in detail is Hermann Weyl [22], in his 1918 work 'Space, Time, Matter' [21]. In this work however, he himself credits an "epoch-making" 1844 work for the systematic treatment of affine geometry in higher dimensions, the 'Ausdehnungslehre' by Hermann Grassmann [8]. Some modern mathematicians also recognize the affine approach in this work, and credit him as the first to develop this area of mathematics [13, 19]. Others are somewhat more critical, stating for instance that Grassmann had created 'empty generalizations', which had yet to be given meaning by later mathematicians [23]. During his lifetime Grassmann's work went largely unnoticed, and it was only towards the end of the 19th century that his work started to gain recognition by mathematicians such as Klein, Lie and Peano [6, 12, 16].

In this thesis, I aim to shed some light on the development of affine geometry. I will focus in particular on Grassmann's work, to dissect what aspects of affine geometry he had developed and what has been over-interpreted by modern readers. To do so, I will start by establishing what I will consider as the 'modern' approach to affine geometry in chapter 2. This will form a reference point to compare the historic approaches to. After this I will go over the approach by Weyl in chapter 3, as his has been credited as the first modern approach. After analyzing what an early modern approach could look like I will move on to Grassmann. In chapter 4 I will go into detail on some of the mathematics Grassmann developed in his Ausdehnungslehre, and attempt to relate it to modern concepts. I will focus in particular on the concepts of points and vectors and how these interact, as well as the higher dimensional spaces he works in and their coordinate systems. In chapter 5 I will discuss the concepts of quotients that Grassmann develops in the second edition of his work. These quotients can take the role of a transformation matrix, and share many properties of square matrices like that of a determinant and eigenvalues. The concept behind them is however rather different, and as such they are an excellent topic for a case study on how Grassmann's approach differs from the modern one. I will use this case study in particular to discuss some of the advantages and disadvantages of Grassmann's approach compared to the modern one.

In order to accurately represent the work done by Weyl and Grassmann, I will often cite segments directly from English translations of their original works before discussing them further. These citations are placed in textboxes to separate them from my original text. Hence, unless stated otherwise, everything contained in textboxes is a direct citation, with the exception of the equation numbering which follows the numbering of this thesis instead. For Weyl's 'Space, Time, Matter' I have used the 1922 translation by Henry L. Brose [21]. In the case of Grassmann I will reference two works, the first being his 1844 'Die lineale Ausdehnungslehre' (from here on referred to as A1) and the second his 1862 'Die Ausdehnungslehre' (from here on referred to as A2) in which he shared his theory in a different format. For both of these I have used translations by Lloyd C. Kannenberg [8, 10].

Grassmann's work contains a lot of unfamiliar terms, both because he wrote his work before vector analysis was fully developed, and because his work outside of mathematics left him limited time to stay up to date with the language of his contemporaries. This made it a difficult read not just for his contemporaries, but also for the modern reader. However, to distinguish between Grassmann's concepts and the modern ones, I will often be using his terms while discussing his work. In appendix A I have included a list of some of the more commonly used terms and their modern interpretation, to serve as a reading aid.

I would like to thank my thesis supervisor, Viktor Blåsjö, for suggesting this interesting topic, the regular feedback and interesting discussions, as well as the support when I was unsure of how to progress on my thesis.

# Chapter 2

# Modern definition of affinity

In this thesis I will discuss several historical views of affinity, and compare them to a modern perspective. However even from our modern perspective there are different ways one can approach the topic of affinity, so for clarity and ease of comparison, I will fix one such definition here to refer back to throughout the rest of the document.

## 2.1 Affine space

**Definition 2.1.1** (Affine space)**.** [20] Let $k$ be a field, and $\mathbb{V}$ be a $k$-vector space. An *affine space* over $\mathbb{V}$ is a set $\mathbb{A}$ together with an operation

$$\begin{aligned} \mathbb{A} \times \mathbb{V} &\to \mathbb{A} \\ P, \vec{v} &\mapsto P \oplus \vec{v} \end{aligned}$$

satisfying the following axioms:

1. (Right identity) For all $P \in \mathbb{A}$, $P \oplus \vec{0} = P$.
2. (Associativity) For all $P \in \mathbb{A}, \vec{v}, \vec{w} \in \mathbb{V}$, $P \oplus (\vec{v} + \vec{w}) = (P \oplus \vec{v}) \oplus \vec{w}$.
3. (Subtraction) For any $P, Q \in \mathbb{A}$ there exists a unique $\vec{v} \in \mathbb{V}$ such that $P \oplus \vec{v} = Q$. I will write $\overrightarrow{PQ}$ to denote this vector.

I will call the elements of $k$ scalars, the elements of $\mathbb{V}$ vectors and the elements of $\mathbb{A}$ points.

Now this definition of affine space has the axioms of vector space, such as the existence of an additive identity, implied in it. Since these had not yet been established in the sources I will discuss, I will also include this definition for ease of comparison:

**Definition 2.1.2** (Vector space)**.** [18] Let $k$ be a field. A *vector space* over $k$ is a nonempty set $\mathbb{V}$ together with two operations

$$\begin{aligned} \mathbb{V} \times \mathbb{V} &\to \mathbb{V} && \text{(addition)} \\ \vec{v}, \vec{w} &\mapsto \vec{v} + \vec{w} \\ k \times \mathbb{V} &\to \mathbb{V} && \text{(scalar multiplication)} \\ k, \vec{v} &\mapsto k\vec{v} \end{aligned}$$

satisfying the following axioms:

1. For all $\vec{u}, \vec{v}, \vec{w} \in \mathbb{V}$, $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$
2. For all $\vec{v}, \vec{w} \in \mathbb{V}$, $\vec{v} + \vec{w} = \vec{w} + \vec{v}$
3. There exists a $\vec{0} \in \mathbb{V}$ such that for all $\vec{v} \in \mathbb{V}$, $\vec{0} + \vec{v} = \vec{v} + \vec{0} = \vec{v}$
4. For each $\vec{v} \in \mathbb{V}$ there exists $\vec{w} \in \mathbb{V}$ such that $\vec{v} + \vec{w} = \vec{w} + \vec{v} = \vec{0}$.
5. For all $a \in k, \vec{v}, \vec{w} \in \mathbb{V}$, $a(\vec{v} + \vec{w}) = a\vec{v} + a\vec{w}$
6. For all $a, b \in k, \vec{v} \in \mathbb{V}$, $(a + b)\vec{v} = a\vec{v} + b\vec{v}$
7. For all $a, b \in k, \vec{v} \in \mathbb{V}$, $(ab)\vec{v} = a(b\vec{v})$
8. For all $\vec{v} \in \mathbb{V}$, $1\vec{v} = \vec{v}$

## 2.2 Affine transformation

**Definition 2.2.1** (Affine transformation)**.** [20] Let $\mathbb{A}$ be an affine space over the $k$-vector space $\mathbb{V}$. An affine transformation on $\mathbb{A}$ is a bijective map $f : \mathbb{A} \to \mathbb{A}$ which induces a linear map $\tilde{f} : \mathbb{V} \to \mathbb{V}$ such that for any point $P$ and any vector $\vec{v}$,

$$f(P \oplus \vec{v}) = f(P) \oplus \tilde{f}(\vec{v})$$

As a linear map, $\tilde{f}$ also satisfies the properties that for all $\vec{v}, \vec{w} \in \mathbb{V}, a \in k$:

$$\begin{aligned} \tilde{f}(\vec{v} + \vec{w}) &= \tilde{f}(\vec{v}) + \tilde{f}(\vec{w}) \\ \tilde{f}(a\vec{v}) &= a\tilde{f}(\vec{v}) \end{aligned}$$

In other words, an affine transformation is a bijective endomorphism on an affine space with an associated map on its vector space such that the two maps preserve the affine properties.

The study of affine geometry, then, is the study of those properties that remain invariant under affine transformations.

# Chapter 3

# Hermann Weyl's 'Space, Time, Matter'

One of the first modern approaches to affine space is that of Hermann Weyl [22], in his 1918 'Space, Time, Matter' [21]. In this book, Weyl introduces affine geometry as a mathematical framework for relativity theory. In the following sections, I will present Weyl's affine geometry and how it compares to our current understanding. Almost all of this is found in §2 of his book, aptly titled "The Foundations of Affine Geometry".

## 3.1 The foundation: points and vectors

Weyl begins his section on affine geometry by describing the objects of study, being vectors and points, and some of their properties, after which he lists the axioms they should satisfy. He seems to work in the context of space as a given, in which we naturally have points which seem to need no further definition. Any displacement or translation is then named a vector, and quickly associated with points, introducing the notation $\overrightarrow{PQ}$ for the displacement that "transfers the point $P$ to the point $Q$". After this, addition is introduced as the result of successive translations. Multiplication and division by an integer are then derived from this concept of addition, and inverses and the nil-vector with their familiar properties are introduced. Following that the concept of multiplication is extended to fractional scalars, and finally through continuity to any real scalar.

After this brief intuitive introduction the axioms are introduced, first those of vectors (separated into addition laws and scalar multiplication laws), and then those of points, or rather those of the interaction between points and vectors. Weyl takes great care in separating points and vectors. They have their own axioms and operations, vectors collectively form a vector field while points collectively form a point-configuration, and vectors have components while points have co-ordinates dependent on a chosen co-ordinate system. Although the two are similar, co-ordinates depend on the origin of the system while components do not. In properties that hold for both, he is careful to state them both and not identify the two. At the same time, they are intimately connected. The relation that a vector transfers one point to another is introduced in the very first paragraph of §2, and it takes an important role in Weyl's view of affine transformations as we shall see in section 3.4.

### 3.1.1 The axioms

Weyl first presents his axioms for vectors, or rather those of vector addition and scalar multiplication, and later follows with those for points and vectors. His axioms for vectors are as follows:

> **Vector axioms, addition and multiplication [21, p. 17]**
>
> Two vectors $\mathfrak{a}$ and $\mathfrak{b}$ uniquely determine a vector $\mathfrak{a} + \mathfrak{b}$ as their sum. A number $\lambda$ and a vector $\mathfrak{a}$ uniquely define a vector $\lambda\mathfrak{a}$, which is "$\lambda$ times $\mathfrak{a}$" (multiplication). These operations are subject to the following laws:—
>
> ($\alpha$) Addition—
>
> (1) $\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}$ (Commutative Law).
>
> (2) $(\mathfrak{a} + \mathfrak{b}) + \mathfrak{c} = \mathfrak{a} + (\mathfrak{b} + \mathfrak{c})$ (Associative Law).
>
> (3) If $\mathfrak{a}$ and $\mathfrak{c}$ are any two vectors, then there is one and only one value of $\mathfrak{x}$ for which the equation $\mathfrak{a} + \mathfrak{x} = \mathfrak{c}$ holds. It is called the difference between $\mathfrak{c}$ and $\mathfrak{a}$ and signifies $\mathfrak{c} - \mathfrak{a}$ (Possibility of Subtraction).
>
> ($\beta$) Multiplication—
>
> (1) $(\lambda + \mu)\mathfrak{a} = (\lambda\mathfrak{a}) + (\mu\mathfrak{a})$ (First Distributive Law).
>
> (2) $\lambda(\mu\mathfrak{a}) = (\lambda\mu)\mathfrak{a}$ (Associative Law).
>
> (3) $1\mathfrak{a} = \mathfrak{a}$.
>
> (4) $\lambda(\mathfrak{a} + \mathfrak{b}) = (\lambda\mathfrak{a}) + (\lambda\mathfrak{b})$ (Second Distributive Law).

These axioms are almost identical to our modern axioms of a vector space as stated in definition 2.1.2. The only notable differences are the inclusion of a "Possibility of Subtraction" axiom, and the absence of modern axioms 3. and 4. (existence of additive identity and additive inverse). However, the Possibility of Subtraction axiom is equivalent to these two modern axioms: We know that it follows from the axioms of a vector space that the inverse of any vector is unique. It follows, then, that for any two vectors $\vec{a}, \vec{c}$ there is a unique vector $\vec{x} = \vec{c} + -\vec{a}$ such that $\vec{a} + \vec{x} = \vec{c}$. Conversely, from Possibility of Subtraction one can obtain the existence of additive identity and inverses. First, for some vector $\mathfrak{a}$ it implies the existence of a vector that I will suggestively call $\mathfrak{o}$, such that $\mathfrak{a} + \mathfrak{o} = \mathfrak{a}$. For any other vector $\mathfrak{b}$ we know there exists some $\mathfrak{x}$ such that $\mathfrak{a} + \mathfrak{x} = \mathfrak{b}$. We then get that

$$\begin{aligned}
\mathfrak{b} + \mathfrak{o} &= (\mathfrak{a} + \mathfrak{x}) + \mathfrak{o} && \text{(def. of } \mathfrak{x}\text{)} \\
&= (\mathfrak{x} + \mathfrak{a}) + \mathfrak{o} && \text{(Commutative Law)} \\
&= \mathfrak{x} + (\mathfrak{a} + \mathfrak{o}) && \text{(Associative Law)} \\
&= \mathfrak{x} + \mathfrak{a} && \text{(def. of } \mathfrak{o}\text{)} \\
&= \mathfrak{a} + \mathfrak{x} && \text{(Commutative Law)} \\
&= \mathfrak{b} && \text{(def. of } \mathfrak{x}\text{)}
\end{aligned}$$

Hence, $\mathfrak{o}$ serves as identity. Using Possibility of Subtraction with any vector $\mathfrak{a}$ and $\mathfrak{o}$ then easily gives us the existence of an inverse $\mathfrak{o} - \mathfrak{a} = -\mathfrak{a}$.

A little further, Weyl states one more vector axiom, namely his "Axiom of Dimensionality", which is conceptually different from how we treat modern vector spaces. Namely, he postulates:

Axiom of Dimensionality [21, p. 19]

**There are $n$ linearly independent vectors, but every $n+1$ are linearly dependent on one another,**
or: The vectors constitute an $n$-dimensional linear manifold.

He states this axiom later as he first has to establish the concept of linear dependence and dimensions, which I will discuss in section 3.2. We can however already observe how the inclusion of this axiom changes the concept of Weyl's spaces compared to our modern view. First of all, the formulation of this axiom betrays a slightly different concept of axiomatization than we typically use. Rather than stating the properties that hold for any vector space, Weyl seems to take an arbitrary but specific space in mind before stating the axioms it satisfies. A more modern statement of this axiom might have been:

> "For any vector space $\mathbb{V}$, there exists a positive integer $n$ such that there exist $n$ linearly independent vectors in $\mathbb{V}$, but every $n+1$ vectors are linearly dependent on one another."

More importantly, this axiom requires that any space has some integer dimension, and therefore rules out infinite-dimensional spaces. Also, although it is not explicitly stated, the further discussion around these axioms suggests that he considers all scalars to be real numbers, hence his concept of vector space is limited to $\mathbb{R}^n$. For the purpose of his work this restriction is very reasonable. In this chapter Weyl is establishing a vector system to discuss space-time, and finite dimensions with real scalars suffice for this purpose.

With the vector axioms out of the way, Weyl proceeds to present his axioms for points and vectors:

Points and Vectors axioms [21, p. 18]

1. Every pair of points $A$ and $B$ determines a vector $\mathfrak{a}$; expressed symbolically $\overrightarrow{AB} = \mathfrak{a}$. If $A$ is any point and $\mathfrak{a}$ any vector, there is one and only one point $B$ for which $\overrightarrow{AB} = \mathfrak{a}$.
2. If $\overrightarrow{AB} = \mathfrak{a}, \overrightarrow{BC} = \mathfrak{b}$, then $\overrightarrow{AC} = \mathfrak{a} + \mathfrak{b}$.

After which he summarizes the objects and relations that occur, which will become relevant when we consider affine transformations:

Fundamental categories of objects, fundamental relations [21, p. 18]

In these axioms two fundamental categories of objects occur, viz. points and vectors; and there are three fundamental relations, those expressed symbolically by—

$$\mathfrak{a} + \mathfrak{b} = \mathfrak{c} \qquad \mathfrak{b} = \lambda\mathfrak{a} \qquad \overrightarrow{AB} = \mathfrak{a} \tag{3.1}$$

The axioms on points and vectors are considerably less concise than those on vectors. The first one in particular does double duty as both a definition of the relation $\overrightarrow{AB} = \mathfrak{a}$ and the axiom of the uniqueness of $B$ in that relation given $A$ and $\mathfrak{a}$. Furthermore, where a modern affine space is typically defined through the addition relation $\oplus : A \oplus \mathfrak{a} = B$, Weyl instead uses the relation $\overrightarrow{AB} = \mathfrak{a}$.

This alternate definition with Weyl's axioms is also used by some modern sources (cf. [15, p. 7]) and is indeed equivalent, as I will show below.

In order to prove the equivalence between Weyl's axioms and the modern definition of affine space in definition 2.1.1, we must first establish the $\oplus$ addition in Weyl's space. Weyl's first axiom on points and vectors easily allows us to define this: Let $A \oplus \mathfrak{a}$ be that unique point $B$ such that $\overrightarrow{AB} = \mathfrak{a}$. Adopting this notation, Weyl's axioms now read:

1. Every point $A$ with any vector $\mathfrak{a}$ determines a point $B$, denoted $A \oplus \mathfrak{a} = B$. If $A$ and $B$ are any two points, there is one any only one vector $\mathfrak{a}$ such that $A \oplus \mathfrak{a} = B$.
2. If $A \oplus \mathfrak{a} = B, B \oplus \mathfrak{b} = C$, then $A \oplus (\mathfrak{a} + \mathfrak{b}) = C$.

First, I will show that these axioms can be deduced from the modern ones. The first axiom now starts simply with the definition of $\oplus$ as an operation instead. Its second part is equivalent to the affine axiom of subtraction. The second axiom follows from the associative axiom of affine spaces:

$$\begin{aligned} A \oplus (\mathfrak{a} + \mathfrak{b}) &= (A \oplus \mathfrak{a}) + \mathfrak{b} \\ &= B + \mathfrak{b} \\ &= C \end{aligned}$$

The reverse is a little bit more involved, as we have to show that the axioms of right identity, associativity and subtraction follow from Weyl's axioms. Subtraction is stated precisely in Weyl's first axiom, so right identity and associativity remain.

For right identity, let $A, B$ be any two points. Then by Weyl's first axiom there exist unique vectors $\mathfrak{a}, \mathfrak{b}$ such that $A \oplus \mathfrak{a} = A$ and $B \oplus \mathfrak{b} = A$. Using Weyl's second axiom, we get that $B \oplus (\mathfrak{b} + \mathfrak{a}) = A$. But then by definition of $\mathfrak{b}$ we get that $\mathfrak{b} = \mathfrak{b} + \mathfrak{a}$, hence, by the vector axioms, $\mathfrak{a} = \mathfrak{o}$. Hence for any point $P$ we have that $\mathfrak{p} = \mathfrak{o}$ for that unique vector such that $P \oplus \mathfrak{p} = P$, thus necessarily that $P \oplus \mathfrak{o} = P$.

Finally, for associativity, Let $A$ be any point and $\mathfrak{a}, \mathfrak{b}$ be any vectors, and suppose $A \oplus \mathfrak{a} = B$ and $B \oplus \mathfrak{b} = C$. Then we have by Weyl's second axiom:

$$\begin{aligned} A \oplus (\mathfrak{a} + \mathfrak{b}) &= C \\ &= B \oplus \mathfrak{b} \\ &= (A \oplus \mathfrak{a}) \oplus \mathfrak{b} \end{aligned}$$

So indeed, with some change in notation, Weyl's affine space is equivalent to our modern concept of an affine space over $\mathbb{R}^n$, and in fact his axioms represent any affine space where the associated vector space is finite-dimensional.

## 3.2 Bases, coordinates and dimensions

In the section above I stated Weyl's axiom of dimensionality without establishing his concept of dimensions. In his work, he presents this axiom a little later, so as to first define dimensions using the other axioms. He once again starts with an intuitive introduction, demonstrating how a straight line, a plane and "all space" can be formed from a point and several vectors. After this introduction he extends this to the formation of a higher dimensional space.

Starting from an arbitrary point $O$ and any non-zero vector $\mathfrak{e}_1$, Weyl forms a straight line as the collection of points $P$ where $\overrightarrow{OP} = \xi_1 \mathfrak{e}_1$ for some real number $\xi_1$. Finding some arbitrary vector $\mathfrak{e}_2$ that is not of the form $\xi_1 \mathfrak{e}_1$ and considering all $P$ such that $\overrightarrow{OP} = \xi_1 \mathfrak{e}_1 + \xi_2 \mathfrak{e}_2$ grants us a plane, conceptualized from the affine perspective as obtained by "sliding one straight line along another". Continuing this process with a third vector (not of the form $\xi_1 \mathfrak{e}_1 + \xi_2 \mathfrak{e}_2$), he then obtains "all space".

In this intuitive build up Weyl relies on the intuition of a 'given' three-dimensional space, but he lets go of this restriction for the definitions that follow. First he introduces the general concept of linear independence:

**Definition: linear independence, $h$-dimensional linear vector-manifold [21, p. 19]**

A finite number of vectors $\mathfrak{e}_1, \mathfrak{e}_2, \ldots \mathfrak{e}_h$ is said to be **linearly independent** if

$$\xi_1 \mathfrak{e}_1 + \xi_2 \mathfrak{e}_2 + \cdots + \xi_h \mathfrak{e}_h \tag{3.2}$$

only vanishes when all the coefficients $\xi$ vanish simultaneously.
With this assumption all vectors of the form (3.2) together constitute a so-called **h-dimensional linear vector-manifold** (or simply vector-field); in this case it is the one mapped out by the vectors $\mathfrak{e}_1, \mathfrak{e}_2, \ldots \mathfrak{e}_h$.

With linear (in)dependence established, Weyl can introduce coordinate systems:

**Definition: co-ordinate system, components, co-ordinates [21, p. 20]**

A point $O$ in conjunction with $n$ linear independent vectors $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_n$ will be called a co-ordinate system (**C**). Every vector $\mathfrak{x}$ can be presented in one and only one way in the form

$$\mathfrak{x} = \xi_1 \mathfrak{e}_1 + \xi_2 \mathfrak{e}_2 + \cdots + \xi_n \mathfrak{e}_n \tag{3.3}$$

The numbers $\xi_i$ will be called its **components** in the co-ordinate system (**C**). If $P$ is any arbitrary point and if $\overrightarrow{OP}$ is equal to the vector (3.3), then the $\xi_i$ are called the **co-ordinates** of $P$.

The above definitions are essentially equivalent to the modern ones, aside from a few concepts that are given a different name. The linearly independent vectors $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_n$ form a basis, which, together with an origin $O$ make up the co-ordinate system. Each vector and point then has components or co-ordinates that uniquely identify them in said co-ordinate system. Weyl also notes that in affine geometry, there is nothing unique about any one co-ordinate system. He proceeds to show that any other co-ordinate system $O'; \mathfrak{e}'_1, \mathfrak{e}'_2, \ldots, \mathfrak{e}'_n$ is equivalent and shows how one can move from one system to another, that is, how one can chance basis and origin, which I will further discuss in section 3.3.

Weyl seems to be fairly comfortable using arbitrary (finite) dimensions, but is aware that this may not be true for his readers. Before introducing $h$-dimensional manifolds he first uses the line, the plane and space to build some intuition of how they would work, and right after his axiom of dimensionality, he reminds his readers that $n = 1, 2, 3$ gives us affine geometry of the straight line, the plane and space respectively. He justifies his $n$-dimensional approach by stating that from a mathematical perspective, this restriction to three dimensions appears to be 'accidental', and should therefore not be used to develop geometry systematically. The straight line, plane and space are to be considered just as special cases.

## 3.3 Change of basis

With the co-ordinate systems established, Weyl explains how one can find the co-ordinates of some point or components of some vector in a co-ordinate system $(O; \mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_n)$, given the co-ordinates or components in another system $(O'; \mathfrak{e}'_1, \mathfrak{e}'_2, \ldots, \mathfrak{e}'_n)$. That is, he derives a change-of-basis formula. In modern notation we typically use matrix multiplication for the change of basis, together with vector addition for the change of origin. Later in his work, Weyl demonstrates some use of this approach. For the introduction of change-of-basis formulas however, he instead uses somewhat more cumbersome summations. Aside from the notation, though, the formulas and their derivations are essentially identical to the modern approach.

On page 21, Weyl derives them as follows: Since the $\mathfrak{e}_i$ form a co-ordinate system, we have equations of the form

$$\mathfrak{e}'_i = \sum_{k=1}^{n} a_{ki}\mathfrak{e}_k \tag{3.4}$$

for all $\mathfrak{e}'_i$, where the $a_{ki}$ form a 'number system' with non-vanishing determinant since the $\mathfrak{e}'_i$ are linearly independent. Now, suppose that $\xi_i$ are the components for some vector $\mathfrak{x}$ in the system $(O; \mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_n)$, and $\xi'_i$ the components of that same vector in the system $(O'; \mathfrak{e}'_1, \mathfrak{e}'_2, \ldots, \mathfrak{e}'_n)$. Substituting the equations (3.4) into $\sum_i \xi_i \mathfrak{e}_i = \sum_i \xi'_i \mathfrak{e}'_i$, he derives the relation

$$\xi_i = \sum_{k=1}^{n} a_{ik}\xi'_i \tag{3.5}$$

Similarly, let $x_i$ be the co-ordinates of a point $P$ in $(O; \mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_n)$, $x'_i$ the co-ordinates of the same point in $(O'; \mathfrak{e}'_1, \mathfrak{e}'_2, \ldots, \mathfrak{e}'_n)$, and write $a_i$ for the co-ordinates of $O'$ in $(O; \mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_n)$. The $x'_i$ then are the components of the vector $\overrightarrow{O'P}$, whereas $x_i - a_i$ are the components of $\overrightarrow{OP} - \overrightarrow{OO'}$. Since $\overrightarrow{O'P} = \overrightarrow{OP} - \overrightarrow{OO'}$, we can once again substitute the equations (3.4) to find that

$$x_i = \sum_{k=1}^{n} a_{ik}x'_k + a_i \tag{3.6}$$

He concludes that this type of notation allows for analytical treatment of affine geometry, where geometrical relations between points and vectors are presented as relations between their co-ordinates and components, of such a kind that they are not destroyed by the co-ordinate transformations described by equations (3.5) and (3.6).

## 3.4 Transformations and matrices

Weyl uses the change-of-basis formulas as a starting point for discussing affine transformations, stating that equations (3.5) and (3.6) can also serve as a representation of affine transformation within one established co-ordinate system. He then proceeds to define a linear or affine transformation:

**Definition: Affine transformations [21, pp. 21–22]**

A transformation, i.e. a rule which assigns a vector $\mathfrak{x}'$ to every vector $\mathfrak{x}$ and a point $P'$ to every point $P$, is called linear or affine if the fundamental relations (3.1) are not disturbed by the transformation: so that if the relations (3.1) hold for the original points and vectors they also hold for the transformed points and vectors:[a]

$$\begin{aligned} \mathfrak{a}' + \mathfrak{b}' &= \mathfrak{c}' \\ \mathfrak{b}' &= \lambda \mathfrak{a}' \\ \overrightarrow{A'B'} &= \mathfrak{a}' \end{aligned}$$

and if in addition no vector differing from $\mathfrak{o}$ transforms into the vector $\mathfrak{o}$. Expressed in other words this means that two points are transformed into one and the same point only if they are themselves identical.

[a] The English translation by Brose states $\overrightarrow{A'B'} = \mathfrak{a}' - \mathfrak{b}'$ as the final relation. I have used the seventh German edition for the version cited here.

Two figures that can be transformed into one another by such a transformation are said to be affine, and considered identical in affine geometry. Affine properties then are precisely those that are preserved by these transformations; linearly independent vectors remain independent, parallels remain parallel.

This entire approach is very similar to our modern one from definition 2.2.1. Assigning points to points corresponds to the map $f$ on the affine space $\mathbb{A}$, and assigning vectors to vectors corresponds to the associated map $\tilde{f}$ on the underlying vector space $\mathbb{V}$. Preserving the first two affine relations simply means that $\tilde{f}$ must be a linear map, and the last property gives us precisely the relation between $f$ and $\tilde{f}$. That is, if we translate that property to $\oplus$-notation like we did before, it now states that if $A \oplus \mathfrak{a} = B$, then $A' \oplus \mathfrak{a}' = B'$. This is identical to the relation $f(P \oplus \vec{v}) = f(P) \oplus \tilde{f}(\vec{v})$ that we require for modern affine maps. Thus, it is the preserving of affine properties that makes this transformation an affine map.

Weyl claims without proof that his requirement of no vector other than $\mathfrak{o}$ transforming to $\mathfrak{o}$, is equivalent to the transformation being injective on points. This is indeed correct, for suppose $A' = B'$ for any two points $A, B$. Then $\overrightarrow{A'B'} = \mathfrak{o}$, but this means that in fact $\overrightarrow{AB} = \mathfrak{o}$ hence $A = B$. So the transformation is indeed injective on points. Furthermore the transformation is also surjective: Any basis $(O; \mathfrak{e}_1, \ldots, \mathfrak{e}_n)$ is transformed into another basis $(O'; \mathfrak{e}'_1, \ldots, \mathfrak{e}'_n)$ since none of the $\mathfrak{e}_i$ get mapped to $\mathfrak{o}$ and the linear independence relation is preserved through preserving the affine relations. Using this new basis any point $P$ can be written as $P = O' + \sum_i p_i e'_i$, and is therefore the result after transforming the point $O + \sum_i p_i e_i$. Hence, the transformation is bijective on points, precisely the last remaining condition on our affine map $f$ to be an affine transformation.

Some slight differences still remain in notation: Weyl considers his transformation to act on both the vectors and the points, rather than there being two related maps on the two different sets. This reflects how Weyl seems to treat the point set and the vector field as two sides of the same coin. Additionally, this transformation is called both affine and linear. This again makes sense from the point of view where it acts both on the vectors and on points. The transformation on vectors indeed fits the modern definition of a linear map, it is only the transformation on points that is affected by a translation and thus still affine but no longer linear.

In this introduction on transformations, each component or co-ordinate is determined by its own sum. In modern notation, these can be condensed into one $n \times n$ matrix with the $a_{ki}$ as its entries, and in the case of points, also the addition of the vector with the $a_i$ as its entries. For the case of linear vector transformations, Weyl does indeed introduce matrices later on:

Definition: Matrix [21, pp. 39–40]

A **linear vector transformation** makes any displacement $\mathfrak{x}$ correspond linearly to another displacement, $\mathfrak{x}'$, i.e. so that the sum $\mathfrak{x}'+\mathfrak{y}'$ corresponds to the sum $\mathfrak{x}+\mathfrak{y}$ and the product $\lambda\mathfrak{x}'$ to the product $\lambda\mathfrak{x}$. In order to be able to refer conveniently to such linear vector transformations, we shall call them **matrices**. If the fundamental vectors $\mathfrak{e}_i$ of a co-ordinate system become

$$\mathfrak{e}'_i = \sum_k a_i^k \mathfrak{e}_k$$

as a result of the transformation it will in general convert the arbitrary displacement

$$\mathfrak{x} = \sum_i \xi^i \mathfrak{e}_i \text{ into } \mathfrak{x}' = \sum_i \xi^i \mathfrak{e}'_i = \sum_{ik} a_i^k \xi^i \mathfrak{e}_k.$$

We may therefore, characterise the matrix in the particular co-ordinate system chosen by the bilinear form

$$\sum_{ik} a_i^k \xi^i \eta_k.$$

I will not go into the concept of "bilinear form" in this definition, but a bit more can be said about the $a_i^k$ that appear throughout. These will eventually come to define the matrix. First they are referred to simply as the components of the matrix (e.g. p. 49), and eventually they are even deemed equal to the matrix, when for instance on page 139 he mentions "the rotation matrix $(a_k^i)$".

The use of matrices explains the mention of determinants in section 3.3 from a modern perspective, and indeed throughout his work we occasionally see the now familiar array notation of a matrix in order to compute a determinant. This is however the only context he uses this array notation, and in most other cases he prefers to refer to either summations or just the components instead. Occasionally (e.g. p. 50), Weyl will refer to a matrix using a capital letter, and use it as a function. This allows him to write $A(\mathfrak{x}) = \mathfrak{x}'$ to say that $A$ transforms $\mathfrak{x}$ into $\mathfrak{x}'$. From this he establishes matrix multiplication, with $BA(\mathfrak{x})$ being the vector one obtains by applying first $A$ and then $B$ to $\mathfrak{x}$, although the components of $BA$ are once again given by a sum. All in all, it does seem like Weyl did have many properties of matrices available to him, although his understanding of them was at times different from how we would approach them now. His introduction using summations rather than matrices, then, seems to have been a conscious choice rather than a limitation in the tools he had available.

# Chapter 4

# Hermann G. Grassmann's 'Ausdehnungslehre'

Although Weyl was one of the first to describe affine geometry in its modern form [22], that is most certainly not where its history starts. For the longest time, however, the study of affine geometry was limited to these affine transformations, often as just a special case of projective transformations. It is all the more surprising, then, that some historians claim that affine geometry was already developed back in the 1840s, by a man called Hermann Günther Grassmann [13, 19]. In the year 1844 Grassmann publishes the first edition of his Ausdehnungslehre (A1) [8]. In this work he establishes 'Extension Theory', an entirely new discipline of mathematics with which he aims to 'complete' a part he deemed missing from mathematics. His work went largely unnoticed, for which he and his admirers give different reasons. Some arguments include that his work was too philosophical for his fellow mathematicians, or that he was simply too far ahead of his time [19]. Indeed, those contemporaries that did receive his work, commented that it was too abstract, lacked intuitive clarity, or that they disliked the philosophical abstractions [17, p. 38]. Grassmann himself lamented that he was never able to obtain a position at a university, which greatly limited his opportunities to share his work [8, p. 19].

In response to the criticisms, Grassmann publishes a second version in 1862 (A2) [10], in which he takes a completely different approach. Since it seemed to be the philosophical approach that was keeping contemporaries from reading A1, in A2 he opted for an Euclidean approach. He also skipped many of the applications in physics that he had included in A1, as well as the sections that he deemed to difficult or analytical [17, p. 70]. Where A1 might have been overly philosophical, A2 jumped to the other extreme. With little to no justification for the concepts introduced in his work, his contemporaries failed to see the potential applications of the theory he developed. It was only in the late 19th and early 20th century that his work became noticed, and he eventually obtained somewhat of a cult following [4].

In this chapter I aim to dissect some of Grassmann's work, particularly that contained in A1, and compare it to our modern approach. In comparing these I hope to show whether Grassmann's posthumous credit for affine geometry is accurate, or whether people may have been too generous in that regard. Any page references in this chapter will refer to A1, unless stated otherwise.

## 4.1 Philosophy behind the Ausdehnungslehre

Grassmann's Ausdehnungslehre (A1) [8] starts off with a philosophical introduction, in which he describes his view of what pure mathematics is and how it should be organized. In his proposed structure he points out that there is still an area that has not been covered, and with his book he aims to develop the groundwork for this missing area of mathematics, completing the picture.

Grassmann describes pure mathematics as the "theory of forms" ("*Formenlehre*"), the study of objects of thought (pp. 23–25). It sets itself apart from other sciences in that its objects of study ('forms') do not exist independently from thought, which groups it together only with logic. It is then separated from logic as the study of the 'particulars', where logic studies the general laws of thought. He proceeds to observe that by this definition, geometry is not a mathematical discipline, as its objects of study (space, lengths, areas) exist independently from our thought. With his extension theory his intention is to generalize geometry into a discipline of pure mathematics, such that geometry will become just one of the applications of this science.

From here he continues to describe two axes along which disciplines of the theory of forms can be categorized, determined by how their object of study is generated from an initial element, and how the generated objects relate to each other. The first axis is that of discrete versus continuous generation (p. 25). On the one hand we have discrete generation, where the form is generated by distinct repeated acts of applying an operation and placing the result. An example of this would be the natural numbers. For continuous generation on the other hand the operation and placing 'blend together' to create a continuous stream of becoming, as we might see in functions, or in Grassmann's terms; magnitudes.

The second axis details how forms relate to others in its particular area of mathematics. Here he distinguishes the study of equals versus the study of differents, calling the former the algebraic form and the latter the combinatorial form (p. 26). In discrete mathematics this covers number theory on the one hand and combinatorics on the other. In continuous mathematics, Grassmann argues that so far only the algebraic form is studied in the form of analysis, the theory of functions. This leaves the continuous combinatoric form to be developed, and this is what Grassmann aims to do in his work A1, his study of 'extensive magnitudes' (p. 27). This partition of mathematics, and its associated areas of mathematics, is summarized in table 4.1.

Grassmann's philosophy of mathematics was heavily influenced by his father, Justus Grassmann [17, p. 236]. Justus had been the one who originally divided the areas of mathematics along these same two axes, although he had had no answer for what might serve as the area of continuous generation of differents. His focus instead was on the area of combinatorics, inspired by Leibniz. He considered combinations to be those objects one obtains by taking the conjunction of different elements [17, pp. 104–106], that is, in some way combining them. He eventually published a combinatorial approach to geometry [11], allowing him to study geometry in a way that was free of metrics. It is in this context that Hermann Grassmann eventually proposes his extension theory as 'continuous combinatorics', the study of conjunctions of continuously generated differents.

| | **Discrete generation** | **Continuous generation** |
|---|---|---|
| **Equal** | Theory of numbers | Theory of functions |
| **Different** | Combinatorics | *Extension theory* |

Table 4.1: Grassmann's partition of mathematics [1, 17]

Later in his career Grassmann seems to abandon this partition of mathematics. Particularly, his Textbook of Arithmetic [9] approaches arithmetic from the perspective of extensive magnitudes, that is, that of extension theory [17, p. 202], a branch that should be opposite to arithmetic in his original partition. A further analysis of Grassmann's changing philosophy and his inspirations can be found in his biography by Petsche [17] as well as several analyses by Cantù [1, 2].

### 4.1.1 Theory of Forms

In Grassmann's view, all branches of mathematics are governed by the same foundational 'truths', which he calls the theory of forms. This encompasses several statements on the concepts of equality, difference, and several operations. As these are considered a given for all his work on extension theory, I will describe some of these concepts here for context of what follows, as he describes them in the first 12 sections of his work (pp. 33–45).

To start off, he defines equality as

**Definition: equality [8, p. 33]**

*Those are equal of which one can always assert the same, or more generally what in any judgment can be substituted one for the other.*
It plainly follows from this that, if two forms are each equal to a third, they are also equal to each other; and that those generated in the same way out of equals are again equal.

With equality established, he discusses some properties of conjunctions, that is, of operations. Any conjunction combines two *factors* to form a *product*, and Grassmann does not limit these terms to multiplication. An important class of conjunctions are the *elementary* ones:

**Definition: elementary conjunction [8, p. 35]**

*If a conjunction is of the type that one can arbitrarily place parentheses around any three of its factors and change the order of any two of its factors without changing the product, then it is also true that the placement of parentheses and the order of factors is unimportant for the product of any number of factors.*
For brevity we call a conjunction satisfying the given conditions *elementary.*

In modern terms, an elementary conjunction is thus an operation that is both associative and commutative.

From here, Grassmann proceeds to separate conjunctions into two classes. On the one hand there are the *synthetic* conjunctions, generating a product from two factors. On the other hand there are their inverses, the *analytic* conjunctions, which aim to recover one of the factors of a synthetic conjunction given the other factor and the resulting product. Using these, he defines addition and subtraction:

Definition: addition/subtraction [8, p. 37]

*If the synthetic conjunction is elementary and the corresponding analytic conjunction is unique, then one can insert or omit parentheses after a synthetic symbol. In this case (provided that that uniqueness is generally valid), we call the synthetic conjunction* addition *and the corresponding analytic function* subtraction.

Hence, addition is defined as any operation that is both associative and commutative, and for which the inverse (subtraction) always yields a unique result. In this case, the brackets after an addition symbol may be omitted: $a + (b - c) = a + b - c$.

Following this, multiplication is defined as a conjunction of 'next higher order', with its only defining property being distributivity over addition (p. 40). Throughout this section, Grassmann derives some elementary properties of addition and multiplication, like the existence of a unique *indifferent form* ($a - a$, independent of $a$) and *negative form* ($-a$) (p. 38).

From a modern perspective, Grassmann's Theory of Forms almost takes the shape of a ring structure. Addition satisfies precisely the properties of addition in a ring, being associative, commutative, having a unique identity and unique inverse. His concept of multiplication is slightly broader, as associativity is not assumed. The level of generality in these definitions makes sense for his intention to apply this to any area of mathematics. In particular, it allows him to study various types of multiplication in his extension theory.

## 4.2 The 1844 lineale Ausdehnungslehre (A1)

After establishing the logical structure he will be working in, Grassmann details first the objects of study in extension theory and then studies their properties. He divides these objects into two types: The first part of his work discusses extensive magnitudes, which includes objects we would now call vectors. The second part studies elementary magnitudes, which capture the concept of weighted points. In the following discussion I will focus mainly on the first part. Towards the end I will discuss how some properties of extensive magnitudes translate to elementary magnitudes, as these will be particularly interesting when it comes to transformations.

Grassmann takes a very philosophical approach in A1, where he often describes the reasoning behind certain properties and slowly builds up to any specific definition or result. These results are then typically stated at the end of such a paragraph. In this discussion of his work I will take the opposite approach, starting by quoting, where possible, this final result, and then analyzing what he meant by it and how it could be interpreted with a modern lens.

### 4.2.1 The foundation: points and displacements

In the first chapter of part one, Grassmann develops the concepts that we could now consider as vectors and vector spaces. In section 13 and 14 (pp. 45–48) he describes how the most basic extensive magnitudes are generated by *elements* and *evolutions*. Elements serve as a generalisation of geometrical points or positions, and evolutions as a generalisation of motion, as it represents the transition of an element from one 'state' to another (p. 46). It is through this evolution that the 'differents', essential to extension theory, are generated. Specifically, he introduces the *extensive structure of first order*:

Definition: extensive structure of first order [8, p. 47]

*By an extensive structure of first order we mean the collection of elements into which a generating element is transformed by a continuous evolution.*
In particular we call the generating element in its first state the initial element, and in its last the final element.

Of particular importance are the infinitesimally small evolutions, which he calls *fundamental evolutions* and which essentially capture only a direction. With these he narrows down the concept of extensive structure:

Definition: elementary extensive structure [8, p. 47]

The *elementary extensive structure* is that which results from the continuous action of the same fundamental evolution.

The definition of extensive structure is a bit too general to translate into geometry, but the elementary extensive structure can represent a line segment. It is the collection of all the positions we travel through when we start at an arbitrary position and move a point in the same direction continuously until some endpoint. Continuing this fundamental evolution indefinitely, and including its opposite evolution, yields a *system (or domain) of first order*. This serves as an abstraction of the line as a collection of points, that is, the one-dimensional space.

When we consider only the means of generation of an elementary extensive structure, rather than its collection of elements, we obtain what Grassmann calls an *extensive magnitude* of first order, or a *displacement*. These correspond geometrically to vectors, as these have only length and direction but no position. As that which is generated equally is considered equal, indeed displacements generated by the same evolution from another starting element are equal, just as vectors of equal direction and length are equal regardless of position.

In a roundabout way we can distill from these definitions the relation between points and vectors. By definition, an extensive structure of first order, and therefore also an elementary extensive structure, has an initial and final element, say $\alpha$ and $\beta$ respectively. As a displacement captures the generation of this structure, it captures in particular how $\alpha$ is transformed into $\beta$. With that in mind, Grassmann begins to write $[\alpha\beta]$ for such a displacement starting in §15.

In this section he establishes addition of displacements:

Definition: addition of similar displacements [8, p. 49]

*If two similar displacements are contiguous, that is, so conjoined that the final element of the first is the initial element of the second, then the displacement from the initial element of the first to the final element of the second is the sum of the two.*

Using his newfound notation for displacements he shows that it indeed satisfies the necessary properties to be an addition, and derives a meaning for a 0 displacement as $[\alpha\alpha]$ for any element $\alpha$, and that of a negative displacement as $-[\alpha\beta] = [\beta\alpha]$.

Technically, Grassmann's definition here does not apply to all displacements. So far he has only established systems of first order, that is, one-dimensional spaces, and he only defines this addition for ‘similar’ displacements, those “generated in the same sense”, or in the opposite sense (p. 39). Geometrically, this corresponds to vectors with parallel directions. To prove that dissimilar, non-parallel displacements can be added in the same way, as he does in §§17–18, he must first define systems of higher order.

### 4.2.2 Bases and dimensions

In §16, Grassmann describes how one can generate systems of any order:

Definition: systems of higher order [8, p. 50]

I consider next two dissimilar fundamental evolutions, and let the first fundamental evolution (or its opposite) arbitrarily transplant an element, and then let the element thus generated likewise be arbitrarily transplanted by the second method of evolution. In this way I can therefore generate an infinite set of new elements from a single element, and the collection of elements so generated I call a system of *second order*. I then assume a third fundamental evolution, which does not lead from that same initial element again to one of the elements of this system of second order, and which I therefore designate as *independent* of those first two, and let this third evolution (or its opposite) arbitrarily transplant an arbitrary element of that system of third order. Since this method of generation has no limitation on its concept, I can in this way proceed to systems of arbitrarily high order.

In his goal of generalizing from geometry, Grassmann avoids the use of any given space, and instead generates higher orders one at a time by conceiving of evolutions independent to the previous ones. The first order, a line, consisted of all the positions a point could arrive at following one direction. By this definition, moving that line in an independent direction we get a collection of infinitely many parallel lines, that is, a plane, the system of second order. Moving that plane in yet another independent direction, the collection of these parallel planes forms all of space, the system of third order. As he doesn’t presume our physical space as given, he is not limited to three spatial directions, and can thus continue this process, collecting parallel spaces to form a fourth order system, and so on.

The fundamental evolutions in this definition can be viewed as some foundation for bases, and from there the order of his systems can be related to dimensions. In §20 he proves that systems of $m$-th order are indeed generated by any $m$ independent displacements:

Independence of evolutions of higher order system [8, p. 60]

*Every displacement of a system of $m$-th order can be represented as a sum of $m$ displacements belonging to $m$ given independent methods of evolution of the system, the sum being unique for each such set.*

And that it’s independent of its choice of initial element:

Independence of starting element of higher order system [8, p. 60]

*Every system of m-th order can be regarded as generated by those same m independent methods of evolution from any arbitrary element; that is, all other elements can be generated from a single such element by those methods of evolution.*

With (arbitrary) bases and (arbitrary) origins of spaces established, it seems like it should be a small step to arrive at a coordinate system. However, the systems he has built have no underlying field, and therefore there are no scalars to function as coordinates. It takes him the better part of four chapters before he introduces ‘numerical magnitudes’ in §68, derived from the inverse of a conjunction he calls outer multiplication.

### 4.2.3 Outer multiplication, numerical magnitudes and coordinates

Grassmann’s goal in A1 is to establish this branch of mathematics fully independently of the other known branches. This means that some familiar concepts such as numbers are put on the sideline in favor of studying the conjunctions between the extensive magnitudes themselves. Even when some numerical concept eventually shows up, it is embedded within the context of extension theory, when he introduces numerical magnitudes. These numerical magnitudes are defined as a ratio or quotient of two similar displacements, and are treated as such throughout. Intuitively, these numerical magnitudes represent how many times longer one vector is compared to another. To understand how Grassmann arrives there, we first need to understand his outer multiplication.

In chapter 2, Grassmann introduces this outer multiplication. He starts out with a geometric concept, generalizes this to his extensive magnitudes, and finally arrives at an abstract property of the multiplication so developed which then serves as its abstract definition. In this section I will follow the same approach.

From a geometric point of view, outer multiplying two vectors yields the area of the parallelogram they enclose, and similarly the product of three vectors gives the volume of the parallelepiped they define. This idea is somewhat reminiscent of our modern cross product. However in Grassmann’s case the product is not another vector, but instead the ‘actual’ two-dimensional area or three-dimensional volume. For example, in figure 4.1, the product of $a$ and $b$ is the parallelogram $\alpha\beta\beta'\alpha'$, and that of $a$ and $-b$ is the parallelogram $\alpha\beta\beta''\alpha''$. These products are also, in a sense, signed. Since $b$ and $-b$ have opposite direction, the products $a.b$ and $a.-b$ will have opposite sign; if the first is positive, the second will be negative.

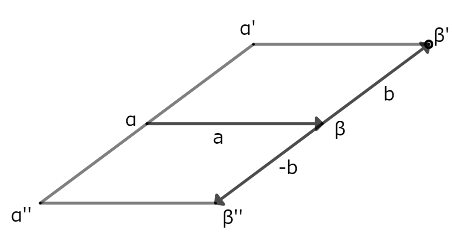



Figure 4.1: Outer multiplication

If one were to imagine a third vector $c$ sticking out of the paper, then $a, b, c$ together define a parallelepiped in the same way $a, b$ define a parallelogram, and this parallelepiped then represents the product $a.b.c$.

If we let go of the restrictions of space, we can continue this process indefinitely with arbitrarily many independent displacements, and indeed this is how Grassmann generalizes this concept to extensive magnitudes. Since the outer product of two displacements is not another displacement but rather a generalization of parallelogram, this product creates a new category of objects which he calls an extension of second order. From there he tentatively generalizes to extensions of $n$-th order:

**Definition: extension of $n$-th order [8, p. 77]**

We now extend this definition to arbitrarily many factors, and tentatively mean by $a{\cap}b{\cap}c\ldots$ where $a, b, c, \ldots$ are arbitrarily many, say $n$, displacements, that extension resulting when each element of $a$ generates the displacement $b$, all elements of the resultant generate the displacement $c$, and so on, and require this extension to be similar to all other parts of the same system of $n$-th order[a]. We call the extension so generated an extension of $n$-th order.

[a]The 1995 translation erroneously said second order here, I have corrected this using the German original.

This definition mirrors that of his systems of higher order, but it generates a section of a system rather than the system in its totality. Geometrically, to generate some extension $a \cap b \cap c \ldots$, one starts with the displacement $a$. This displacement marks a line segment, a section of a line. Moving all points on the line segment $a$ along the displacement $b$ yields a collection of parallel line segments, which together form a parallelogram, a section of a plane. Moving all points on the parallelogram along a third displacement $c$ yields a collection of 'parallel' parallelograms, which together form a parallelepiped, a section of space. It is this process of generation which eventually leads to the $n$th order extensions, as a part of an $n$th order system.

This concept relies on the concept of similarity and what it means for higher order extensions to be similar, that is, 'generated in the same way'. He therefore also includes that any $n$th order magnitude belonging to the same $n$th order system, is similar to one another. This does not immediately seem intuitive; the parallelograms in figure 4.2a seem to be entirely unrelated, yet Grassmann considers them 'similar' because they lie in the same plane. However, parallels only

(a) A 2D view

(b) A 3D view

Figure 4.2: similar parallelograms

become visible when we go one dimension up. And indeed, in the 3D view in figure 4.2b, their shared property becomes visible: although they are completely different parallelograms, they lie on parallel planes, and are thus generated by the same two directions. This is the property that Grassmann means by 'similarity'.

The definition of higher order extensions that Grassmann gave is only tentative as he has yet to show that this operation is indeed a multiplication, which is also why he uses the $\cap$ notation. In the following section he shows that it indeed satisfies the properties of multiplication that he established for the theory of forms, and writes the product of $a, b$ as $a.b$ instead.

In §§33–34, finally, he arrives at the abstract property that will define outer multiplication:

**Definition: outer multiplication [8, p. 81]**

[...] for this particular type of multiplication we have obtained the law that "if a factor includes a summand that is similar to one of the adjacent factors, then that summand can be dropped," which already incorporates the result that if two adjacent factors are similar, the product is zero.
This law, in combination with the general multiplicative relation to addition for forms, fixes all further laws regarding the particular type of multiplication we consider here, and thus can be taken as its fundamental law. We call this type of multiplication *outer*, and take as its particular symbol the point (period), retaining simple juxtaposition as the general sign for multiplication.

Geometrically, this law is the generalisation of the fact that in a situation like figure 4.3, $a.b$, that is, the parallelogram $\alpha\beta\beta'\alpha'$, has an equal area to $(a + b_1).b$, the parallelogram $\alpha\gamma\gamma'\alpha'$. The term $b_1$, being similar to $b$, can thus be 'dropped' from the multiplication $(a + b_1).b$ and we can say that $(a + b_1).b = a.b$. In particular, this means that $b.b_1 = b.0 = 0$, and in fact the product is only non-zero when all factors are independent. This is the reason he chose the name *outer* product, as it only has a meaningful result when choosing factors outside of the previous ones.

Grassmann spends all of chapter 3 developing the properties of the higher order extensions, and so it takes until §60 in chapter 4 for Grassmann to define the inverse of outer multiplication, being outer division. A special case of this occurs when dividing two similar displacements, say $a$ and $a_1$. The quotient $\frac{a_1}{a}$ must be such that, when outer multiplied by $a$, it yields $a_1$. Outer multiplication

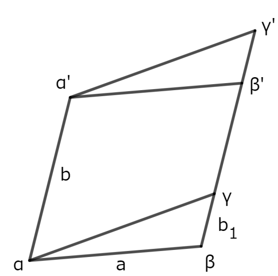



Figure 4.3: Outer multiplication

however has the property that the order of the product is the sum of the orders of its factors. In order to extend this property to fractions, and thus to hold for the equality $\frac{a_1}{a}.a = a_1$, the order of $\frac{a_1}{a}$ must be 0 [8, see footnote p. 125]. It is these 0-order magnitudes, representing the proportion between two similar displacements, that Grassmann defines as *numerical magnitudes* in §68. Intuitively, these represent scalars $\lambda$ such that $\lambda a = a_1$. Equality of numerical magnitudes is defined through their relation with outer multiplication: two numerical magnitudes $\frac{a_1}{a}, \frac{b_1}{b}$ with $a, b$ independent, are considered equal when $a_1.b = a.b_1$ (p. 122). Since $\frac{b_1}{b}b = b_1$ and equals can be substituted for each other, it then also holds that $\frac{a_1}{a}b = b_1$, which fully defines the scalar multiplication. Furthermore, the numerical magnitude $\frac{a}{a}$ for arbitrary $a$ can serve as 1, since for any $b$, $\frac{a}{a}.b$ equals that magnitude $b_1$ such that $a.b = a.b_1$, thus in fact $b_1 = b$. Grassmann proceeds to define addition and multiplications of these objects, and eventually proves that these numerical magnitudes satisfy all laws of arithmetic. I will omit these definitions and proofs since they would require further discussion on the higher order extensive magnitudes.

Now that we finally have an extensional equivalent of numbers, and thus scalars, a coordinate system can be established, and Grassmann does so in §§87–88. Grassmann defines these in such a way that they apply to extensions of any order, making for a rather convoluted definition. Hence after citing his definitions I will narrow this down to displacements.

**Definition: Reference system [8, p. 147]**

I call the $n$ displacements $a, b, \ldots$ that define a system of $n$-th order (and thus are all mutually independent) the *reference measures* of first order or the *fundamental measures* of the system, inasmuch as each displacement of the system is expressed through them; the assembly of them a *reference system*, the products of $m$ fundamental measures (retaining their original order) *reference measures of $m$-th order*, the reference measure of $n$-th order the *principal measure*. Finally, we call the system of reference measures of $m$-th order a *reference domain of $m$-th order*, and in particular the system of fundamental measures, *reference axes* (coordinate axes).

**Definition: Reference terms, indicators [8, p. 148]**

[...] each extension of $m$-th order belonging to a system of $n$-th order can be interpreted as a sum of terms similar to the reference measures of $m$-th order belonging to that system. We now call these terms *reference terms* of that magnitude, so that each magnitude appears as the sum of its reference term; the numerical magnitudes resulting if the reference terms of a magnitude are divided by the corresponding (similar) reference measures are the *indicators* of the magnitude, so that each magnitude therefore appears as a multiple sum of the reference measures of the same order.

If we're only concerned with displacements, we will only have to concern ourselves with the fundamental measures. So let $e_1, \ldots, e_n$ be $n$ independent displacements that define a system of $n$-th order, then these take the role of reference measures. The collection of these, a basis of the space, Grassmann calls a reference system. We know that any displacement $x$ in the system can be written as a sum of multiples of the reference measures: $x = \alpha_1 e_1 + \cdots + \alpha_n e_n$. The terms $\alpha_i e_i$ are what Grassmann calls the reference terms. Finally, taking the quotients $\frac{\alpha_i e_i}{e_i}$ results in $n$ numerical magnitudes, which Grassmann calls the indicators of the displacement. It is these indicators of the displacement which correspond to the coordinates of our vector, the $\alpha_i$ in the earlier expression. In this way, Grassmann's reference systems almost define a coordinate system. This system however lacks an origin, and as such it cannot give the coordinates of any points.

### 4.2.4 Elementary magnitudes

Where part one of A1 is concerned with extensive magnitudes, generalizing vectors, the second part discusses the concept of elementary magnitudes, which generalize weighted points. In this part, Grassmann demonstrates how many of the concepts he developed for extensive magnitudes translate to elementary magnitudes. In this section I will describe some of the concepts and their consequences from a more modern lens, without quoting Grassmann directly as I have done in the previous sections.

As with extensive magnitudes, the concept of elementary magnitude is very broad and not immediately intuitive. In §98 however, Grassmann demonstrates that elementary magnitudes of first order can simply be considered as weighted points, or as he calls them, multiple elements, the weight being a numerical magnitude. The element and weight together form the elementary magnitude, so every elementary magnitude $A$ is of the form $x\sigma$ with a weight $x$ and an element $\sigma$. An elementary magnitude with weight 1 is then called a simple element. The one exception to this rule is the magnitude with weight 0, which will be represented by displacements, for reasons that will become clear when we consider addition.

Grassmann's argument in §98 also shows how addition works on these objects. Suppose one wants to compute the sum $a\alpha + b\beta + \dots$, where $\alpha, \beta, \dots$ are elements, or points, and $a, b, \dots$ their corresponding weights. Then the outcome is defined as that element $\sigma$ with weight $x$ such that the weight $x$ is equal to the sum of the $a, b, \dots$, and the element $\sigma$ satisfies the following relation for any arbitrary point $\rho$:

$$a[\rho\alpha] + b[\rho\beta] + ... = (a + b + \dots)[\rho\sigma]$$

or equivalently

$$[\rho\sigma] = \frac{a[\rho\alpha] + b[\rho\beta] + ...}{a + b + \dots}.$$

Hence, the sum essentially takes the shape of a weighted average.

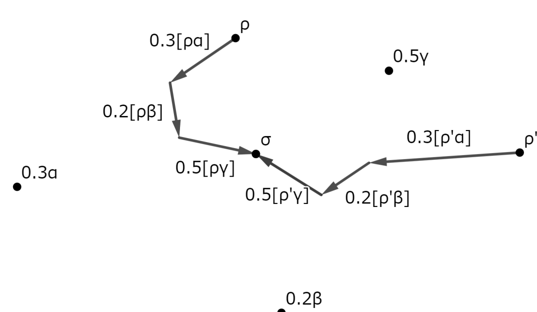



Figure 4.4: Addition of elementary magnitudes

As an example, consider figure 4.4, which represents the addition $0.2\alpha + 0.3\beta + 0.5\gamma$. One takes an arbitrary point $\rho$, and adds to $\rho$ the displacements $0.2[\rho\alpha], 0.3[\rho\beta], 0.5[\rho\gamma]$ to arrive at $\sigma$, which will be in the same location no matter what point $\rho$ was chosen.

The one exception to this rule of addition is the case where $a + b + \cdots = 0$, in which case an element $\sigma$ satisfying the above equations will no longer exist. In fact, in this case it is not the element $\sigma$ that remains equal regardless of the choice of $\rho$, but rather the displacement $[\rho\sigma]$. Consider for example $\beta - \alpha$. We would be looking for a $\sigma$ such that, for any $\rho$,

$$\begin{aligned}
[\rho\sigma] &= [\rho\beta] - [\rho\alpha] \\
&= [\rho\beta] + [\alpha\rho] \\
&= [\alpha\rho] + [\rho\beta] \\
&= [\alpha\beta]
\end{aligned}$$

Clearly no one such $\sigma$ exists, and Grassmann opts to instead define $\beta - \alpha = [\alpha\beta]$, showing that this is consistent with the further definitions. In fact, it is consistent with the relation $\alpha + [\alpha\beta] = \beta$, where we can now consider both points and displacements as elementary magnitudes. Finally, the weights of elementary magnitudes allow us to easily define scalar multiplication: multiplying an elementary magnitude $a\alpha$ by a numerical magnitude $m$ simply gives the same element $(\alpha)$ with a weight of $m \cdot a$ instead. This is still consistent with displacements as well, as indeed $2[\alpha\beta] = 2(\beta - \alpha) = 2\beta - 2\alpha$ will have twice the length of $[\alpha\beta]$.

With elementary magnitudes established, Grassmann proceeds to show how many of the properties of extensive magnitudes also extend to elementary magnitudes. In §107 he defines elementary magnitudes as independent when one cannot be written as a multiple sum of the others. Geometrically, this occurs when three points do not all lie on the same line, or four points do not all lie on the same plane. Following that, $n$ independent elementary magnitudes form an *elementary system of $n$-th order*, as the collection of elements that are dependent on the $n$ independent elements. Since one needs four points to define a three-dimensional space, this means that space can be interpreted both as a third order extensional system and a fourth order elementary system. More generally, the displacements in any elementary system of $n$-th order belong to an extensional system of $n-1$-th order, for suppose $\alpha, \beta, \gamma, \ldots$ are the $n$ mutually independent elements of the elementary system. Then any displacement $\rho - \sigma$ in this system can be written as

$$\rho - \sigma = a\alpha + b\beta + c\gamma + \ldots$$

where $a + b + c + ... = 0$, hence $a = -b - c - \ldots$. Substituting that, we get that

$$\begin{aligned}
\rho - \sigma &= b(\beta - \alpha) + c(\gamma - \alpha) + ... \\
[\sigma\rho] &= b[\alpha\beta] + c[\alpha\gamma] + ...
\end{aligned}$$

and thus, indeed, a system of $n-1$ displacements with $\alpha$, in a sense, acting as the origin.

Grassmann also defines an outer product of elementary magnitudes in §108 using the same abstract property he used for extensional magnitudes (see section 4.2.3), and which therefore has all the same properties as were proven for the extensional version. This elementary product relates to the extensional product the same way the elementary systems relate to extensional ones: the product of three points, say $\alpha.\beta.\gamma$, is the parallelogram with these three points as three of its vertices, making for a third order elementary magnitude. It corresponds to the second order extensional magnitude $[\alpha\beta].[\alpha\gamma]$, as Grassmann establishes in §115. In general, any $n$-th order elementary magnitude will correspond in this way to an $n-1$-th order extensional magnitude.

In §§116-117 Grassmann extends the concept of reference systems, that is, coordinate systems, to elementary magnitudes, using the previously established addition and scalar multiplication. In §117 he mentions that, when limited to space, we find the type of coordinates that Möbius used in his Barycentrische Kalkül [14], the simplest version of which we obtain when we consider only simple elements, or points with weight 1. In this case, suppose $\alpha, \beta, \gamma, \delta$ are four independent reference magnitudes. Then any point $\rho$ in space can be written as a sum

$$\rho = a\alpha + b\beta + c\gamma + d\delta \qquad \text{with } a + b + c + d = 1,$$

similar to how in figure 4.4 on page 24, $\sigma$ was written as a sum of $\alpha, \beta, \gamma$. The $a, b, c, d$ then are $\rho$'s indices, and they add up to 1 to ensure that $\rho$ is once again a point with weight 1. As it follows that $a = 1 - b - c - d$, this can also be written as

$$\begin{aligned} \rho &= (1 - b - c - d)\alpha + b\beta + c\gamma + d\delta \\ &= \alpha + b(\beta - \alpha) + c(\gamma - \alpha) + d(\delta - \alpha) \\ &= \alpha + b[\alpha\beta] + c[\alpha\gamma] + d[\alpha\delta] \end{aligned}$$

and we have thus found our 'typical' system with an origin ($\alpha$) and three base vectors as a full coordinate system.

With the elementary magnitudes and their reference systems established, we can finally consider Grassmann's concept of the affine transformation.

### 4.2.5 Affinity

While most of the history of affine geometry has been primarily concerned with affine transformations, Grassmann only discusses affinity in a few sections towards the end of his work.

In the 1844 edition, the concept of affinity is mainly treated as an equivalence relation between two *assemblies*, or collections, of magnitudes. In chapter 4 of part 2, Grassmann observed that any numerical relation that holds between magnitudes, is preserved when one considers their 'shadows' in a system of lower order. The reverse, however, is not generally true. For example, when three points lie on the same line in space, their projections on any plane will be collinear as well. However, figure 4.5 shows the reverse is not true. The shadows $\alpha', \beta', \gamma'$ on the plane of the points $\alpha, \beta, \gamma$

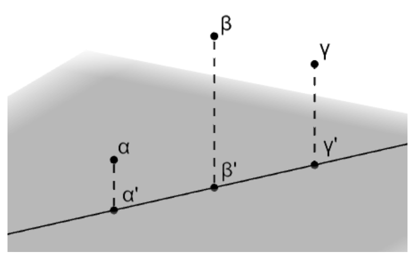



Figure 4.5: Shadows of vectors

might be collinear, while the original points are not. Observing this one-sided relation, Grassmann proceeds to consider those relations that do preserve numerical relations both ways, and calls those affine:

> Definition: affine assemblies [8, p. 248]
>
> This relation only appears in its complete generality, however, if the correspondence is reciprocal, that is if every numerical relation obtaining between magnitudes of one series, whichever it is, also prevails between the magnitudes of the other series; and two such assemblies of corresponding magnitudes standing in this reciprocal relation to each other we call *affine*.

Although Grassmann defines this relation between arbitrary collections of magnitudes, this can easily be extended to entire systems to obtain a concept of affine transformation. In §154 Grassmann describes that to construct an assembly affine to a given one, one must identify a set of $n$ independent magnitudes out of which all other magnitudes in the assembly can be obtained. The second assembly is then built by associating these magnitudes by another $n$ arbitrary magnitudes with the only requirement that these, too, are mutually independent. The other magnitudes are associated to magnitudes built from these $n$ magnitudes in the same way the original magnitude was in the original assembly. As an example, suppose one has a set $S$ of magnitudes, and each of these magnitudes can be written as a multiple sum of the three independent magnitudes $\alpha, \beta, \gamma$. An affine assembly is then constructed by finding three independent magnitudes $\alpha', \beta', \gamma'$, and associating $\alpha$ with $\alpha'$, $\beta$ with $\beta'$ and $\gamma$ with $\gamma'$. Any other magnitude $\rho = a\alpha + b\beta + c\gamma$ in the set $S$ will then be associated to $\rho' = a\alpha' + b\beta' + c\gamma'$.

This concept of $n$ independent magnitudes can easily be generalized to the reference system of an $n$-th order system. Two assemblies then are affine if they transform into one another by changing the reference system, and properties are affine precisely when they can be expressed as a numerical relation of the indicators. Indeed, in §161 Grassmann proves that his initial definition of affinity corresponds to that of having equal indicators in different reference systems.

In §159, Grassmann discusses affinity in the plane, in which case three independent point magnitudes $\alpha, \beta, \gamma$, correspond to three independent point magnitudes $\alpha', \beta', \gamma'$. This correspondence preserves the property of three points being collinear, and in fact Grassmann also uses the term collinearly related to denote the same relation. The relation is however more general, as the reference measures are not restricted to simple points. The reference measures in one assembly might be associated to magnitudes in the other with a different weight, or even to ones with weight 0, that is, displacements. In this case, points get associated to directions and directions get associated to points, in which case properties like parallel lines are no longer preserved. Grassmann acknowledges this in §164, where he states that "Our definition of affinity therefore coincides with the usual one so long as it is applied to the same magnitudes, that is to simple points (with equal weights)".

## 4.3 Grassmann's second edition (1862) (A2)

In 1862 Grassmann publishes 'Die Ausdehnungslehre. Vollstandig und in strenger Form bearbeitet' [10] (A2). In this work he once again publishes his extension theory, but this time with a far more mathematical approach. Along with the change in style, Grassmann has also changed some of the mathematical structure in A2. The strict divide that he made between extensive and elementary magnitudes in A1 is entirely gone, and he instead treats them as the same object wherever possible. He also lets go of his aim to build extension theory completely independently from existing areas of mathematics like analysis [17, p. 76], which removed many obstacles. Particularly, where in A1 numbers (or numerical magnitudes) were first introduced in chapter 4, in A2 he makes use of them immediately. Magnitudes are now built up from *units*:

**3. DEFINITION (Unit) [10, p. 3]**

I define as a *unit* any magnitude that can serve for the numerical derivation of a series of magnitudes, and in particular I call such a unit an *original unit* if it is not derivable from another unit. The unit of numbers, that is *one* I call the *absolute* unit, all others *relative*. Zero can never be a unit.

Several independent units form a *system of units*, and any magnitude derived from them is called an extensive magnitude:

**5. DEFINITION (Extensive magnitude) [10, p. 4]**

I define as an *extensive magnitude* any expression that is derived from a system of units (none of which need be the absolute unit) by numbers, and I call the numbers that belong to the units the *derivation numbers* of that magnitude; for example the polynomial

$$\alpha_1 e_1 + \alpha_2 e_2 + \dots$$

or

$$\sum \alpha e \quad \text{or} \quad \sum \alpha_r e_r,$$

where $\alpha_1, \alpha_2, \dots$ are real numbers and $e_1, e_2, \dots$ form a system of units, is an extensive magnitude, specifically the one derived from the units $e_1, e_2, \dots$ by the numbers $\alpha_1, \alpha_2, \dots$ belonging to them.

The collection of all extensive magnitudes derived from a system of units is called a *domain (of* $n$*th order)* in A2, rather than a *system* (of $n$th order) [10, p. 8].

The concept of weighted points is eventually introduced, but only in the context of applications to geometry [10, ch. 5]. This time he specifies that every point has a position and a coefficient (weight). The displacements, as points of weight 0, are given a position at infinity.

For multiplication he also takes a different approach. First of all, he chooses a different notation, as the product of two magnitudes $a, b$ is now written $[ab]$. Further, since all extensive magnitudes can be represented by their derivation from a system of units, this notation can be used to define a general product from its distributivity over addition [10, p. 19]:

$$\left[\sum_r \alpha_r e_r \cdot \sum_s \beta_s e_s\right] = \sum_{r,s} \alpha_r \beta_s [e_r e_s]$$

Different choices for what the $[e_r e_s]$ mean then lead to different types of products. Setting $[e_r e_s] = -[e_s e_r]$ for all $r, s$ (from which it naturally follows that $[e_r e_r] = 0$ for all $r$), yields what Grassmann calls the *combinatorial product*. He later defines this somewhat more generally:

**52. DEFINITION (Combinatorial product) [10, p. 29]**

If the factors of a product $P$ are derived from a system of units, and every pair of products of the units that result from the interchange of the last two factors yield zero upon summation, but every product that includes only different units as factors is different from zero, then I call that product $P$ *combinatorial*, and those factors of it its *elementary factors*; that is, if $b$ and $c$ are *units*, and $A$ is an arbitrary *series of units*, then the above definition is expressed by the formula

$$[Abc] + [Acb] = 0$$

He then proceeds to show that, in fact, the combinatorial product is such that interchanging any two elementary factors (that is, factors that are not themselves products), causes the sign to change. This turns out to be equivalent to the outer product from A1. Indeed, in A1, Grassmann proves that the outer product is such that interchanging two adjacent factors will change the sign [10, p. 82]. Conversely, in A2 he shows that the combinatorial product does not change if, to one of the factors, a multiple of one of the other factors is added [10, 35], meaning that for instance $[ab] = [a \cdot (b + qa)]$. This was precisely the fundamental property of outer multiplication in A1. Finally, he defines *units of $m$th order* as the products of $m$ 'original' units, and *magnitudes of $m$th order* as any magnitude derived from the $m$th order units, which completes the connection to the outer product of A1. The term outer product does appear in A2 as well, and is closely related: the outer product of two higher order magnitudes is the combinatorial product of their underlying factors [10, p. 45].

The entire section on affinity in A1 has been removed in A2. Anything related to geometric relationships has been condensed to a remark [10, p. 221]. This remark is in regards to a whole section on 'quotients', a concept Grassmann introduced in A2 which, through a modern lens, might be interpreted as transformation matrices. In the next chapter I will further discuss these quotients and their relation to matrices.

# Chapter 5

# Quotients as Grassmann's concept of transformation matrix

In modern algebra, matrices play a major role, as they can be used to represent any linear map. In Grassmann's time, the concept of matrices, or arrays, definitely wasn't unknown, but it didn't yet see much application beyond representing systems of linear equations. Its use for transformations had been very limited. Between the publications of A1 and A2 the theory of matrices and their use for transformations was greatly expanded by Arthur Cayley, however it is unclear whether Grassmann was aware of this development [5]. It is therefore not strange that, although Grassmann does mention the determinant of "$n$ series of $n$ numbers" in A2 [10, p. 33], no actual matrices appear in his work. He does however introduce a more general concept of a quotient which, as we shall see, behaves like a matrix in many ways. In the following chapter I will detail how Grassmann defines his quotient and develops some of its properties and applications, comparing to the modern matrix properties throughout. All page and section numbers in this chapter will thus refer to A2 [10].

Like the rest of this work, Grassmann attempts to define his quotients and their properties for the most general concept of (first order) magnitudes. For readability, I will only interpret these as displacements, that is, vectors. When we get to applications, I will extend this interpretation to points as well.

In 377, Grassmann defines his quotient:

**377. DEFINITION (Quotient) [10, p. 207]**

If $a_1, a_2, \ldots, a_n$ are magnitudes of first (or $(n-1)$th) order in a principal domain of $n$th order, and *stand in no numerical relation to one another*, I mean by the *fraction (quotient)*

$$Q = \frac{(b_1, b_2, \ldots, b_n)}{(a_1, a_2, \ldots, a_n)}$$

the expression which, multiplied by $a_1, a_2, \ldots, a_n$, yields the values $b_1, b_2, \ldots, b_n$ respectively, so that

$$\frac{(b_1, b_2, \ldots, b_n)}{(a_1, a_2, \ldots, a_n)} a_r = b_r.$$

I call $(a_1, a_2, \ldots, a_n)$ the *denominators* of the fraction, $(b_1, b_2, \ldots, b_n)$ their corresponding *numerators*, and set two fractions, or two expressions numerically derived from fractions, equal to one another if and only if they yield equals when they are multiplied by every magnitude of first {or $(n-1)$th} order. If in addition the numerators are magnitudes of first (or $(n-1)$th) order, *and stand in no numerical relation to one another*, I call the fraction *invertible*, and in this case, if

$$Q = \frac{(b_1, b_2, \ldots, b_n)}{(a_1, a_2, \ldots, a_n)},$$

I denote the inverted fraction by $\frac{1}{Q}$, that is, I set

$$\frac{1}{Q} = \frac{(a_1, a_2, \ldots, a_n)}{(b_1, b_2, \ldots, b_n)}.$$

So, what we have here is a set of $n$ independent vectors $a_1, a_2, \ldots, a_n$ in an $n$-dimensional space, a set of 'values' $b_1, b_2, \ldots, b_n$, and an object $Q$ that, when applied to any $a_r$ through multiplication, yields the corresponding $b_r$. While the term 'value' is rather vague and seems to imply numbers, Grassmann actually uses the term for just about any object and expression, including magnitudes. In the remark after 377 he clarifies that he will always consider the denominators magnitudes of first (or $(n-1)$th) order, unless stated otherwise. Hence for the sake of comparison I will also consider the $b_i$ to be displacements. The object $Q$ then acts as a linear transformation on a basis $a_1, \ldots, a_n$. How $Q$ acts on any other displacement follows by writing this displacement as a numerical relation of the $a_i$ and applying the distributive properties of multiplication that he established earlier. In 378 he uses this to prove that indeed defining a quotient on a system of units uniquely determines how it acts on all displacements, and thus that the $Q$ defined in 377 functions as a well-defined map on the whole space.

By the above definition, scalars behave as a special case of a quotient. Multiplying any vector by a scalar $\rho$ or a matrix $\rho I_n$ yields the same result, so from Grassmann's perspective $\rho$ and $\rho I_n$ are equal, as they both represent the quotient $\frac{(\rho a_1, \rho a_2, \ldots, \rho a_n)}{(a_1, a_2, \ldots, a_n)}$. He points out this particular fact in his remark after 383 (p. 210), and this will become relevant to understand his notation in section 5.2.

In the remark after 377 Grassmann states that the denominators could, in fact, have been magnitudes of arbitrary higher order, "but one would then exchange the indubitable advantage of greater simplicity for the doubtful advantage of barren generality." It might be precisely this generality where both the strength of his approach, and the weakness in its reception lies. For even now to compare this work to modern mathematics, I narrow his 'narrow' definition down further to only first order, and specifically only to the first order object of displacements. Grassmann's approach actually applies to more, including more abstract, objects, but in doing so it becomes harder to grasp intuitively, without many apparent applications.

In 379 (p. 208), Grassmann demonstrates how to add quotients with the same denominators and how to multiply them by scalars. Using the general properties of multiplication he shows that

$$\beta \frac{(b_1, b_2, \ldots)}{(a_1, a_2, \ldots)} + \gamma \frac{(c_1, c_2, \ldots)}{(a_1, a_2, \ldots)} + \cdots = \frac{((\beta b_1 + \gamma c_1 + \ldots), (\beta b_2 + \gamma c_2 + \ldots), \ldots)}{(a_1, a_2, \ldots)}.$$

This equation makes sense from the perspective of a fraction; to add fractions, the denominators must be equal, and the numerators are added (in this case; component-wise), and scalar multiplication gets applied to the numerator.

In 380 (p. 208) he proves that a quotient can be rewritten to have any set of independent displacements as its denominators, which allows us to write all of our quotients from the same original units $e_1, e_2, \ldots, e_n$. A quotient then tells us where all of the $e_i$ get mapped to, which is precisely what the columns of a modern transformation matrix do. In other words, a matrix contains only the information in the numerator of the quotient, as the denominator is automatically assumed to be 'the' basis relative to which everything is computed. With this translation the properties of 379 also boil down to the scalar multiplication and component-wise addition of matrices.

The result of 381 (p. 209) completes this translation, as here he introduces what are essentially the components of a matrix, and from it a rather unintuitive way of writing any quotient as a sum of these. To do so, he first introduces fractional units. These are the $n^2$ quotients $E_{r,s}$ that multiplied by $e_r$ give $e_s$ and multiplied by any other $e_i$ give 0. In matrix terms: $E_{r,s}$ is the matrix with a 1 in the $r$'th column, $s$'th row, and 0s everywhere else. As multiplying a fraction $Q$ with some $e_i$ results in some displacement of the form $\sum_{b=1}^{n} \alpha_{i,b} e_b$, it can be decomposed into all of its fractional units:

$$Q = \sum_{1 \leq a,b \leq n} \alpha_{a,b} E_{a,b}$$

Once again, in matrix terms; the $\alpha_{a,b}$ are the components of the matrix that represents $Q$.

So far, we have a matrix-like object that acts on displacements, with components with two indices, addition and scalar multiplication of these objects, but no multiplication of the objects itself. The objects are also limited to the equivalents of square matrices, which gives us two more potential properties: the determinant and eigenvalues.

## 5.1 Determinant

The equivalent of the determinant will be given by the *power* of a quotient, although this equivalence is somewhat hidden in the use of other definitions. This power is first introduced in 383:

**383. DEFINITION. (Power of a fraction) [10, p. 210]**

The relative product of the numerators of a fraction whose denominators form the system of original units I call the *power* of the fraction, and I denote the power of the fraction $Q$, if the number of its numerators is $n$, by $[Q^n]$, that is, if

$$Q = \frac{(a_1, a_2, \ldots, a_n)}{(e_1, e_2, \ldots, e_n)},$$

and $e_1, e_2, \ldots, e_n$ form the system of original units, I set

$$[Q^n] = [a_1 a_2 \ldots a_n].$$

To see what is computed here, we must trace back what it means to compute the relative product $[a_1 a_2 \ldots a_n]$. The definition of relative product in 94 (p. 53) is dependent on the orders of the factors and the order of the domain they are a part of. In the case that the orders (dimensions) of the factors sum to the order of the space, as is true for our case, the relative product is equal to the outer product. As we have seen in the previous chapter, this comes down to what he, in A2, calls the combinatorial product. So, to compute the power of a fraction, we have to compute the combinatorial product of the $a_i$.

This does not immediately seem helpful, as the combinatorial product was defined by its properties rather than a practical computation. Conveniently though, in 62 (p. 33) Grassmann introduces determinants entirely separately from his quotients, and the result of 63 will give exactly the relation between this determinant and combinatorial products. The determinant in 62 is that of "$n$ series of any $n$ numbers", and in his definition he essentially describes the Leibniz formula for determinants:

62. DEFINITION. (Determinant) [10, p. 33]

62. DEFINITION. By the *determinant* of $n$ series of any $n$ numbers one means, if one denotes the $r$th number of the $s$th series by $\alpha_r^{(s)}$, that polynomial one obtains from the product $\alpha_1^{(1)}\alpha_2^{(2)}\dots\alpha_n^{(n)}$ by interchanging the lower indices in all possible ways one by one, leaving the upper indices unchanged, then furnishes these products with a $+$ or $-$ sign, according as the number of those pairs of indices that are oppositely ordered relative to those above is even or odd, and adds this collection of terms. One denotes this determinant by $\sum \pm\alpha_1^{(1)}\alpha_2^{(2)}\dots\alpha_n^{(n)}$, that is, one sets

$$\sum \pm\alpha_1^{(1)}\alpha_2^{(2)}\dots\alpha_n^{(n)} = \sum(-1)^u\alpha_r^{(1)}\alpha_s^{(2)}\dots\alpha_w^{(n)},$$

where $r, s, \dots, w$ are equal to the numbers $1, 2, \dots, n$ taken in any order, the sum refers to all possible orderings of this type, and $u$ denotes the number of index pairs below that are oppositely ordered relative to those above.

In 63 he shows how the combinatorial product of $n$ factors of the form $\alpha_1^{(i)}a_1 + \dots + \alpha_n^{(i)}a_n$ relates to that of $a_1, \dots, a_n$. This turns out to use this determinant, giving the following relation:

$$\begin{aligned}[(\alpha_1^{(1)}a_1 + \dots + \alpha_n^{(1)}a_n)(\alpha_1^{(2)}a_1 + \dots + \alpha_n^{(2)}a_n)\dots(\alpha_1^{(n)}a_1 + \dots + \alpha_n^{(n)}a_n)]\\ = \sum \pm\alpha_1^{(1)}\alpha_2^{(2)}\dots\alpha_n^{(n)}\cdot[a_1a_2\dots a_n]\end{aligned}$$

With this in mind, let's look back at the power of a quotient. In 381 we had established the indices of our matrix, such that if

$$Q = \frac{(a_1, a_2, \dots, a_n)}{(e_1, e_2, \dots, e_n)} = \sum_{1\le a,b\le n} \alpha_{a,b}E_{a,b},$$

we have for each $i$ that $a_i = \alpha_{i,1}e_1 + \dots \alpha_{i,n}e_n$. The power $[Q^n] = [a_1 \dots a_n]$ then becomes

$$[(\alpha_{1,1}e_1 + \dots \alpha_{1,n}e_n)\dots(\alpha_{n,1}e_1 + \dots \alpha_{n,n}e_n)]$$

Which, by 63, is equal to

$$\sum \pm\alpha_{1,1}\alpha_{2,2}\dots\alpha_{n,n}\cdot[e_1e_2\dots e_n]$$

That is, the determinant of the $\alpha_{i,i}$ times the combinatorial product of the original units. If we could set the combinatorial product of the $e_i$ equal to 1, we could then conclude that

$$[Q^n] = det(\alpha_{i,i}) = det(Q)$$

This final identification might seem problematic. We had established earlier that the outer product of $n$ units forms a magnitude of $n$th order, whereas 1, a numerical magnitude, should be of 0th order. This is however where another part of the definition of relative product comes into play: in 94 the product of the $e_i$ is *defined* to 1. He resolves the discrepancy in orders in 95, where he shows that the order of a product should in fact be considered modulo the order of the domain. He proves that this preserves all of the properties he previously established and indeed this means that the product of $n$ displacements in a domain of $n$-th order is a number.

Although Grassmann does not precisely derive the multiplicative rule of determinants, he does get pretty close. In his definition of the power of a quotient he only defined it for those quotients with the original units as the denominator. In the result following that, he proves how to calculate it for general quotients, and shows how it follows the fractional structure again. Namely, for any quotient $Q = \frac{(b_1,\dots,b_n)}{(a_1,\dots,a_n)}$, its power is equal to

$$[Q^n] = \frac{[b_1 \dots b_n]}{[a_1 \dots a_n]}.$$

If we consider that $[a_1, \dots, a_n]$ is the power of $A = \frac{(a_1,\dots,a_n)}{(e_1,\dots,e_n)}$ and similarly $[b_1, \dots, b_n]$ is the power of $B = \frac{(b_1,\dots,b_n)}{(e_1,\dots,e_n)}$, we can consider what their corresponding matrices do. Since $Ae_i = a_i$ and $Qa_i = b_i$, the *matrix* product $QA$ would map $e_i$ to $b_i$, that is, it would represent the same mapping as multiplying by $B$ does. In terms of matrices then, we have $QA = B$ hence $\det Q \cdot \det A = \det B$, which leads us to the result $\det Q = \frac{\det B}{\det A}$ that Grassmann found.

Weyl, in his 'Space, Time, Matter', makes use of Grassmann's notation for the power of a quotient to obtain the multiplication theorem of determinants [21, p. 139], proving that he had indeed found inspiration in Grassmann's work.

## 5.2 Eigenvalues

With the determinant established, we can consider the concept of eigenvalues and how to determine them. This is the matrix concept that appears most directly in A2, as Grassmann defines in 387:

387. DEFINITION. (Principal number) [10, p. 214]

If a fraction $Q$ multiplied by a nonzero magnitude of first order yields a multiple of this magnitude, say the $\rho$-fold of it, so that

$$Qx = \rho x,$$

I call the coefficient $\rho$ (whether it is real or imaginary) a *principal number* of the fraction $Q$, and the domain to which all those magnitudes $x$ belong that satisfy that equation the *principal domain* of that principal number.

Translating 'fraction' in this definition to 'matrix' and 'nonzero magnitude of first order' to 'vector', we have here precisely the definition of an eigenvalue (principal number) and the eigenspace (principal domain) associated to it. It is even specified that the eigenvalues may be imaginary.

In 388 Grassmann describes how one can compute these eigenvalues, and both the method and its derivation are essentially identical to our modern approach. The identity $Qx = \rho x$ he rewrites to $0 = (\rho - Q)x$, as in his notation a scalar is also a quotient; in modern notation this would read $0 = (\rho I - Q)x$. Since $(\rho - Q)$ is a quotient, we can let $c_i$ for $1 \leq i \leq n$ be displacements such that $(\rho - Q)e_i = c_i$, and write our quotient in a fractional form: $(\rho - Q) = \frac{(c_1,\ldots,c_n)}{(e_1,\ldots,e_n)}$. If we then write $x = \sum \xi_i e_i$, we get that $(\rho - Q)x = \sum \xi_i c_i = 0$. Since $x$ is by definition non-zero this means that the $c_i$ (or, the columns of the matrix $(\rho I - Q)$) are dependent. The product of dependent factors equals 0, and so since $[(\rho - Q)^n] = [c_1 \ldots c_n]$, the power of $(\rho - Q)$ and hence the determinant of $(\rho I - Q)$ is equal to 0. From this identity he derives what we would now call the characteristic polynomial of $Q$:

$$\alpha_0 \rho^n - \alpha_1 \rho^{(n-1)} + \cdots + (-1)^n \alpha_0 = 0,$$

whose roots give us the $n$ principal numbers. Grassmann then concludes this section with the fact that the product of the $n$ principal numbers is equal to the power of $Q$.

What follows from the principal numbers does look somewhat different from the modern approach. Grassmann is not particularly interested in the eigenvectors belonging to an eigenvalue, but rather in the principal domain (eigenspace) belonging to it. How he obtains this space also looks different, and relies on the fact that his quotients are not tied to one particular basis.

Suppose we have some quotient $Q = \frac{(a_1,\ldots,a_n)}{(e_1,\ldots,e_n)}$ and we have found some principal number $\rho$. Then there exist one or more independent displacements $x_i$ that, multiplied by $\rho - Q$, give 0, say $m$ of them. These displacements can then replace part of the denominators in $\rho - Q$, that is, if we sort the $e_i$ appropriately, we can write

$$\rho - Q = \frac{(c_1, \ldots, c_n)}{(e_1, \ldots, e_n)} = \frac{(c_1, \ldots, c_{n-m}, 0, \ldots, 0)}{(e_1, \ldots, e_{n-m}, x_{n-m+1}, \ldots, x_n)}.$$

These $x_i$, then, form the basis of the principal domain belonging to $\rho$, and in modern notation these are the eigenvectors belonging to $\rho$.

It is earlier in 386 (p. 213) where Grassmann shows how to rewrite a quotient in this form, which in general is possible whenever the numerators are linearly dependent. Since we know $\rho$ is an principal number, we know that the $c_i$ will be dependent. We can then order them in such a way that the first $n - m$ form an independent system and the remaining $m$ are dependent on those. That means, that for every $i > n - m$, we can write

$$c_i = \alpha_{1,i} c_1 + \cdots + \alpha_{n-m,i} c_{n-m}.$$

This means that multiplying the quotient $\rho - Q$ by either the displacement $e_i$ or the displacement $\alpha_{1,i} e_1 + \cdots + \alpha_{n-m,i} e_{n-m}$ will give $c_i$ as the product. The displacement

$$x_i = \alpha_{1,i} e_1 + \cdots + \alpha_{n-m,i} e_{n-m} - e_i$$

multiplied by $\rho - Q$ then has 0 as the product and is thus part of the principal domain of $\rho - Q$. It is also independent of $e_1, \ldots, e_{n-m}$ and can thus replace $e_i$ in the denominator of the quotient. Following the same procedure for all $n - m < i \leq n$, we find the $m$ displacements $x_i$ that span the principal domain.

The method used here, is, in the end, not all that different from what we would do in modern notation. To find which of the $c_i$ are dependent on the others, and to find the $\alpha_{j,i}$ that relate them, the modern approach would be to solve the equation $(\rho I - Q)x = 0$, since the $c_i$ form the columns of the matrix $(\rho I - Q)$. This is the same equation we solve to find our modern eigenvectors. It is the conceptual idea behind the procedure, and the different notation it allows for, that really defines the difference.

## 5.3 Applications

Now that we have established some of the parallels between quotients and matrices, the question of application remains. What does Grassmann use his quotients for, and how do they hold up compared to the use of matrices?

In 382 (p. 210) and its following remark, Grassmann establishes that the quotient can represent systems of linear functions in multiple variables, although he does not show how this would help in solving the system. More interesting in the context of affine geometry is his second remark after 390 (p. 221), where he discusses the quotient as representing a collinear relationship.

If we continue to view the numerators and denominators of a quotient as vectors, then any quotient represents a linear transformation, just like matrices do. As stated before however, this is not the only possible interpretation. We can also consider points in space as first order magnitudes. With the numerators and denominators considered as points, the quotients represent a far broader set of transformations, without the need to 'augment' them like we do with matrices. In this discussion I will limit these to simple points, as they relate most directly to geometry.

In the remark on page 221, Grassmann discusses several properties of collinear relationships that one can derive from their quotient representation, and he defines several 'special' geometric relationships by the properties of their quotient. This is where the affine relationship appears in A2, where it is defined by the requirement that "to the infinitely distant points of each system there also correspond infinitely distant points of the other". Indeed, this would map 'directions' to directions, and since collinearity is preserved, this means that parallel lines remain parallel. I will show that any *invertible* quotient of finitely distant points also preserves infinitely distant points.

Let $A, B, C, D$ and $A', B', C', D'$ be any two sets of independent, finitely distant points, and let $E$ be an infinitely distant point. We can then derive $A', B', C', D', E$ from $A, B, C, D$:

$$\begin{aligned} A' &= a_1A + a_2B + a_3C + a_4D & \qquad a_1 + a_2 + a_3 + a_4 &= 1 \\ D' &= b_1A + b_2B + b_3C + b_4D & b_1 + b_2 + b_3 + b_4 &= 1 \\ C' &= c_1A + c_2B + c_3C + c_4D & c_1 + c_2 + c_3 + c_4 &= 1 \\ D' &= d_1A + d_2B + d_3C + d_4D & d_1 + d_2 + d_3 + d_4 &= 1 \\ E &= e_1A + e_2B + e_3C + e_4D & e_1 + e_2 + e_3 + e_4 &= 0 \end{aligned}$$

The quotient $Q = \frac{(A',B',C',D')}{(A,B,C,D)}$ then maps the point $E$ to

$$\begin{aligned}
E' =& e_1A' + e_2B' + e_3C' + e_4D' \\
=& e_1(a_1A + a_2B + a_3C + a_4D) + e_2(b_1A + b_2B + b_3C + b_4D) \\
&+ e_3(c_1A + c_2B + c_3C + c_4D) + e_4(d_1A + d_2B + d_3C + d_4D) \\
=& (e_1a_1 + e_2b_1 + e_3c_1 + e_4d_1)A + (e_1a_2 + e_2b_2 + e_3c_2 + e_4d_2)B \\
&+ (e_1a_3 + e_2b_3 + e_3c_3 + e_4d_4)C + (e_1a_4 + e_2b_4 + e_3c_4 + e_4d_4)D
\end{aligned}$$

The coefficients of which add to

$$\begin{aligned}
& e_1a_1 + e_2b_1 + e_3c_1 + e_4d_1 + e_1a_2 + e_2b_2 + e_3c_2 + e_4d_2 \\
&+ e_1a_3 + e_2b_3 + e_3c_3 + e_4d_4 + e_1a_4 + e_2b_4 + e_3c_4 + e_4d_4 \\
=& e_1(a_1 + a_2 + a_3 + a_4) + e_2(b_1 + b_2 + b_3 + b_4) + e_3(c_1 + c_2 + c_3 + c_4) + e_4(d_1 + d_2 + d_3 + d_4) \\
=& e_1 + e_2 + e_3 + e_4 \\
=& 0
\end{aligned}$$

Thus $Q$ maps infinitely distant points to infinitely distant points. An entirely analogous derivation for $\frac{1}{Q}$ shows that indeed infinitely distant points in each system correspond to infinitely distant points in the other. Therefore any invertible quotient of finitely distant, simple points, represents an affine transformation.

Grassmann follows the introduction of the affine relationship by naming several more special relations. For instance, he proposes a "special type of affinity" where the product of the principal values, that is the discriminant, is equal to 1, in which case areas and volumes are preserved. Unfortunately, aside from this one remark, Grassmann does not explore the representation of transformations any further.

## 5.4 Comparing quotients and matrices

All in all, Grassmann's approach provides an interesting alternative to our modern one. One of its strengths lies in its generality; its general approach is the exact same regardless of whether its applied to vectors, points, or any other generic object of first (or $(n-1)$th) order. In my comparison to the modern approach I restricted the points to simple, finitely distant points, which put some restrictions on the quotients of points compared to those of vectors - allowing for multiple and infinitely distant points resolves these differences. Had Grassmann's approach found more popularity, it is conceivable that the concept of multiple ('weighted') points would have been further developed, and mass point geometry might have become incorporated in geometry as a whole.

The notation of quotients itself is both a strength and a weakness. The quotient notation is a sensible way to denote a linear map; a linear operation is fully defined by how it acts on any basis. If the operation acts by multiplication, then the operation itself is found by division, that is as the quotient of those known values. Several of its properties naturally follow from this definition as well. There is a freedom in being able to easily define an operation as how it acts on any basis, without immediately having to consider how it acts on some 'standard' basis. This freedom in basis, and the explicit notation of the one used, allows for some operations on quotients that are either not possible in matrices, or where our notation obscures what is happening geometrically. For instance, we typically apply elementary matrix operations to the rows of a matrix, for instance with Gaussian elimination. If we were to apply these to columns of a matrix rather than rows, the result essentially changes the basis we start from and tells us where this different basis gets mapped to. For matrices this is entirely hidden in the algebraic operations, whereas in Grassmann's notation we can see the numerators change and the denominators change accordingly, as he describes in 380. A particular application of this is how the eigenvectors of the eigenvalue $\rho$ of a quotient $Q$ show up in the denominator when one looks at the quotient $\rho - Q$ and adjusts it so that the numerator contains 0's.

As an example, suppose we want to find the eigenvectors of the matrix

$$\begin{pmatrix} -2 & -1 & 0 \\ 0 & 1 & 1 \\ -2 & -2 & -1 \end{pmatrix}$$

belonging to the eigenvalue 0. In modern notation, one would apply Gaussian elimination to

$$\left(\begin{array}{ccc|c} -2 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ -2 & -2 & -1 & 0 \end{array}\right)$$

to eventually arrive at

$$\left(\begin{array}{ccc|c} -2 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

and use this to conclude that the eigenvectors are given by any multiple of $(1, -2, 2)^t$. This process is perfectly valid algebraically, but has very little connection to the geometric transformation the matrix might represent.

To compare this to Grassmann's approach, we much first write this matrix as a quotient. The matrix above maps $e_1$ to $a_1 = (-2, 0, 2)^t$, $e_2$ to $a_2 = (-1, 1, -2)^t$ and $e_3$ to $a_3 = (0, 1, -1)^t$. The matrix can then be written as the quotient $\frac{(a_1, a_2, a_3)}{(e_1, e_2, e_3)}$. The question of the eigenvectors then becomes a question of which vectors get mapped to 0. To solve this, we must write $a_3$ as a sum of $a_1, a_2$, and indeed we can write that $a_3 = -\frac{1}{2}a_1 + a_2$. It follows that both $e_3$ and $-\frac{1}{2}e_1 + e_2$ get mapped to $a_3$, and thus that $x = -\frac{1}{2}e_1 + e_2 - e_3$ gets mapped to 0. As it is also independent of $e_1$ and $e_2$, the quotient can be rewritten to $\frac{(a_1, a_2, 0)}{(e_1, e_2, x)}$, and any multiple of $x$ is an eigenvector for the eigenvalue 0.

This method lacks the convenient algorithm that Gaussian elimination provides. The quotient notation however shows more clearly what the eigenvector means geometrically, as we end up rewriting the quotient such that it directly reveals the vector that gets mapped to 0. This is only possible because this notation is not limited to the standard basis. This method also does not require the coefficients of the matrix, the only thing we need to know is the relation between the $a_i$.

The generality and lack of coefficients does come at a price. Where the concepts and ideas behind it can be fairly intuitive if we let go of our modern framework, the same cannot be said for working with these objects in a practical context. To multiply any point or other object with a quotient, one needs to know the coordinates of this object with regards to whichever basis had been chosen for this particular quotient. That is, to multiply a vector $x$ with the quotient $\frac{(b_1, \ldots, b_n)}{(a_1, \ldots, a_n)}$, one would first have to write $x$ as a multiple some of the $a_i$. A solution could be to rewrite any quotient to having a standard basis as its denominators, but in doing so we are ignoring precisely the strength of generality that it had.

This problem also extends to Grassmann's version of determinant. Although he shows that these can be computed for any quotient regardless of numerators, no clear-cut computation is ever provided for the relative products one has to compute. The closest we get is in the relation between those products and the "determinant of $n$ series of $n$" numbers. To determine this, however, we still end up needing the coefficients of our numerator and denominator relative to some standard basis. Since we need to compute this for both numerator and denominator, this means twice the work with precisely those coefficients that this notation otherwise avoids altogether.

One more apparent limitation is that the quotients are only equivalent to square matrices, as any $n$ objects must map to $n$ objects. This is however not as bad as it seems, as the numerator does not need to be independent. The map can indeed not map into a subset of a higher dimensional space. However when the objects in the numerator are linearly dependent, they span a lower dimensional space and the quotient essentially maps into that space. In fact one can rewrite the quotient in the same way as when determining the eigenvectors, to find the kernel of the map in the denominator.

All in all, the quotient notation that Grassmann introduces does have some advantages compared to our modern transformation matrix. However I do not think that these advantages outweigh the limitations that they have, particularly in their lack of coefficients making any numerical approach unnecessarily difficult.

# Chapter 6

# Conclusion

In 1918, we find the first systematic approach to affine geometry in Hermann Weyl's 'Space, Time, Matter' [21]. He makes use of a rather modern axiomatic approach, and even where his axioms diverge from the now common ones, correspondences are easily established. In fact some modern sources still use his axioms, now referred to as "Weyl's axioms" [15]. Some conceptual differences still remain. He does not differentiate as strongly between the affine space and its underlying vector space as we do now, although he does make mention of both a vector field and a point configuration. His spaces are also limited to finite dimensions, the generalization to infinite-dimensional spaces being a more modern one.

In his work, the axiomatic approach has a clear purpose in establishing a rigorous foundation for the later chapters, in which he uses the affine geometry to discuss the general theory of relativity. He points this out on page 159, where he states that "We have set forth these details with pedantic accuracy so as to be armed at least with a set of mathematical conceptions which have been sifted into a form that makes them immediately applicable to Einstein's principle of relativity for which our powers of intuition are much more inadequate than for that of Galilei." Indeed, this rigor serves as a foundation to work on theories where intuition is no longer adequate, and might even lead one astray. The quest for rigor is however not his only aim. In the preface, Weyl mentions his desire to create a systematic presentation of the general theory of relativity, which he says was lacking up to this point. This indicates his more descriptive aim, where he takes great care in how he presents his material so the reader can follow along, with great attention to the foundations. On the topic of affine geometry this is particularly noticeable in how he moves back and forth between the abstract, context-free axioms, and the discussions before and after where he does introduce context and encourages the reader to build their intuition. [3]

Hermann Grassmann's Ausdehnungslehre [8,10] is quite a different story. First of all, the intentions behind this work were very different. Through his theory of forms, Grassmann had a strong conviction of how mathematics should be built up, and as such creating that foundation served to fulfill an architectural aim. More than that, though, Grassmann aimed to develop a new branch of mathematics. In his foreword he argues how abstraction and generalization reveal the patterns and elegance of mathematical theories, which in turn made it easier to connect different concepts and eventually apply them. As such, his abstract approach served as a tool for conceptual analysis, a tool for understanding his new branch of mathematics [3]. This was however not a goal he managed to achieve, as it turned out that it was precisely this approach that stood in the way of his contemporaries understanding his work.

From a modern perspective, it might seem that the abstract approach also hindered Grassmann himself. Through careful reading we may certainly uncover many affine concepts in Grassmann's work, and he deserves particular credit for the early treatment of $n$-dimensional spaces. Many properties however get lost in the overly general treatment, and the desire in A1 to develop extension theory independently from other areas of mathematics. Grassmann did not seem to realize the significance of some of the more specific concepts he developed, which is particularly noticeable when many properties of vectors are overshadowed by the study of higher dimensional magnitudes. Even the numerical magnitudes, which ended up functioning as a generalization of numbers, could not be fully understood without these higher dimensional magnitudes.

The question of whether Grassmann was indeed the first to develop affine geometry deserves, I believe, a two-fold answer. On the one hand, he did indeed develop a new approach to geometry and mathematics in general, and its foundation uses many affine concepts. It is entirely possible that, as mathematics progressed, Grassmann's work did inspire mathematicians like Weyl that brought affine geometry to the area of mathematics that we know today. On the other hand, the work that Grassmann wrote is not in itself a foundation for affine geometry, and to say that it is would be disingenuous. However it is hard to say what extension theory might have developed into if Grassmann had gotten earlier recognition and more time to expand on it. In chapter 5 I discussed some of the advantages Grassmann's notation of quotients might have had compared to our modern transformation matrices. Possible follow-up research could involve extending this comparison to other areas of extension theory. Higher order magnitudes in particular seem to have no meaningful equivalent in modern affine geometry. Further investigation on what application these could have might reveal whether we lost a brilliant theory in neglecting Grassmann's work, or whether the eventual approach that affine geometry took was indeed the superior one.

# Appendix A

# ‘Translations’ of Grassmann’s terms

Below is a list of terms used in Grassmann’s work, and how they can be interpreted in modern affine geometry. The page number refers to the page where I discuss this correspondence in this thesis. These correspondences are not a direct translation. In many cases they behave similarly, but the concept behind them is different. In some cases, I have chosen a modern concept that is more narrow than the one Grassmann considers.

| **Grassmann** | **Modern** | **Page** |
|---|---|---|
| conjunction | (binary) operation | 16 |
| displacement | vector | 18 |
| domain ($n$th order) | $n$-dimensional space | 28 |
| element | point/position | 17 |
| elementary extensive structure | line segment | 18 |
| elementary magnitude (first order) | weighted point | 24 |
| evolution | motion | 17 |
| extensive magnitude (first order) | vector | 18 |
| extensive magnitude ($n$th order) | $n$-dimensional parallelepiped | 21 |
| fundamental evolution | direction | 18 |
| indicators of a magnitude | coordinates of a vector | 24 |
| magnitude | vector/point/plane/space/etc. | |
| magnitude (first order) | 1-dimensional object (point/vector) | 30 |
| numerical magnitudes | (real) numbers/scalars | 23 |
| power of a quotient | determinant of a matrix | 32 |
| principal domain | eigenspace | 34 |
| principal number | eigenvalue | 34 |
| quotient/fraction | square matrix | 30 |
| reference measures | unit vectors | 23 |
| reference system | basis of a space | 23 |
| similar displacements | parallel vectors | 19 |
| similar extensions | parallel vectors/parallelograms/etc | 21 |

| | | |
|---|---|---|
| system (first order) | one-dimensional space | 18 |
| system ($n$th order) | $n$-dimensional space | 19 |
| system of units | basis of a space | 28 |

# Bibliography

[1] Paola Cantù, *Grassmann's epistemology: multiplication and constructivism*, From past to future: Graßmann's work in context: Graßmann bicentennial conference, september 2009, 2010, pp. 91–100.

[2] ______, *Grassmann's concept structuralism*, The Prehistory of Mathematical Structuralism (2020), p. 21.

[3] Paola Cantu, *What is axiomatics?*, Annals of Mathematics and Philosophy **1** (2023), pp. 1–24.

[4] Michael J Crowe, *A history of vector analysis: The evolution of the idea of a vectorial system*, University of Notre Dame Press, 1967.

[5] John A Dossey, Charles Vanden Eynden, Albert D Otto, and Lawrence E Spence, *Discrete mathematics*, 4th ed., Addison-Wesley Educational Publishers, Inc., 2001.

[6] Friedrich Engel, *Hermann grassmanns gesammelte mathematische und physikalische werke: Grassmann's leben*, Vol. 3, BG Teubner, 1911.

[7] Leonhard Euler, *Introductio in analysin infinitorum*, Vol. 2, Apud Marcum-Michaelem Bousquet & Socios, 1748.

[8] Hermann Grassmann, *Die lineale ausdehnungslehre ein neuer zweig der mathematik: dargestellt und durch anwendungen auf die übrigen zweige der mathematik, wie auch auf die statik, mechanik, die lehre vom magnetismus und die krystallonomie erläutert*, 2nd ed., Open Court Publishing Company, 1844/1878. English translation by L.C. Kannenberg, 1995.

[9] ______, *Lehrbuch der arithmetik für höhere lehranstalten*, Th. Chr. Fr. Enslin, 1861.

[10] ______, *Die ausdehnungslehre. vollständig und in strenger form bearbeitet*, Vol. 1, American Mathematical Society, 1862. English translation by L.C. Kannenberg, 2000.

[11] JG Grassmann, *Zur physischen krystallonomie und geometrischen combinationslehre, stettin, bei friedr*, Heinr. Morin (1829).

[12] F. Klein, A. Sommerfeld, and P. Furtwängler, *Ausgewählte kapitel der zahlentheorie*, 1896.

[13] F. William Lawvere, *Review of [hermann günter grassmann, a new branch of mathematics, "the ausdehnungslehre of 1844," and other works, translated by lloyd c. kannenberg, with foreword by albert c. lewis, open court, 1995. hermann günter grassmann, extension theory, "the ausdehnungslehre of 1862," translated and with a foreword and notes by lloyd c. kannenberg, history of mathematics, vol. 19, american mathematical society/london mathematical society, 2000.]*, Historia Mathematica **32** (2005), no. 1, pp. 99–106.

[14] August Ferdinand Möbius, *Der barycentrische calcul ein neues hülfsmittel zur analytischen behandlung der geometrie dargestellt und insbesondere auf die bildung neuer classen von aufgaben und die entwickelung mehrerer eigenschaften der kegelschnitte angewendet von august ferdinand mobius professor der astronomie zu leipzig*, Verlag von Johann Ambrosius Barth, 1827.

[15] Katsumi Nomizu and Takeshi Sasaki, *Affine differential geometry: geometry of affine immersions*, Cambridge university press, 1994.

[16] Giuseppe Peano, *Geometric calculus, according to the ausdehnungslehre of h. grassmann*, Birkhauser Boston, 1888. English translation by L.C. Kannenberg, 2000.

[17] Hans-Joachim Petsche, *Hermann grassmann: Biography*, Birkhäuser, 2009. English translation by M. Minnes, 2009.

[18] Steven Roman, S Axler, and FW Gehring, *Advanced linear algebra*, Vol. 3, Springer, 2005.

[19] Christoph J Scriba and Peter Schreiber, *5000 years of geometry: Mathematics in history and culture*, Birkhäuser, 2010. English translation by J. Schreiber, 2015.

[20] Agustí Reventós Tarrida, *Affine maps, euclidean motions and quadrics*, Springer, 2011.

[21] Hermann Weyl, *Space, time, matter*, 4th ed., Dover Publications, inc., 1918/1921. English translation by H.L. Brose, 1922.

[22] Edmund Taylor Whittaker, *From euclid to eddington: A study of conceptions of the external world*, Dover Publications, 1958.

[23] Hans Wussing, *The genesis of the abstract group concept: a contribution to the history of the origin of abstract group theory*, MIT Press, 1969. English translation by A. Shenitzer, 1984.