Part I

Background Topics

1 Linear Algebra

1.1 *Preface*

THE GOAL OF THIS CHAPTER is not to teach you, the reader, linear algebra from scratch - nor to be a thorough source of information on the topic. Rather, my aim is to introduce important "advanced" concepts for those who took a basic linear algebra course as part of an undergraduate university program. These concpets should help you gain a basic knowledge of the topics needed for understanding the rest of the background material, as well as the topic of spinors itself.

My approach to teching topics in linear algebra - and in mathematics as a whole - is to first build an intuition and only then formalize and generalize the ideas as needed. In my personal experiences, when I was studying linear algebra I completely failed to understand it (and indeed, failed the course) until it "clicked" for me in regards to 2- and 3-dimensional real spaces, i.e. - visible geometry. After that I didn't even have to study for exams anymore, as everything became clear enough to grasp and develop on the spot even during an exam (except for later, more advances concepts). That is why, for example, I absolutely adore courses and study materials of the topic[1] which use animation, such as *3Blue1Brown* great video essay series Essence of linear algebra[2].

There are very few proofs in this chapter, and those that are shown are not completely rigorous. For more in-depth materials, see the last section (further read). With that out of the way - let's begin!

[1] And other mathematical topics as well.
[2] Temporary sidenote which should become a citation for the mentioned 3B1B video series

1.2 Change of Coordinates

In introductory linear algebra courses you should have learned about change of coordinate systems: a coordinate system is just another name for a basis set of whatever vector space is used (in this section it's $\mathbb{R}^n$). A change of coordinate system is the transformation of vectors from being represented in one basis set $B = \{e_1, e_2, \dots, e_n\}$ to being represented in another basis set $\tilde{B} = \{\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n\}$. Since such transformations are linear they are commonly represented in a matrix form.

In this section we will discuss *how* vectors and their components change under change of basis sets. There are many components involved in these kind of transformations, which causes them to be quite confusing. I will therefore color code the equations consistently as a visual guide. In addition, I will always introduce the $\mathbb{R}^2$ case first, before giving the generalized form for $\mathbb{R}^n$.

1.2.1 Change of basis set in $\mathbb{R}^2$

Suppose we use the standard basis set to represent $\mathbb{R}^2$:

$$B = \{e_1, e_2\} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \tag{1.1}$$

and we want to change our coordinate system to use the following basis set:

$$\tilde{B} = \{\tilde{e}_1, \tilde{e}_2\} = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{4} \end{bmatrix} \right\}. \tag{1.2}$$

(the two basis sets are shown in Figure 1.1)

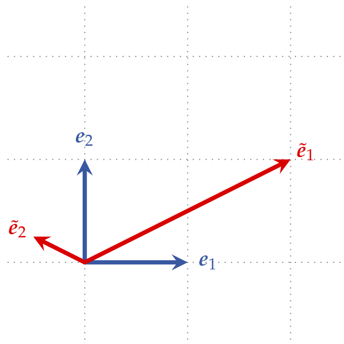

Figure 1.1: The standard basis set B and a new basis set $\tilde{B}$ shown together.

$$F = \begin{array}{c} \begin{array}{cc} \tilde{e}_1 & \tilde{e}_2 \\ \downarrow & \downarrow \end{array} \\ \begin{bmatrix} 2 & -\frac{1}{2} \\ 1 & \frac{1}{4} \end{bmatrix} \end{array}. \tag{1.3}$$

Now, if we want to transform $\tilde{e}_1$ and $\tilde{e}_2$ into $\tilde{e}_1$ and $\tilde{e}_2$ using F, we simply multiply them by F:

$$\begin{aligned} \tilde{e}_1 &= F e_1, \\ \tilde{e}_2 &= F e_2. \end{aligned} \tag{1.4}$$

To make **??** more concise, we can collect the two vectors into a matrix disguised as a row vecor:

$$e = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} e_1, e_2 \end{bmatrix}. \tag{1.5}$$

We then get that **??** can be written in vector-matrix notation as

$$\tilde{e} = \begin{bmatrix} e_1, & e_2 \end{bmatrix} = \begin{bmatrix} e_1, & e_2 \end{bmatrix} \begin{bmatrix} 2 & -\frac{1}{2} \\ 1 & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} 2e_1 + e_2, & -\frac{1}{2}e_1 + \frac{1}{4}e_2 \end{bmatrix}. \quad (1.6)$$

The reverse transformation can be calculated by applying the inverse transformation $\boldsymbol{F}^{-1}$ on Equation 1.6:

$$\begin{bmatrix} \tilde{e}_1, & \tilde{e}_2 \end{bmatrix} \boldsymbol{F}^{-1} = \left(\begin{bmatrix} e_1, & e_2 \end{bmatrix} \boldsymbol{F} \right) \boldsymbol{F}^{-1} = \begin{bmatrix} e_1, & e_2 \end{bmatrix}. \quad (1.7)$$

(where in our case $\boldsymbol{F}^{-1} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} \\ -1 & 2 \end{bmatrix}$)

To summarize: given a set B of basis vectors in $\mathbb{R}^2$, we can transform them into the basis vectors set $\tilde{B}$ by using the forward transformation $\boldsymbol{F} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$. To get back the original basis vectors from the transformed vectors we use the inverse transfomation $\boldsymbol{F}^{-1} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}^{-1}$. If the matrix $Forw$ is invertible - i.e. if its two columns are not linearly depended (equivalently, if $F_{11}F_{22} \neq F_{12}F_{21}$), then $\boldsymbol{F}^{-1} = \frac{1}{F_{11}F_{22} - F_{12}F_{21}} \begin{bmatrix} F_{22} & -F_{12} \\ -F_{21} & F_{11} \end{bmatrix}$.

1.2.2 *The more general case,* $\mathbb{R}^n$

In $\mathbb{R}^n$ the transformations behave in a similar way: given the transformation rule that each new basis vector $\tilde{e}_i \in \tilde{B}$ is a linear combination of the old basis vector set B, i.e.

$$\begin{aligned} \tilde{e}_1 &= F_{11}e_1 + F_{12}e_2 + \cdots + F_{1n}e_n, \\ \tilde{e}_2 &= F_{21}e_1 + F_{22}e_2 + \cdots + F_{2n}e_n, \\ \vdots &= \vdots \\ \tilde{e}_n &= F_{n1}e_1 + F_{n2}e_2 + \cdots + F_{nn}e_n, \end{aligned}$$

we can write the transformation in matrix form as

$$\begin{bmatrix} \tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n \end{bmatrix} = \begin{bmatrix} e_1, e_2, \ldots, e_n \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} & \ldots & F_{1n} \\ F_{21} & F_{22} & \ldots & F_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ F_{n1} & F_{n2} & \ldots & F_{nn} \end{bmatrix}. \quad (1.8)$$

Per new basis vector $\tilde{e}_i$ Equation 1.8 has the form

$$\tilde{e}_j = \sum_{k=1}^{n} F_{kj} e_k. \quad (1.9)$$

Similarly, the inverse operation is given by

$$\begin{aligned}
e_1 &= F^{-1}_{11}\tilde{e}_1 + F^{-1}_{12}\tilde{e}_2 + \cdots + F^{-1}_{1n}\tilde{e}_n,\\
e_2 &= F^{-1}_{21}\tilde{e}_1 + F^{-1}_{22}\tilde{e}_2 + \cdots + F^{-1}_{2n}\tilde{e}_n,\\
\vdots &= \vdots\\
e_n &= F^{-1}_{n1}\tilde{e}_1 + F^{-1}_{n2}\tilde{e}_2 + \cdots + F^{-1}_{nn}\tilde{e}_n,
\end{aligned}$$

its matrix form is

$$\left[e_1, e_2, \ldots, e_n\right] = \left[\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n\right] \begin{bmatrix} F^{-1}_{11} & F^{-1}_{12} & \ldots & F^{-1}_{1n} \\ F^{-1}_{21} & F^{-1}_{22} & \ldots & F^{-1}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ F^{-1}_{n1} & F^{-1}_{n2} & \ldots & F^{-1}_{nn} \end{bmatrix}, \tag{1.10}$$

and per basis vector the transformation is

$$e_i = \sum_{j=1}^{n} F^{-1}_{ji}\tilde{e}_j. \tag{1.11}$$

By subtituting Equation 1.9 into Equation 1.11, we get

$$\begin{aligned}
e_i &= \sum_{j=1}^{n} F^{-1}_{ji}\tilde{e}_j \qquad (1.12)\\
&= \sum_{j=1}^{n} F^{-1}_{ji}\left(\sum_{k=1}^{n} F_{kj}e_k\right)\\
&= \sum_{k=1}^{n} \underbrace{\left(\sum_{j=1}^{n} F_{kj}F^{-1}_{ji}\right)}_{\text{this is just a number!}} e_k.
\end{aligned}$$

?? just tells us something we already know: each basis vector e_i equals to a linear combination of the *same set* of basis vectors. This must mean that for $k = i$ the number in paranthesis is one, and for any other value of k it is zero - i.e. it equals δ_{ik}:

$$e_i = \sum_{k=1}^{n} \delta_{ik}e_k. \tag{1.13}$$

In turn, Equation 1.13 means that the matrix $\boldsymbol{F}\boldsymbol{F}^{-1} = \boldsymbol{I}_n$, the identity matrix in $\mathbb{R}^n$ - and thus the matrices $\boldsymbol{F}$ and $\boldsymbol{F}^{-1}$ are eachother's inverses.

1.3 Dual Vectors and Dual Spaces

1.3.1 Linear measurements and rulers (or: why do we care about dual vectors?)

Frequently in linear algebra we want to measure vectors. A measure in this context is a way to assign each vector a real number which somehow reflects its properties (i.e. direction and/or magnitude). Mathematically, we are looking for a function $\phi : \mathbb{R}^n \to \mathbb{R}$ - we feed it a vector, and it returns some measurement.

One common way to measure vectors is using the norm: the norm of a vector in $\mathbb{R}^n$ using the standard basis set $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ is given by

$$\|\boldsymbol{v}\| = \sqrt{v_1^2 + v_2 + \cdots + v_n^2}, \tag{1.14}$$

where the numbers $v_1, v_2, \ldots, v_n$ are the components in each of the respective basis vectors.

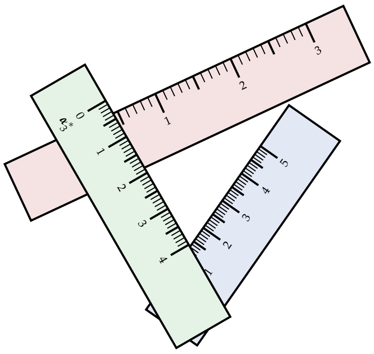

Figure 1.2: Three rulers, $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*, \boldsymbol{\alpha}_3^*$, each with its own orientation and frequency of graduation lines.

While the norm does tell us something useful about a vector, it has two drawbacks: it doesn't tell us anything about the vector's orientation in space, and even worse - it is not linear. We can try and simplify the calculation of the norm by dropping the square root and calculating the *square norm*, but that too isn't linear, nor does it tell us anything about the vector's orientation.

Another way of measuring vectors is by using *rulers*. Rulers are nothing more than a set of graduation lines with an orientation in space (**??**). We can therefore represent a ruler using a vector: the magnitude of the vector is the frequency of the graduation lines, while its direction is the direction of the ruler (which is orthogonal to the graduation lines). To not confuse ruler vectors with "regular" vectors, I will denote ruler vectors using a greek letter and an asterisk, e.g. $\boldsymbol{\alpha}^*$.

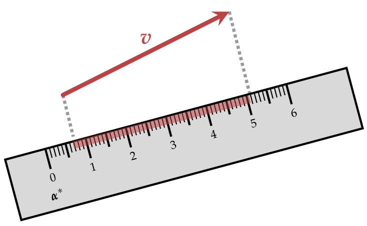

Figure 1.3: Measuring a vector using a ruler by projecting the vector onto the ruler. The result of the projection is drawn as a red line on the ruler.

Now, we usually rotate rulers to align with the orientation of the magnitude we wish to measure - however here we want our rulers to also measure orientation. Therefore, instead of rotating the ruler to align with a vector we wish to measure, we *project* the vector on the ruler and then take our measurement (**??**).

Since projection in linear algebra is calculated via the inner product, the measurement m of a vector $\boldsymbol{v}$ using a ruler $\boldsymbol{\alpha}^*$ is given by

$$m = \langle \boldsymbol{\alpha}^*, \boldsymbol{v} \rangle. \tag{1.15}$$

However, since we can represent a ruler using a vector, we can use the vector representation in **??**. And again, to distinguish ruler vectors from "regular" vectors written in an excplicit form, we write the ruler vectors as row vectors:

$$\langle \boldsymbol{\alpha}^*, \boldsymbol{v} \rangle = \begin{bmatrix} \alpha_1, & \alpha_2, & \ldots, & \alpha_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \tag{1.16}$$
$$= \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n.$$

Another reason for representing rulers as row vectors is that it corresponds well to matrix multiplication: if we regard a row vector as a $1 \times n$ matrix, and a column vector as an $n \times 1$ matrix, the resulting product is a 1×1 matrix, which is equivalent to a scalar. This will become handy when we discuss higher-order rulers (which in turn can handle higher-order "vectors").

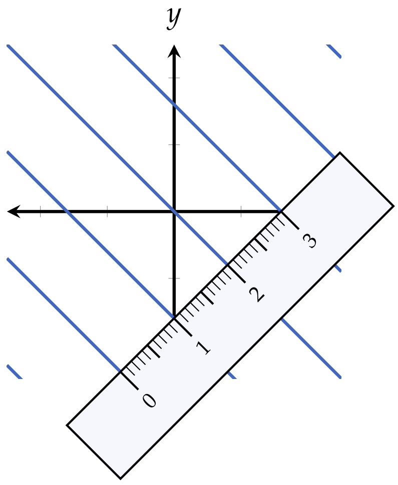

Figure 1.4: The graphical representation of a ruler shown next to the ruler, such that the infinite set of lines drawn match the ruler's graduation marks.

A convinient visualization for rulers in $\mathbb{R}^2$ is their representation as a set of parallel lines which are the extensions of the ruler's graduation marks. We draw the first line going through the origin and orthogonal to the orientation of the vector representation of the ruler. Then, we draw a line for each integer multiple of the frequency of the ruler's graduation marks. This results in a set of evenly spaced and parllel lines, which fit precisely with the main graduation mark of the ruler they represent (**??**).

In $\mathbb{R}^3$ a ruler is represented in a similar way, but with *planes* instead of lines (**??**). And in general, the rulers for a space $\mathbb{R}^n$ can be represented using an infinite set of $(n-1)$-dimensional *hyperplanes*.

From **??** we can see that any ruler $\boldsymbol{\alpha}^*$ can be written as the left half of an inner product, i.e. $\langle \boldsymbol{\alpha}^*, \cdot \rangle$. This hints that the ruler is *acting* on vectors, yielding a scalar. Looking at the explicit form of a ruler acting on a vector, we can see that not only do rulers represent linear functions, but they in-fact represent *all* the possible *linear* functions of the type $\phi : \mathbb{R}^n \to \mathbb{R}$. Such functions are called *linear forms*, or more specifically in this case a *1-form*.

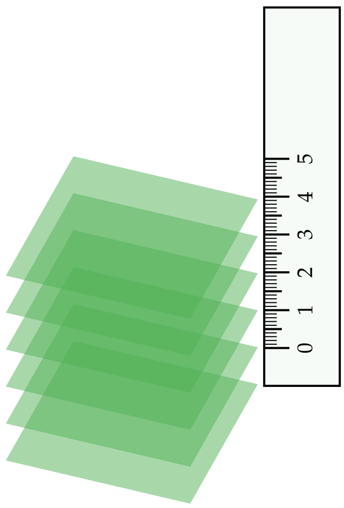

Figure 1.5: Representation of rulers in $\mathbb{R}^3$ as planes.

Since rulers can be represented as vectors but aren't exactly the same as "regular" vectors, they are also called *dual vectors*, which is the name I will use for them from now on. For any vector space $\mathbb{R}^n$, the dual vectors form a vector space of their own, called the *dual space to* $\mathbb{R}^n$.

The linearity of the dual vectors space simplifies many calculations - foremost, it allows us to construct dual vectors from other dual vectors using linear combinations. See **??** below.

Example 1.1 Linear combination of dual vectors

Suppose we have two dual vectors

$$\boldsymbol{\alpha}^* = \begin{bmatrix} 1, & -4, & 0 \end{bmatrix}, \qquad (1.17)$$
$$\boldsymbol{\beta}^* = \begin{bmatrix} 2, & 1, & 3 \end{bmatrix}.$$

We can use $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$ to create a third dual vector γ^*:

$$\begin{aligned}\gamma^* &= 5\begin{bmatrix} 1, & -4, & 0 \end{bmatrix} + 2\begin{bmatrix} 2, & 1, & 3 \end{bmatrix} \\ &= \begin{bmatrix} 5+4, & -20+2, & 0+6 \end{bmatrix} \\ &= \begin{bmatrix} 9, & -18, & 6 \end{bmatrix}.\end{aligned} \qquad (1.18)$$

The application of γ^* on any vector $\boldsymbol{v} \in \mathbb{R}^3$ is then given by

$$\langle \gamma^*, \boldsymbol{v} \rangle = 9v_1 - 18v_2 + 6v_3, \qquad (1.19)$$

which is exactly the application of $\boldsymbol{\alpha}^*$ on the vector five times followed by the application of $\boldsymbol{\beta}^*$ twice (try it for yourself!).

1.3.2 *Introducing some formalism*

Now that we know *why* we use dual vectors, we can formalize the concept and find useful properties.

Definition 1.1 Dual space

Given a vector space V over a field $\mathbb{F}$, its *dual space*, denoted V^*, is the set of all the linear functions $\phi : V \to \mathbb{F}$. The elements of V^* are called *dual vectors*.

π

A dual space equipped with a closed addition operation between any two of its elements and a closed product between its elements and the elements of the field $\mathbb{F}$ is itself a *vector space*, with similar structure to the space V.

Challenge 1.1 Dual space as a vector space

Show that a dual space euipped with addition and scalar product as defined above is indeed a vector space (Use the definition XXX).

?

To write:

1. Examples of dual vectors of functions?..

1.3.3 Basis sets and coordinate transformations

1. Dual basis: converting from a basis set in V to its dual in V^*.
2. Covariance of dual vectors basis change vs. contra-varience of vectors.

1.4 Further Reading

2 Geometric Algebra

2.1 *Preface*

This is a temp text.

3 *Abstract Algebra*

3.1 *Preface*

This is a temp text.

4 Lie Groups and Algebras

4.1 Preface

This is a temp text.

Part II

Spinors