TENSORS

For Inquiring Minds

Yury Deshko

Tensors For Inquiring Minds
Yury Deshko www.srelim.com
ISBN 978-1-7948-2018-0
Copyright © 2025

9 781794 820180

All rights reserved.
Imprint: Lulu.com

To Mom

With Love

That only son can have

Contents

Acknowledgment

My sincere gratitude goes to all reviewers of the early drafts of the book for their valuable feedback. In particular, I want to thank Dr. Alex Rylyakov, Dr. Mikhail Makouski, Prof. Anton Kananovich, and Dr. Mohammad Teimourpour. To Dr. Teimourpour I must give separate thanks for numerous discussions, helpful suggestions on material presentation, and hospitality.

My friends, discussing the book with you was both illuminating and fun.

Finally, special acknowledgment must be given to my son, Daniel, for his help with fixing colors in many figures.

Yury Deshko
Weehawken, New Jersey
2024

Preface

This book is the result of lectures delivered to curious, motivated, and studious high schoolers. The lectures ran during the years 2019-2024 in various formats, but mostly in class during a three week summer school organized by Columbia University Pre-College Programs. Additionally, the same lectures were taught remotely to selected students of Ukrainian Physics and Mathematics Lyceum.

The material has been designed to be accessible to people with solid background in high-school algebra and physics (mostly mechanics). Several years of teaching to a relatively diverse set of students proved that nearly all material can be efficiently absorbed by most, provided diligent work is done on exercises and problem. The last fact confirms a well-known truism: *No real learning occurs without practice.*

Exercises are essential part of this book. They are carefully selected to help readers get better understanding of the material and they are also fully solved. The difficulty of the exercises varies from simple to quite challenging.

This book *is not a standard textbook* and lacks the breadth and rigor present in many excellent introductions into Quantum Physics. The best way to view this book is as a *bridge* between elementary and popular books and the more challenging college-level textbooks.

At Any Cost

To explain the subtitle of this book let us refer to the letter written by Max Karl Ernst Ludwig Planck to an American physicist

This book explores *tensors* – a type of mathematical objects that extends the notion of numbers and vectors. The method of exploration is deliberately chosen to resemble a journey. Starting from familiar grounds of numbers and operations with numbers, a reader will re-examine familiar concepts in a new light and then will arrive at new concepts gradually, connecting the dots along the way.

Although the topic of the book is mathematical, the exploration will lack proper mathematical rigor, aiming instead at simplicity, clarity, and the use of helpful analogies.

This book is **not intended** to substitute more serious textbooks on linear algebra or tensor algebra. Hopefully, the main benefit of reading this book – either before, or after, or in addition to other books on the subject – is that it should help lower the *"mental barrier"* we all encounter when learning new concepts, especially abstract mathematical concepts.

To comprehend and enjoy the material of this book the reader should have a solid knowledge of basic high-school algebra and an open and inquiring mind. The book is a bit longer that it could have been because all derivations are detailed and all exercises are fully solved.

Some sections are marked with an asterisk, for example **Transposition***. Those sections contain material that is either optional or a bit more advanced that usual. These sections can be skipped without significant impact on the main message of the book.

1. Introduction

Abstract In this chapter.

Quantum physics methods play an increasingly important role in many domains of science. Modern mathematical tools are numerous and require serious effort to master. The algebra and calculus of *tensors* are good examples of this.

The goal of this book is to explain tensors by showing them in action and in relation to less complicated mathematical objects, such as vectors and numbers. Understanding numbers and vectors is essential for understanding tensors, therefore the former two are discussed in details.

The development of concepts will happen the following direction:

$$\text{Numbers} \quad \rightarrow \quad \text{Vectors} \quad \rightarrow \quad \text{Tensors}.$$

We will start with reviewing numbers as the simplest mathematical objects, and will consider operations on numbers – functions – in a general and more abstract way than one usually does in school. Many abstract concepts related to numbers and functions will be useful for studying vectors and tensors.

From numbers and numeric functions we will move on to vectors. Vectors are closely connected to numbers and can't even be properly defined without the latter. Vectors are more powerful than numbers and represent the next step in the hierarchy of mathematical objects. Vectors and functions on vectors (operations or operators) provide many new concepts that are crucial for understanding of tensors.

Careful study of vectors and functions on vector (operators) will inevitably lead us to tensors. Tensors and vectors are as intimately related, as vectors and numbers. In fact, having studied the basics of tensor algebra, we will see that numbers, vectors, and tensors are conceptually connected. We will be able to recognize that numbers are[1] very reduced tensors; numbers are tensors of rank zero. In a similar sense, vectors are not completely reduced tensor; vectors are tensors of rank one. Therefore, the progression of the topics from numbers to tensors can be viewed as follows:

$$\begin{array}{lclcl} \text{Numbers} & \to & \text{Vectors} & \to & \text{Tensors.} \\ \text{Tensors}^{(0)} & \to & \text{Tensors}^{(1)} & \to & \text{Tensors}^{(2+)}\,. \end{array}$$

Here the superscript in parentheses indicates the rank of the tensor[2].

As we move from numbers to tensors, the level of abstraction increases. To a significant degree, the difficulty of understanding tensors is due to high level of abstraction used in the definition of tensors as mathematical objects. Abstraction is the price we pay for more powerful and versatile tools. But more powerful tools are needed as scientists address more and more advanced problems.

The inventions of numbers, algebra, and then calculus were monumental breakthroughs. The transition from numbers to vectors and then to tensors is a more natural process that occurred rather quickly on the scale of the history of science.

1.1 Who Needs Tensors?

Although physicists make heavy use of tensors, today thousands of scientists – not only physicists – use the mathematical methods of tensors. Tensor mathematics (the algebra and calculus of tensors) is a *tool*; it is a fitting tool for some problems, and not too fitting for others. This situation is quite analogous to *vector algebra and calculus*.

Tensor literacy will enrich you and will open doors to new problems and new methods of their analysis. The following historical episode illustrates the point well.

In October of 1912, Albert Einstein wrote in a letter to his physicist friend Arnold Sommerfeld:

> ***Einstein on General Relativity***
> I am now exclusively occupied with the problem of gravitation theory

[1] In a certain sense which will become clear after reading the book.

[2] Don't worry if the concept of *rank* seems unclear right now – it will be explained in due time.

> and hope, with the help of a local mathematician friend, to overcome all the difficulties. One thing is certain, however, that never in my life have I been quite so tormented. A great respect for mathematics has been instilled within me, the subtler aspects of which, in my stupidity, I regarded until now as a pure luxury. Against this problem [of gravitation] the original problem of the theory of relativity is child's play.

In the period from 1905 to 1916 Einstein was feverishly working on the General Theory of Relativity – the next best theory of gravity since Newton. The mathematics of general relativity is based on the calculus of tensors, created by Italian mathematicians Ricci-Curbastro and Levi-Civita roughly a decade before Einstein started working on the problem of gravity.

To overcome the mathematical difficulties, Einstein used the help of his friend and former classmate Marcel Grossmann, who was an expert in tensor calculus and non-Euclidean geometry. The general theory of relativity was the first physical theory to use the power of tensors (in combination with profound physical insights) to achieve remarkable breakthrough. Since then, the methods of tensor calculus and non-Euclidean geometries have been used in many physical theories and problems.

In the end of the book (Section **??** on page **??**) a less dramatic example is given. The example describes a real-world situation when the understanding of vectors and tensors lead to significant practical benefits.

1.2 Naive Notion of Tensors

We may think of tensors as some kind of "super-numbers." In what sense tensors are numbers and what makes them "super?"

Similar to numbers, tensors can be added and subtracted. Also, tensors can be "scaled" by multiplying them by "normal" numbers, like 2 or π.

Unlike numbers, tensors support richer set of operations. Given two tensors, we can "kind-of-multiply" them to get either a simple number as the result (*scalar product*, see Section 5.1), or we can get another tensor, somewhat "bigger" than the original two (*tensor product*, see Section 5.8.)

Looking at the evolution of the concept of number, we can see the series of steps to higher levels of *generality*, *efficiency* , and *abstraction*: From natural numbers, to whole numbers, to fractions, to real and then to complex numbers. At each step new *mathematical objects* are introduced that can be added and multiplied in a "usual way."

Tensors come as the result of quite natural evolution of "number-like mathematical objects." Tensors extend the notion of numbers, all the way through vectors into a new and very powerful realm. If numbers are "bare quantities," and vectors are "quantities with direction" (e.g., velocity in physics), then tensors are "quantities with shape."

Tensors are naturally and closely connected with numbers and vectors. In fact, numbers and vectors *are tensors!*

> ***Tensors Naively Defined***
> Tensors are *mathematical objects* that cover and extend the concepts of numbers and vectors. As more powerful mathematical objects, tensors support many algebraic operations, including addition, subtraction, and scaling by a number.
>
> Tensors might be viewed as "quantities with shape," in analogy with vectors – "quantities with direction."
>
> Numbers and vectors represent the lowest "tiers" (called *ranks*) in the hierarchy of tensors.
>
> Finally, tensors generalize the idea of *linear functions* (see subsection **??** on page **??** and also Section 4.2). Tensors of higher ranks can "act" on tensors of lower ranks in a very simple way, resembling familiar multiplication.

1.3 Example Definitions

Now what are tensors more rigorously? Can we give a short definition to this concept? Let us take a look at several examples and see whether they shed sufficient light. The definitions given below differ from each other, but they simply convey *the same idea in different ways.*

The Encyclopedia of Mathematics [3] provides the following definition:

Definition 1.1 ***Tensors Definition 1***

Tensor on a vector space V over a field k is an element t of the vector

[3] `https://encyclopediaofmath.org/wiki/Tensor_on_a_vector_space`

space

$$T^{p,q}(V) = (\otimes^p V) \otimes (\otimes^q V^*),$$

where $V^* = \mathrm{Hom}(V, k)$ is the dual space of V.

To understand this defintion we first need to understand what *vector space* is, what *field* is, what *dual* means, and what is going on with superscripts and circles (e.g., in $\otimes^q$).

Wolfram Math World[4] provides another view of tensors:

Definition 1.2 ☞ ***Tensors Definition 2***

An nth-rank tensor in m-dimensional space is a mathematical object that has n indices and m^n components and obeys certain transformation rules.

Here we encounter new concepts, such as: *rank of a tensor*, m*-dimensional space*, *indices*, *components*, and some kind of *transformation rules*. They all will be discussed later in the book.

Yet another definition can be found in the book *Encyclopedia Of Mathematics*[5]

Definition 1.3 ☞ ***Tensors Definition 3***

Just as a *vector* is a mathematical quantity that describes translations in two- or three-dimensional space, a tensor is a mathematical quantity used to describe general transformations in n-dimensional space. Precisely, if the locations of points in n-dimensional space are given in one coordinate system by $(x^1, x^2, \dots, x^n)$ and in a transformed coordinate system by $(y^1, y^2, \dots, y^n)$ (it is convenient to use superscripts rather than subscripts), then a "rank 1 contravariant tensor" is a quantity T, with single components, that transforms according to the rule:

$$T^i_{new} = \sum_{r=1}^{n} \frac{\partial y^i}{\partial x^r} T^r .$$

Again we face a wall of new concepts: *vectors*, *transformations*, n*-dimensional space*, *contravariant tensors* and their *components*. Without clear understanding of these concepts, it is impossible to learn what tensors are.

The definitions given above are typical. They provide a good way to *end* the study of tensors, to summarize everything learned about them.

[4] `https://mathworld.wolfram.com/Tensor.html`

[5] *Encyclopedia Of Mathematics*, James Tanton, Facts On File, Inc, 2005

However, they are not good starting points. It is better to arrive at the concept of tensor gradually, going from numbers, through vectors, to tensors. This path starts in the next chapters. But before we begin, a couple of preliminary remarks are needed.

1.4 Diagrams

Sometimes to illustrate mathematical concepts and *relations between them*, we will use diagrams. Diagrams are helpful in highlighting some general features of *mathematical structures*.

Let us study an example, shown in the Figure 1.1. A collection of cars in a parking lot – Figure 1.1(a) – can be schematically represented as a set of *points* Λ – Figure 1.1(b). Each point of the set Λ corresponds to a certain car in the parking lot. Apart from the set of cars Λ, the Figure 1.1(b) contains other sets, denoted as K, N, and B. These sets correspond to various values of cars' properties, such as color (set K), mileage (set N), and so on. The set B is a very important set – called *Boolean set*. Boolean set has just two points, one labeled t for True, the other labeled f for False.

Reduced to a point of the set Λ, a car looses its individuality and those parameters that make it unique (make, color, milage, etc.) This is remedied by using other sets, such as the set of various colors K or the set N of numbers that can represent mileage, and so on.

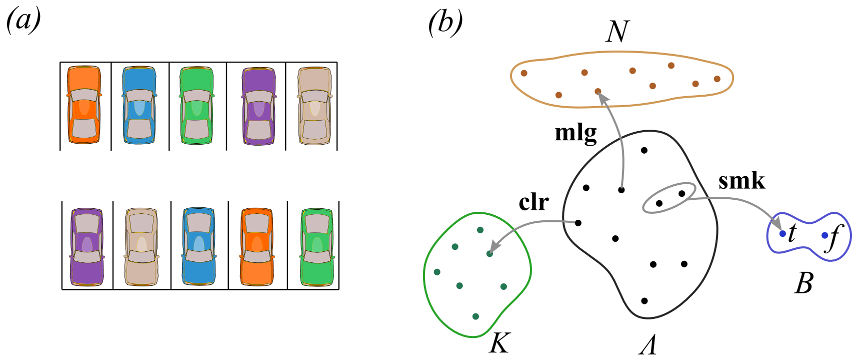


Fig. 1.1: Diagrams are used to graphically represent sets of objects and relationships between them. Arrows can connect (map) elements of one set with another. Such mappings may have names: **mlg** returns mileage for a given car, **clr** – color, and **smk** determines whether two cars are of the same make.

A particular property of a car-point can then be represented using

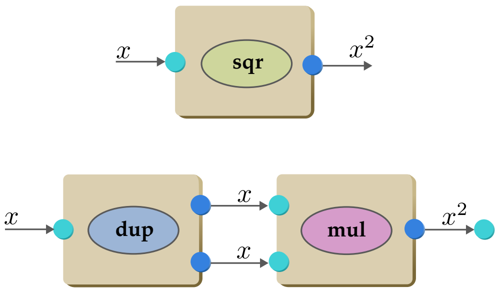


Fig. 1.2: Schematics can be used to represent functions, operators, their compositions and structure.

an arrow that connects the car-point to another point in the relevant set. We say that such an arrow *maps* points of one set into another set. The Figure1.1(b) shows three maps: **mlg** gives the mileage for each car from the set Λ, **clr** gives the color for each car, and **smk** compares whether two cars have the same make.

Exercise 1.1
Extend the diagram from the Figure 1.1(b), adding a set of different car makes (e.g., Ford, Toyota, Fiat, etc.) Come up with a mapping from this set into the Boolean set B.

1.5 Schematics

To illustrate the concepts of functions, operators, their structures and properties, we will be using schematics like the one shown in the Figure 1.2.

A simple schematic element is represented as a box with inputs and outputs. A box can have a name (label) which describes what the function does to its input. The number of inputs and outputs can vary depending on the complexity of a function.

Various "boxes" can be combined (or *composed*) to create a more complex structure. The Figure 1.2 shows how the outputs of a function **dup** (duplicate its input) are connected to the inputs of a function (**mul**) that multiplies its inputs. The result is a function that squares its input.

1.6 Sets and Tuples

We will use many notational conventions in this book. Most of them will be typical for mathematics: "+" denotes summation of two quantities, "=" means equality of two quantities, "*" – a multiplication, and parentheses in "3*(x+y)" are used to indicate the order in which operations should be performed (first add and then multiply).

At a certain point we will be discussing "assemblies"[6] of quantities. The simplest example – all natural numbers:

$$1, 2, 3, 4, 5, \ldots, n, \ldots .$$

We can also consider all letters of some alphabet:

a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z.

Both these examples can be formally considered as *sets* – an assembly of objects of similar kind. A set can given a special name when it is referred to often. For example, the set of natural numbers is denoted as

$$\mathbb{N} = \{1, 2, 3, \ldots\} .$$

Note the use of curly braces – they indicate that we are talking about a set. Thus, for a set $\mathbb{S}$ with elements a, b, x, and y we write

$$\mathbb{S} = \{a, b, x, y\} .$$

The concept of a set is one of the most basic in mathematics and the notation for sets is very standardized.

Another useful concept is called *tuple*. A tuple is a series of quantities that are related in a certain way, but can be of different kinds. It is better to study examples:

$$(3, \text{'a'}) \quad \text{– couple} ,$$

$$(\vec{a}, \vec{b}, \theta) \quad \text{– triple} ,$$

$$(x, 4, \vec{a}, \sin) \quad \text{– quadruple} .$$

A series of n quantities is called n-tuple.

The most familiar use of tuples is the representation of coordinates of points. In two dimensional plane:

$$(x, y) = (2, 5) \quad \text{– couple (pair)} .$$

[6]Collections, ensembles, groups, families – these terms already have reserved specific mathematical meaning, although they express similar idea.

In three dimensional space:

$$(x, y, z) = (0, 1, -2) \quad \text{– triple}\,.$$

Notice the difference between the coordinate tuples and the examples given above: The tuples with coordinates contain only numbers, whereas in the examples above, tuples contain quantities of various types. The point is that tuples *can* contain quantities of the same type, but in general they do not. Another important difference between sets and tuples is the importance of the order. Consider the set with two elements:

$$\mathbb{S} = \{0, 1\} = \{1, 0\}\,.$$

The order in which we write the elements does not matter. In contrast, the order is important for tuples:

$$(0, 1) \neq (1, 0)\,.$$

This inequality becomes obvious if we interpret these couples as coordinates of points in a plane. The first pair correspond to the point on the x axis, while the second pair corresponds to the point on the y axis.

Finally, the tuples can be used to concisely describe *mathematical structures*. As a simple example, consider the set of whole numbers $\mathbb{Z} = \{0,\ \pm1,\ \pm2,\ \ldots\}$. We can do many arithmetic operations with these numbers, but let us focus only on the operation of addition. Then we can summarize our interest using a triple as follows:

$$(\mathbb{Z}, +, 0)\,.$$

This expression says that we are studying a set of whole numbers $\mathbb{Z}$ equipped with a single operation "+". Moreover, we recognize that for this operation there is a special element 0 with "neutral" behavior:

$$n + 0 = 0 + n = n$$

for all numbers n from the set $\mathbb{Z}$. We will encounter more examples of this sort later in the book.

Chapter Highlights

- *Natural evolution of mathematical objects from numbers, through vectors, leads to tensors.*
- *Each successive tier of mathematical object in the progression "numbers, vectors, tensors" is more abstract and more powerfull.*
- *Numbers, vectors, and tensors are all conceptually connected.*
- *Just like the use of vectors opens up new methods for solving abstract and applied problems, so the use of tensors opens up new, even more powerful, methods for solving problems in various domains of science.*
- *Diagrams and schematics are helpful to illustrate various mathematical relations and structures.*
- *Set is an assembly of objects of similar kind. It is one of the basic concepts in mathematics.*
- *Tuple is an ordered sequence of elements related to each other by a common context. The elements of a tuple can be of different kind.*

2. Physics

Numbers are powerful mathematical objects. They are used to solve an endless list of problems that involve *quantities*. As mathematics and sciences progressed, natural numbers evolved into whole numbers, then into rational numbers and beyond.[1]

At a certain stage, problems of physics needed a mathematical tool to describe quantities with arbitrary direction. For example, a motion of a body involves velocity – a physical quantity describing how fast the body is moving and in what direction. Quite rapidly vectors led to tensors. Tensors had to be invented because there are many important problems where tensors are very natural. Examples will be given in the last chapter of the book.

On the surface, numbers, vectors, and tensors are rather different. However, they have a lot in common. Tensors are "numbers on steroids" in the sense that all things you can do with numbers, you can do with tensors, and even more.

Before we turn to tensors, we should familiarize ourselves with vectors. And before that, we must review the main concepts associated with numbers.

2.1 Power of Abstraction

Mathematics is a remarkably effective and universal discipline, its methods and results can be applied in a wide range of fields. In part, the universality of mathematics stems from the *general* and *abstract* nature of mathematical concepts. Let us illustrate this using an example.

---

[1] A superb account of this process is given in the book *"Number: The Language of Science"* by Tobias Dantzig.

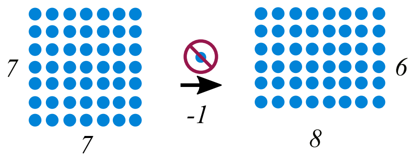


Fig. 2.1: 49 objects can be arranged in a square 7x7. 48 objects can be arranged as a rectangle of 6x8.

An astute farmer notices that 49 sacks of grains can be arranged in a square with each side having 7 sacks (see the Figure 2.1). When one sack is used up, the remaining 48 sacks can be arranged as a rectangle 6 by 8 sacks.

The farmer realizes that this curious fact has nothing to do with either grains or sacks. The same observation could be made about buckets, chairs, people, and so on. As the first step of generalization, the farmer states that 49 *objects* can be arranged as a 7 by 7 square, while 48 objects can be arranged as a 6 by 8 rectangle. The farmer also notes that $48 = 49 - 1$, whereas $6 = 7 - 1$ and $8 = 7 + 1$. She writes down the newly discovered relation as follows:

$$7 * 7 \text{ obj } - 1 \text{ obj } = (7 - 1) * (7 + 1) \text{ obj } ,$$

where *obj* is the denotation of *any* object.

As the grain is used up, the farmer discovers two more relations:

$$6 * 6 \text{ obj } - 1 \text{ obj } = (6 - 1) * (6 + 1) \text{ obj}$$

and

$$5 * 5 \text{ obj } - 1 \text{ obj } = (5 - 1) * (5 + 1) \text{ obj } .$$

At this point the farmer makes an educated guess, stating that a more general relation must exist:

$$n * n - 1 = (n - 1) * (n + 1) \,. \tag{2.1}$$

In the last expression the reference to objects is dropped, the expression is written simply in terms of *numbers*.

A deeper analysis reveals that the relation given by (2.1) exists for *any quantities* that obey usual rules of addition and multiplication. This includes rational numbers, real numbers, complex numbers (see Section

??), and even operators! The relation

$$x^2 - 1 = (x - 1)(x + 1) \tag{2.2}$$

holds true because of the way we define *rules for manipulation* – addition and multiplication in this case – of number-like objects, regardless of what those number-like objects represent in a particular problem.

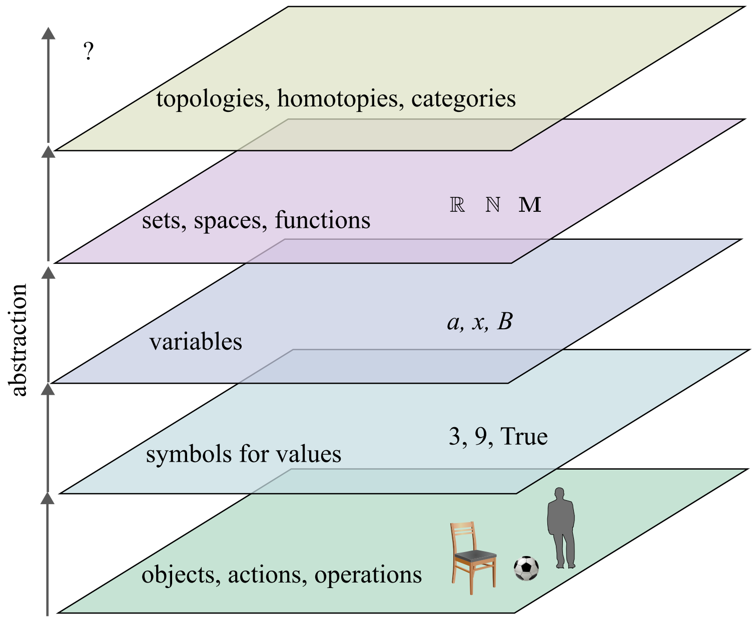


Fig. 2.2: Mathematical thinking has many levels of abstraction. Going to higher levels of abstractions results in higher efficiency and more powerful ideas and tools.

The path from "sacks of grain" to a variable x is the path from concrete, specific objects to *abstract* entities that are the product of creative imagination. This path to higher levels of abstraction is illustrated in the Figure 2.2. As we move to higher levels of abstraction, our mathematical tools become more powerful and more universally applicable: From everyday arithmetic, to economics, to general relativity, and quantum gravity.

Using more abstract mathematical objects requires serious mental effort. To reach the highest levels one needs to do mathematics professionally. However, any profession can benefit from *some* level of abstraction, and to understand vectors and tensors, we must go beyond usual high-school level. One of the goals of this book is to help readers build tolerance and appreciation of more abstract aspects of mathematics.

2.2 Terminology Barrier

Every high school student has a working knowledge of bilinear operators over associative and commutative fields, but hardly anyone of those students is aware of this fact. We refer here to the ability to simply add and multiply numbers. This demonstrates that even familiar and basic notions may look complicated when "dressed in unfamiliar clothing."

When we learn new mathematical concepts, especially at a higher level of abstraction, we often encounter what might be called a *terminology barrier*: a concept seems more difficult if it is formulated in a new language, without sufficient connections to already familiar concepts, and without clear examples of how the concept can be applied.

New terminology is unavoidable when learning new concepts. There will be a number of new mathematical concepts and definition introduced in this book. To lower the terminology barrier, every new mathematical concept will be illustrated with examples and the connections to already familiar concepts will also be given. Additionally, it is recommended to do the following exercise every time a new concept with unusual terminology is introduced:

Dealing With New Concepts

- Take a critical look at a new name and notation.
- Think whether the new name or notation looks like something you know. Is the resemblance helpful or misleading?
- Be creative and try to come up with your own notation or word to describe the new concept.

Remember: Symbols and names are not essential. What is important is the set of *relations* of a new concept to other concepts. The relations show how the concept fits and functions within the larger framework.

Demonstrations of this approach can be found in the rest of the book.

2.3 Algebra of Numbers

We will start with the “familiar” numbers:

$$0, \pm 1, \pm 2, \ldots, \pm n, \ldots$$

This endless collection, considered as one entity, is called the *set of whole numbers*; it is denoted as $\mathbb{Z}$.

The set $\mathbb{Z}$ is not a formless heap of elements. On the contrary, it has a rich *structure*. The structure of any set, including the set of whole numbers, is revealed through various *relations* between all or some of its elements. Here are several examples:

- 1 and -1 are related, and so are 2 and -2, and generally, n and $-n$.
- Relation exists between 1 and 2, 2 and 4, 3 and 6, and generally, between n and $2n$.
- The pairs $(1,2)$, $(2,3)$, $(3,4)$, and so on, illustrate an important relation of order that exists in $\mathbb{Z}$.
- The pairs $(2,1)$, $(3,1)$, $(4,2)$, $(5,1)$, $(6,3)$ and similar ones unite a number and its greatest divisor not equal to the number itself.

The list can be continued indefinitely, but the general idea of relations should be clear. Such number-to-number relation can be schematically represented as a box with an input and an output, as shown in the Figure 2.3. A few important relations are given descriptive names: **neg**, **dbl**, **suc**, and **gsd** are examples from the Figure 2.3; they correspond to negation, doubling of a number, finding the successor of a given number, and finding the greatest divisor smaller than the number itself.

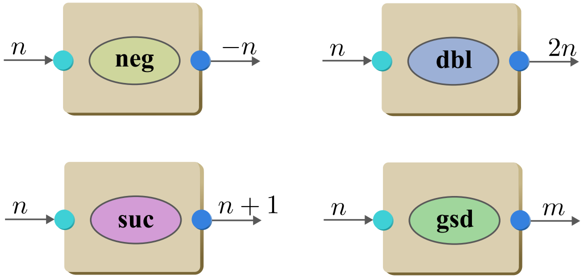
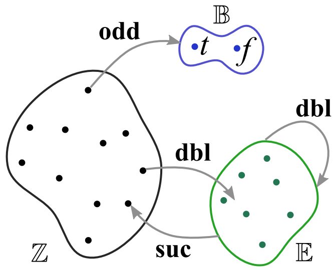


Fig. 2.3: Relation between elements of a set can be schematically represented using boxes with inputs and outputs. Here the relations between numbers are given descriptive names: **neg** is negation, **dbl** is doubling, **suc** is getting the successive number, **gsd** is the greatest divisor of a number smaller than the number itself.

Fig. 2.4: Relations can be viewed on the level of sets. A function maps (connects) one set with another in a meaningful way. For example, **dbl** maps every integer from $\mathbb{Z}$ into the set of even numbers $\mathbb{E}$.



2.3.1 Functions

What we have just described is the idea of a *function*, or numeric function of a single *argument*, to be precise. A function of a single argument connects every *argument* (input) to a certain *value* (output), establishing a *relation* between a pair of elements.

Another view on relations is illustrated in the Figure 2.4. Relations between elements can be "elevated" to the level of sets and depicted as arrows connecting one set (*domain*) to another set (*range*). Symbolically we can write:

$$\mathbb{Z} \xrightarrow{\textbf{dbl}} \mathbb{E}\,,$$

where $\mathbb{E}$ denotes the set of all even numbers, **dbl** is the name of the function that doubles its argument.

The relations of the type "one number to one number" – considered above – can be generalized to "several numbers to one number" or "one number to several numbers" or even "several numbers to several numbers." The schematic representation of such relations are given in the Figure 2.5.

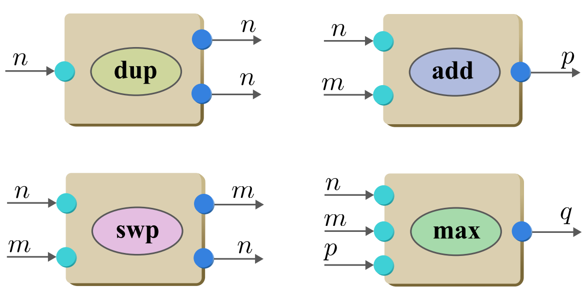


Fig. 2.5: Relations of several elements to several elements: **dup** duplicates its input, **add** calculates the sum, **swp** swaps the order of the arguments, **max** returns the maximum of three input numbers.

Exercise 2.1
Think how you would represent the generalized relations of the types given in the Figure 2.5 at the level of sets? What kind of diagrams would you draw?

Binary Functions

An important function of two variables is the familiar **add** relation:

$$\mathbf{add}\, n\, m = n + m\,.$$

The right-hand side of this equality is just another way of writing the expression involving function with two input arguments. It is a special case of a more general rule, which can be written as follows:

$$f\ n\, m = n \circledcirc m\,.$$

On the left we have a *prefix* notation, where a function f is *applied* to two arguments. On the right, we have an *infix* notation and a special symbol placed between the first and the second argument. Several familiar examples are:

- **mul** $n\, m = n * m$ – multiplication.
- **pow** $n\, m = n{}^\wedge m$ – power.
- **sub** $n\, m = n - m$ – subtraction.
- **div** $n\, m = n/m$ – division.

The functions **mul**, **pow**, **sub**, and **div** are all examples of *binary* functions – functions of two arguments.

Exercise 2.2
Think of a your own example of a binary function (function of two arguments). Create its infix variant.

Exercise 2.3
Show that

$$\epsilon_{ij} a_i a_j = 0\,.$$

Chapter Highlights

- *The power of mathematical concepts and methods increases with the level of abstraction.*
- *Learning new concepts often involves learning new terminology. The latter can create an artificial mental barrier.*
- *"Usual" numbers form a mathematical structure. The structure is revealed through various relations that exist between numbers.*
- *Relations between numbers are expressed using the concept of functions and operations (e.g., addition). Each operation is characterized by its arity – the number of arguments it accepts as an input.*
- *Functions can be represented schematically as boxes with inputs and outputs.*
- *Functions that act on natural numbers can be written using index notation (e.g.,* $f\ i = f_i$*).*
- *Linear functions represent the simplest but still powerful and useful kind of functions.*
- *Functions can be composed to create new functions.*
- *A function with several inputs is said to be partially applied when not all its inputs are populated.*
- *The same function can be represented in various ways: Graphical, as a symbolic formula, as a table. The function is not reduced to any of its representations.*
- *The power of abstract mathematical thinking comes, in part, from efficient notation. Einstein's Summation Rule (ESR) is a good example of this.*

3. Mathematics

In the previous chapter we learned about numbers and various relations between them. As a particular class of relations we discussed functions. We introduced *binary* and *unary* functions and different ways functions can be combined (*composed*) to produce new functions. We also learned that functions can be represented in various ways and that none of those different representations defines the concept of function completely. Each representation of a function highlighted some important aspect of it.

Vectors, which will be introduced in this chapter, also allow different representations. We will start with a particular model of vector quantities – *arrows*. It is important to remember that while this model illustrates the concept of vectors, it does not define vectors completely. In other words, arrows are particular examples of vectors, but vectors are more that just directed line segments.

To arrive at the definition of vectors we must explore their properties more fully. This will be the goal of current chapter.

3.1 Arrows

To arrive at the idea of vectors we will start with simple geometrical objects – arrows in a plane, as illustrated in the Figure 3.1.

Symbolically, we will denote vectors by placing an arrow over letters:

$$\vec{a}\,,\vec{b}\,,\vec{c}\,,\ldots\,,\vec{\alpha}\,,\vec{\beta}\,.$$

The length of an arrow $\vec{a}$ is denoted by the same letter without an arrow:

$$\textbf{length}\,\vec{a} = a\,.$$

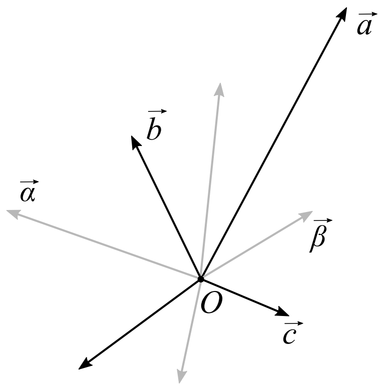


Fig. 3.1: Set of arrows starting at the same origin point O. All imaginable arrows taken as one set form the arrow space $\overset{\Rightarrow}{A}$.

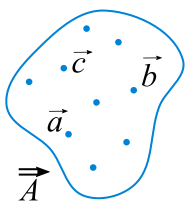


Fig. 3.2: All arrows, considered as unified collection of objects of similar nature, can be viewed as a mathematical set. Each arrow is an element of this set, denoted as a point.

The set of all possible arrows – called *arrow space* (or *vector space*) – we will denote as

$$\overset{\Rightarrow}{A} = \{\vec{a}\,, \vec{b}\,, \vec{c}\,, \ldots\,, \vec{\alpha}\,, \vec{\beta}\, \ldots\}\,.$$

Diagrammatic representation of the set of arrows is given in the Figure 3.2.

Functions on Arrows

The function **length** is a unary function that accepts arrows as input (argument) and returns a numeric value – the lengths of the arrow.

On the level of sets, the **length** function *maps* the set of all vectors into the set of real numbers:

$$\overset{\Rightarrow}{A} \xrightarrow{\textbf{length}} \mathbb{R}\,.$$

In the next chapter we will encounter other types of functions on arrows. Some of them will return numbers, like **length** does, and

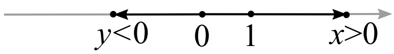


Fig. 3.3: Real numbers can be represented by arrows oriented along a fixed line – *number line.*

others will return arrows, mapping *arrow space* into its own copy:

$$\overset{\Rightarrow}{A} \longrightarrow \overset{\Rightarrow}{A}\,.$$

Arrows and Numbers

Arrows in a plane include, in some sense, the notion of numbers. Indeed, as shown in the Figure 3.3, real numbers may be represented as points on a line – number line. Positive numbers correspond to arrows pointing to the right of the origin (number zero), negative numbers correspond to the arrows directed to the left.

3.2 Algebra of Arrows

Next, we will establish similarities between arrows and numbers by exploring possible *algebraic operations* on vectors.

3.2.1 Combining Arrows

Two arrows can be *combined* in a natural way to yield the third arrow. Using either the head-to-tail approach or a parallelogram method, the vectors $\vec{a}$ and $\vec{b}$ can be combined graphically, as illustrated in Figure 3.4.

Chapter Highlights

- *Arrows in a plane provide a simple model for vectors.*
- *Arrows can be manipulated in ways analogous to numbers: Two arrows be added, an arrow can be "scaled" (stretched or compressed). Arrows form an algebra.*
- *Basis is an extremely important concept. Basis is a set of objects (arrows) that can be used to "build" all other similar objects (arrows). At the same*

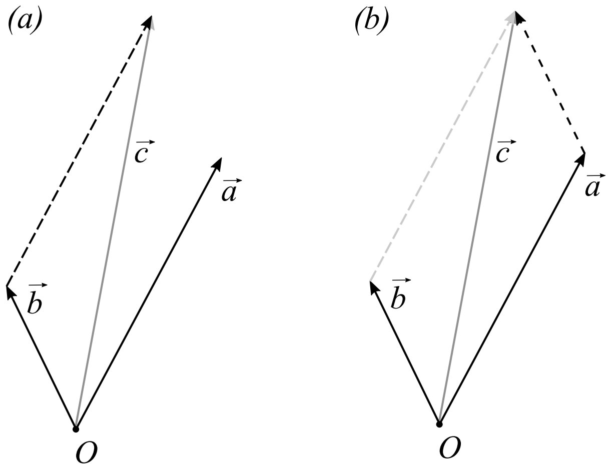


Fig. 3.4: Two arrows can be combined to produce new arrow. One of the simplest way to do this is to arrange two arrows "tail-to-tip". This operation is called *addition of arrows*: $\vec{a} + \vec{b} = \vec{c}$.

time, basis can not be used to build itself – basis arrows are independent.

- *Basis can be chosen in infinite number of ways. There is no special basis. Different bases might be useful for different problems.*
- *Given a basis, arrows can be specified by writing their components (using index notation) relative to the basis.*
- *Einstein's summation rule is very useful for manipulating expressions involving components of arrows.*
- *The exact values of the arrow's components depend on the basis. Changing the basis changes the values of components, while the arrow remains the same. This is one of defining properties of vectors.*
- *When basis is changed, components of the same arrow transform in a very specific way. Depending on exact form of transformation, we can speak of two kind of vectors: contravariant and covariant. Arrow-like vectors are example of contravariant vectors.*

4. Classical Physics

The concept of *operators* extends the idea of functions. An unary numeric function f takes some numeric value x as an input and produces another numeric value y:

$$f\,x = y \qquad \text{or} \qquad x \xrightarrow{f} y\,.$$

In mathematical jargon, f *maps* x into y.

Similarly, we can study functions that *operate* on arrows/vectors, yielding other arrows/vectors. The parallel between unary functions over numbers and unary operators over vectors is highlighted in the Figure 4.1.

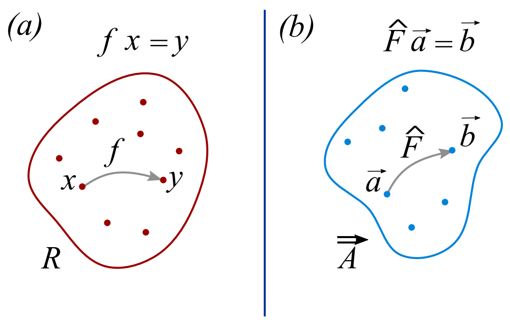


Fig. 4.1: Operators extend the idea of functions. (a) An unary function f can be applied to a number x to produce another number y. (b) An unary operator $\widehat{F}$ can be applied to a vector $\vec{a}$ to yield another vector $\vec{b}$.

4.1 Operators on Arrows

An action of an operator F on arrows can be represented symbolically as an equation:

$$F\vec{a} = \vec{b}\,.$$

Often a "hat" is placed on top of an operator[1], to emphasize that it is different from numeric function:

$$\boxed{\widehat{F}\vec{a} = \vec{b}\,.}$$

Simple Operators

It is easy to come up with examples of operators:

- Unit operator (or *identity* operator), such that

$$\widehat{I}\vec{a} = \vec{a}\,.$$

- "Zeroing" operator that maps every vector into a zero vector:

$$\widehat{0}\vec{a} = \vec{0}\,.$$

- Scaling operator, such that

$$\widehat{S}_5\,\vec{a} = 5\vec{a}\,.$$

- Rotation operator, such that

$$\widehat{L}\vec{a} = \vec{b}\,,$$

where $\vec{b}$ is rotated by $45°$ counter-clockwise relative to the vector $\vec{a}$.

To fully describe an operator, we must describe how it acts *on every* arrow. In general, this requires an *infinite* amount of information, since there are infinite number of arrows and the action of F on different arrows might be different[2]. Describing the action of a general operator

[1] In Quantum Mechanics, for example.

[2] Simple operators given before are special cases when it is easy to describe action of

in graphical terms using arrows is a hopeless task. This is the case where component representation of arrows saves the day.

If we fix a basis $\{\vec{e}_i\}$, then every vector gets an algebraic representation as a set of components:

$$\vec{a} \quad \xrightarrow{\{\vec{e}_k\}} \quad a_i\,,$$

$$\vec{b} \quad \xrightarrow{\{\vec{e}_k\}} \quad b_j\,.$$

Now to describe the action of an operator $\widehat{F}$ on the vector $\vec{a}$ we can specify the components of the result b_j for arbitrary components of the input a_i. For vectors in a plane, we have

$$b_1 = F_1\, a_1\, a_2$$

and

$$b_2 = F_2\, a_1\, a_2\,.$$

Thus, a pair of binary numeric functions F_1 and F_2 is sufficient to describe an operator. A situation significantly simplifies in the case of very important *linear operators* which we will discuss soon in Section .

Examples

Let us take a closer look at a couple of operators. While studying these examples we must keep in mind that the relations between components are *specific to basis* and will change if we change the basis. The question of how exactly the relation between components changes will be addressed later in Section 4.7 for the simplest types of operators.

The first operator acts on the components as follows:

$$b_1 = a_1 + a_2\,,$$

$$b_2 = a_1 * a_2\,.$$

To illustrate these relations visually, we can start with arrows of equal lengths but pointing in all directions. The tips of such arrows will lie on a circle, as shown in the Figure 4.2 using blue dotted lines. The tips of all output arrows $\vec{b} = \widehat{F}\,\vec{a}$ lie on a curve shown in the Figure 4.2 using solid red line. In the first example the arrows tracing the circle of the

operators on *all* vectors.

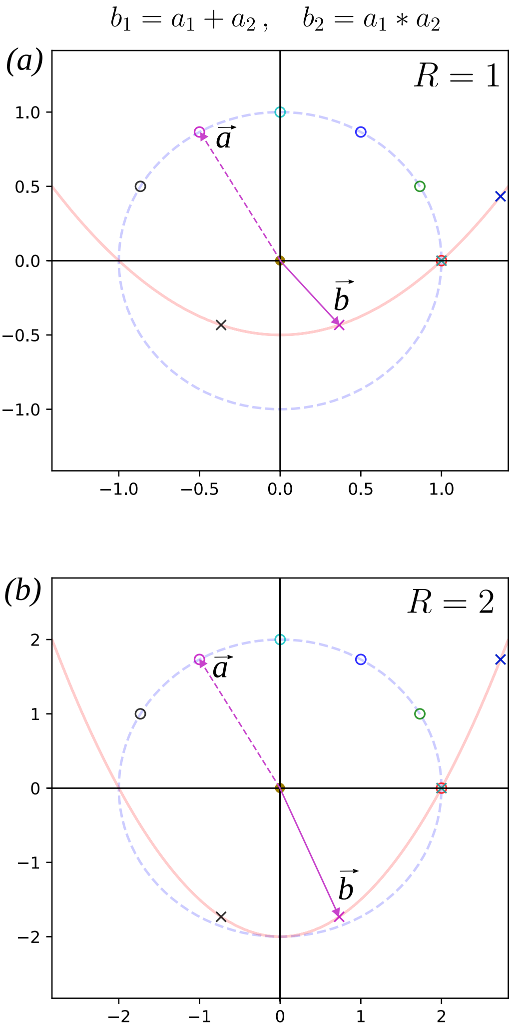


Fig. 4.2: The action of the operator $\widehat{F}$ on planar arrow-vectors $\vec{a}$. To demonstrate how $\widehat{F}$ acts on arrows of different directions and lengths, we consider what happens to the circles of two different radii.

radius $R = 1$ are transformed into arrows tracing a curve that looks like parabola.

Exercise 4.1
Show that when the components of the output arrow $\vec{b}$ are given by

$$b_1 = a_1 + a_2\,,$$

$$b_2 = a_1 * a_2\,,$$

then the circe with the radius R becomes a parabola described by the equation

$$b_2 = b_1^2/2 - R^2/2\,.$$

In the second example the operator $\widehat{G}$ acts on the components as follows:

$$b_1 = a_1 - a_2\,,$$

$$b_2 = -a_1^2 * a_2\,.$$

The effect of this operator on the arrows forming a circle is illustrated in the Figure 4.3. It seems that increasing the radius of the circles does not substantially change its "image" (solid red line) and only "stretches" the output curve both horizontally and vertically.

Nonlinear and Linear Operators
The operators $\widehat{F}$ and $\widehat{G}$ are examples of rather complicated operators. They are *nonlinear* because they lack the simple property of *linearity* (we will learn about it soon in Section 4.2.)

Linear operators connect components of input and output vectors in a simple way:

$$b_1 = A\,a_1 + B\,a_2\,, \quad b_2 = C\,a_1 + D\,a_2\,.$$

Here A, B, C, D are numbers. Different linear operators differ only in the values of these numbers. Despite their simplicity, linear operators are powerful and widely used.

Since two arrows can differ only in their magnitude (length) and directions, the action of an operator can be represented by a combination

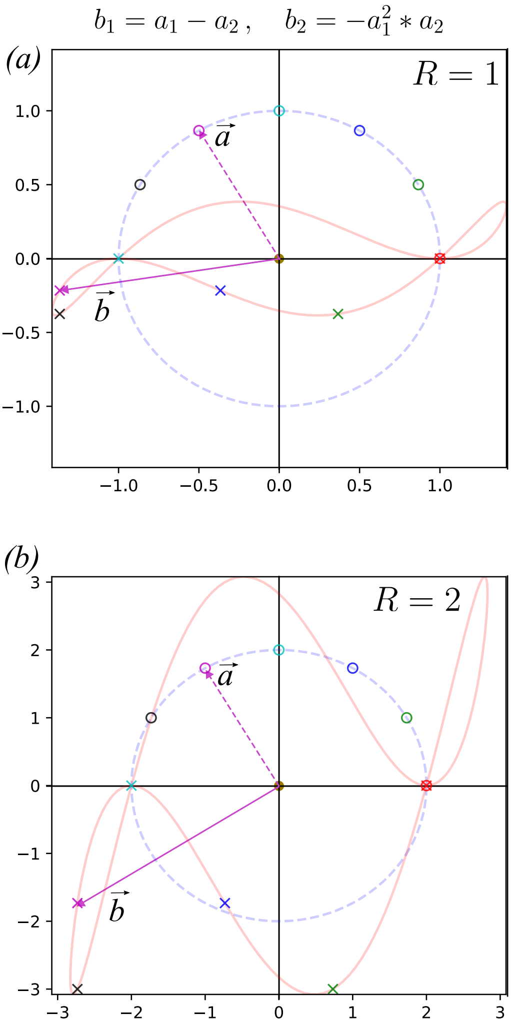


Fig. 4.3: An action of an operator $\widehat{G}$ on a planar vector $\vec{a}$. To demonstrate how $\widehat{G}$ acts on arrows of different directions and lengths, we consider what happens to the circles of two different radii.

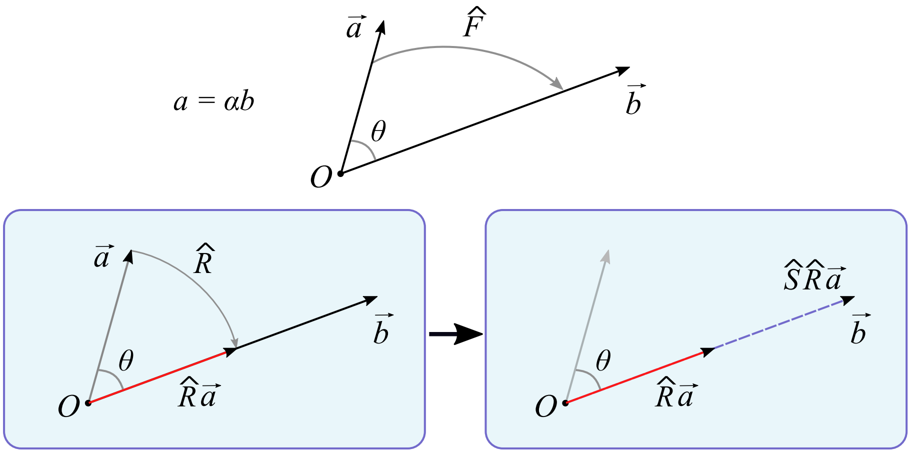


Fig. 4.4: An action of an operator $\widehat{F}$ on a planar vector $\vec{a}$ can be described by the sequence of rotation $\widehat{R}$ and scaling $\widehat{S}$: $\widehat{F} = \widehat{S} \circ \widehat{R}$.

of two steps: rotation and scaling; see the Figure 4.4. The order of scaling and rotation does not matter:

$$\widehat{F} = \widehat{S} \circ \widehat{R} = \widehat{R} \circ \widehat{S}\,.$$

As our next step, we will study rotation operators in more details.

4.1.1 Rotation Operator

A simple non-trivial[3] operator rotates a vector by some angle, as demonstrated in the Figure 4.5. Strictly speaking, there are infinite number of such operators, one for each rotational angle θ.

An important property of any rotation operator is that it preserves some relations between arrows. For example, given a vector

$$\vec{c} = \vec{a} + \vec{b}\,,$$

its "image" $(\widehat{R}\,\vec{c})$ can be constructed from rotated vectors $(\widehat{R}\,\vec{a})$ and

[3]An example of a trivial operator is the identity operator $\widehat{I}$ which does not change the input vector: $\widehat{I}\vec{a} = \vec{a}$. Another example is the operator that always returns "zero"-arrow: $\widehat{0}\,\vec{a} = \vec{0}$.

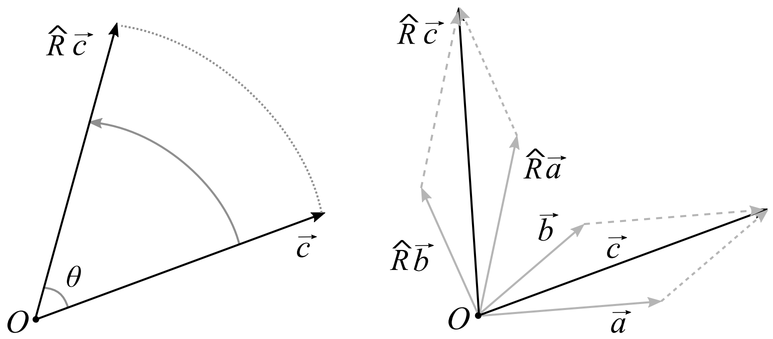


Fig. 4.5: The action of the rotation operator $\widehat{R}$ on a planar vector $\vec{a}$ results in the rotation by a specified angle θ. Such rotation posseses an important property: $\widehat{R}\,(\vec{a} + \vec{b}) = \widehat{R}\,\vec{a} + \widehat{R}\,\vec{b}$.

$(\widehat{R}\,\vec{b})$:

$$\widehat{R}\,\vec{c} = \widehat{R}\,(\vec{a} + \vec{b}) = (\widehat{R}\,\vec{a}) + (\widehat{R}\,\vec{b}).$$

This property is the consequence of two features of the rotation operator: 1) it does not change the length of a vector; 2) it rotates every vector by the same amount, thus keeping the relative angle between any two vectors intact, as shown in the see Figure 4.5.

Another important property of rotation operators is an obvious one:

$$\widehat{R}\,(\alpha\vec{a}) = \alpha(\widehat{R}\,\vec{a})\,.$$

Rotation operators represent a subset of a larger set of important operators – *linear operators.*

4.2 Linear Operators

Linear operators are operators with a simple but important property of *linearity*. Two things are required for a linear operator $\widehat{L}$:

$$\boxed{\widehat{L}\,(\vec{a} + \vec{b}) = (\widehat{L}\,\vec{a}) + (\widehat{L}\,\vec{b})}$$

and

$$\boxed{\widehat{L}\,(\alpha\vec{a}) = \alpha(\widehat{L}\,\vec{a})\,.}$$

Rotation operators satisfy these requirements, as we established in the previous section.

To represent linear operators numerically using basis arrows requires very little information. Indeed, for a general arrow $\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2$ the action of the operator $\widehat{L}$ can be written as

$$\widehat{L}\,(a_1\vec{e}_1 + a_2\vec{e}_2) = [\widehat{L}\,(a_1\vec{e}_1)] + [\widehat{L}\,(a_2\vec{e}_2)] = a_1(\widehat{L}\,\vec{e}_1) + a_2(\widehat{L}\,\vec{e}_2)\,.$$

Thus, to define the action of the linear operator $\widehat{L}$ on an arbitrary arrow $\vec{a}$, it is sufficient to define its action on the basis arrows $\vec{e}_1$ and $\vec{e}_2$.

Operator $\widehat{L}$ acts on arrows to yield other arrows. Thus, we can write

$$\widehat{L}\,\vec{e}_1 = \vec{f}_1 \text{ and } \widehat{L}\,\vec{e}_2 = \vec{f}_2\,.$$

Like any other vectors, the vectors $\vec{f}_1$ and $\vec{f}_2$ can be written in terms of the basis vectors:

$$\vec{f}_1 = l_1\vec{e}_1 + l_2\vec{e}_2\,,$$
$$\vec{f}_2 = l_3\vec{e}_1 + l_4\vec{e}_2\,.$$

These equations can be written more compactly if we improve the notation. Firstly, instead of numbers $l_1\,, l_2\,, l_3\,, l_4$ we will write

$$\widehat{L}\,\vec{e}_1 = L_{11}\vec{e}_1 + L_{12}\vec{e}_2$$

and

$$\widehat{L}\,\vec{e}_2 = L_{21}\vec{e}_1 + L_{22}\vec{e}_2\,.$$

Notice how the subscript indices match nicely with the indices of the basis arrows: The first subscript index of L_{ij} corresponds to the basis vector being acted on, while the second index corresponds to the basis vectors being multiplied by L_{ij}. The second step is to use the summation agreement:

$$\widehat{L}\,\vec{e}_1 = L_{1j}\vec{e}_j,$$
$$\widehat{L}\,\vec{e}_2 = L_{2j}\vec{e}_j.$$

Finally, we can write the most compact form:

$$\boxed{\widehat{L}\,\vec{e}_i = L_{ij}\vec{e}_j.} \tag{4.1}$$

In summary, the action of a linear operator $\widehat{L}$ on the basis arrows (and, consequently, on *any* arrow) is completely determined by its components L_{ij}– just four numbers for arrows in a plane[4].

The components of a linear operator $\widehat{L}$ are specific to the basis. This is completely analogous to the components of arrow-vectors. Indeed, the use of a basis translates arrows and linear operators from the graphical world of drawings into the algebraic world of numbers:

$$\vec{a} \xrightarrow{\{\vec{e}_1,\vec{e}_2\}} a_i\,,$$

$$\widehat{L} \xrightarrow{\{\vec{e}_1,\vec{e}_2\}} L_{ij}\,.$$

Simple Linear Operators

Four simple types of linear operators can be defined without specifying their components:

- Unit operator, such that $\widehat{I}\,\vec{a} = \vec{a}$.
- "Zeroing" operator that maps every vector into a zero vector:

$$\widehat{0}\,\vec{a} = \vec{0}\,.$$

- Scaling operators, such that $\widehat{S}\,\vec{a} = \alpha\vec{a}$ for some specified value α.
- Rotation operators, such that $\widehat{R}_\theta\,\vec{a} = \vec{b}$, where $\vec{b}$ is simply rotated by specified angle θ.

To find the components of any operator we must see how it acts on basis vectors. Let's see how this works for the scaling operator $\widehat{S}$ described above:

$$\widehat{S}\,\vec{e}_1 = \alpha\vec{e}_1 + 0\vec{e}_2\,,$$

$$\widehat{S}\,\vec{e}_2 = 0\vec{e}_1 + \alpha\vec{e}_2\,.$$

From these equations we can read the components S_{ij} and write them in matrix form:

$$\widehat{S} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}.$$

[4] In a space of N dimensions the number of components is, in general, equal to N^2.

Similar approach can be used to find the components of any linear operator.

Exercise 4.2

Consider an operator $\widehat{N}$ which "normalizes" an arrow – returns an arrow of unit length pointing in the same direction as the original one. For example:

$$\widehat{N}\,\vec{a} = \vec{u}_a = \frac{\vec{a}}{a}\,.$$

Is it a linear operator?

$\widehat{J}$-operator

Operator that rotates any vector by 90 degrees has a special notation (it will be heavily used in Sections **??** and **??**):

$$\widehat{R}_{\pi/2} = \widehat{J}\,.$$

It is instructive to see how this operator acts on orthonormal basis $\{\vec{e}_i\}$:

$$\widehat{J}\,\vec{e}_1 = \vec{e}_2 = 0\,\vec{e}_1 + 1\,\vec{e}_2$$

and

$$\widehat{J}\,\vec{e}_2 = -\vec{e}_1 = -1\,\vec{e}_1 + 0\,\vec{e}_2\,.$$

From the last two expressions we can find the components of the $\widehat{J}$-operator:

$$J_{ij} = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\,.$$

Here we used *matrix* form of presenting the components of linear operators. In this form, the first index of the components corresponds to the row in the matrix-table, and the second index corresponds to the column.

Another important property of this operator can be seen when we act on any vector twice:

$$\widehat{J}\,(\widehat{J}\,\vec{a}) = -\vec{a}\,.$$

Indeed, rotating any vector twice by 90 degrees results in total rotation by 180 degrees – the direction of the original vector is flipped. What we showed is that the sequence $\widehat{J} \circ \widehat{J}$ is the same as the operator $(-\widehat{I})$:

$$\boxed{\widehat{J} \circ \widehat{J} = -\widehat{I}\,.}$$

4.3 Plotting Linear Operators

A numeric unary function

$$f\,x = y$$

can be represented graphically as a plot "$y\,vs\,x$". In this plot we indicate a value y *for each value of* x.

Linear operators of the type

$$\widehat{L}\,\vec{a} = \vec{b}$$

also allow graphical representation. We can, for example, draw an arrow $\vec{b}$ *for each value of* $\vec{a}$. The input vector $\vec{a}$ can vary both its lengths and direction. For a linear operator $\widehat{L}$ the change in length of the input vector $\vec{a}$ is handled trivially:

$$\widehat{L}\,(\alpha\vec{a}) = \alpha(\widehat{L}\,\vec{a})\,.$$

In words: To find the action of a linear operator on a scaled vector we first apply the operator to the original vector and then scale the result. It also follows that all vectors $\widehat{L}\,(\alpha\vec{a})$ corresponding to different values of the scale factor α are *parallel* – they are all parallel to the vector $\vec{b} = \widehat{L}\,\vec{a}$.

The preceding considerations show that to describe the action of a linear operator on various input vectors we can focus only on vectors with unit lengths, but pointing in all possible directions. The tips of all such vectors form a unit circle, as illustrated in the Figure 4.6.

For linear operators it is sufficient to plot only the half of the circle, since

$$\widehat{L}\,(-\vec{a}) = -(\widehat{L}\,\vec{a})\,.$$

In other words, the missing half is the inverted copy of the original half. An illustration of this is given in the Figure 4.7

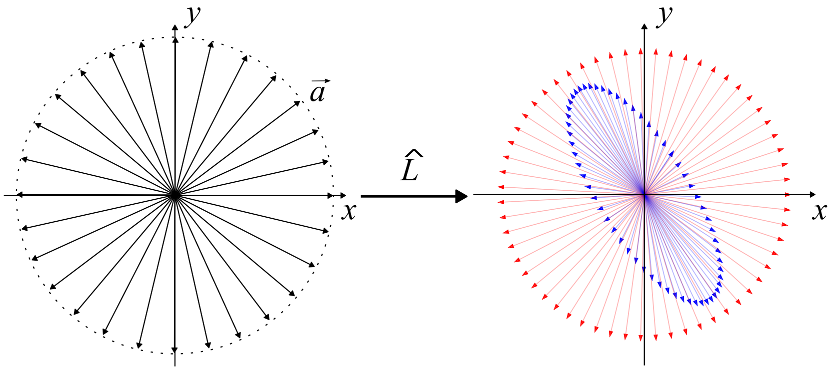


Fig. 4.6: Input vectors for a linear operator can vary in direction. In this example the components of the operator are: $L_{11} = 2/10$, $L_{12} = 3/10$, $L_{21} = 5/10$, $L_{22} = -7/10$.

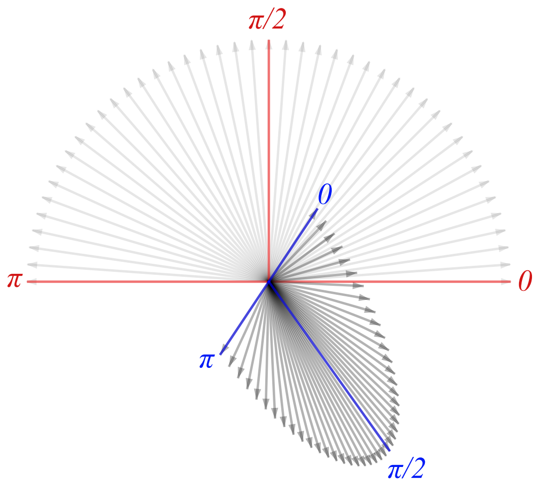


Fig. 4.7: To graphically describe a linear operator we can specify how it acts on vectors from the top half of the circle (the bottom is found by inverting the top). In this example the components of the operator are: $L_{11} = 2/10$, $L_{12} = 3/10$, $L_{21} = 5/10$, $L_{22} = -7/10$.

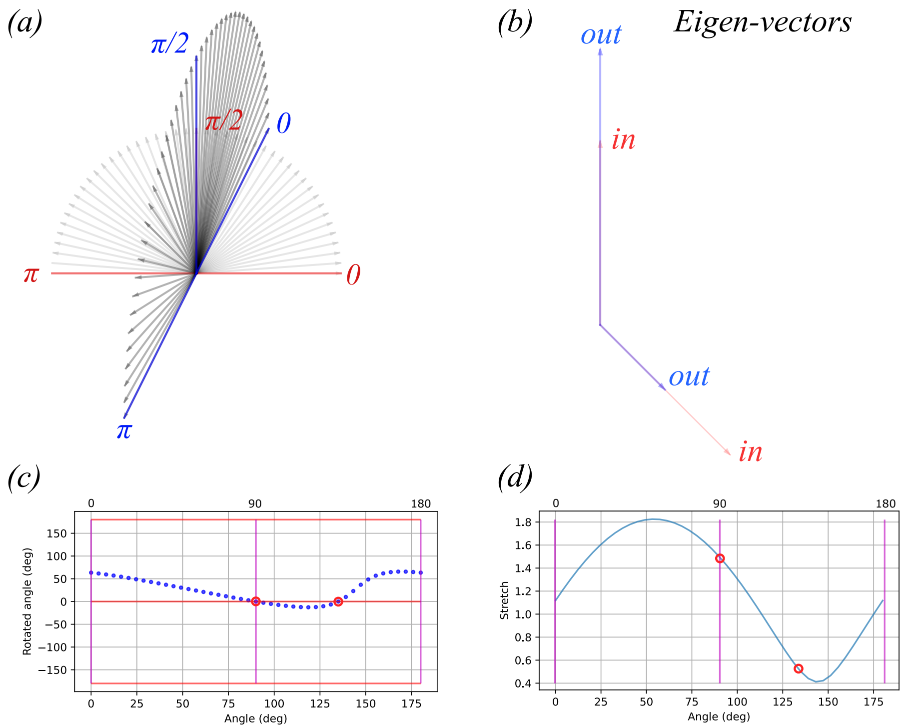


Fig. 4.8: Graphical representation of a linear operator $\widehat{L}$ with components: $L_{11} = 5/10\,, L_{12} = 2\,, L_{21} = 0\,, L_{22} = 3/2\,.$ (a) The effect of the operator on unit-length arrows with direction from $0°$ to $180°$; (b) Two special arrow-vectors that do not change their direction under the action of the operator $\widehat{L}$. The operator simply scales the arrows. Such vectors are called *eigen-vectors* of the operator; (c) Horizontal axis: initial orientation of the unit-length arrow; vertical axis: rotation angle of each unit-length arrow. Two special angles are highlighted. For each special angle the arrow is not rotated. Such arrows are called eigen-vectors of the operator $\widehat{L}$; (d) Horizontal axis: initial orientation of the unit-length arrow; vertical axis: the scaling factor of each unit-length arrow.

Another linear operator, with components $L_{11} = 5/10\,, L_{12} = 2\,, L_{21} = 0\,, L_{22} = 3/2$, is graphically represented in the Figure 4.8. In the Figure 4.8(a) the action of the operator on unit vectors from the top half-plane is shown, with the missing part (bottom half-plane) is easily constructed by inverting the transformed part through the origin. Although not easily seen, there are two pairs of vectors[5] that are transformed rather simply by this linear operator – they are scaled without rotation. Two such vectors are shown in the Figure 4.8(b). Such vectors are called *eigen-vectors* of a given operator. Eigen-vectors are discussed in more details in the next section.

Another way to graphically represent a linear operator is to plot two functions: 1) How much an input unit vectors gets rotated; 2) How much an input unit vector gets stretched. This is done in the Figure 4.8(c) and (d). From the Figure 4.8(c) it can be seen that there are two unit vectors that are not rotated by the operator (rotation angle is zero for the input vectors at 90 degrees and at about 130 degrees or 310=130+180 degrees).

Eigen-vectors are important and finding them for a given linear operator is an often encountered problem. It is called *eigen-vector problem* or *eigen-problem* for short.

4.4 Eigen-Problem*

Eigen-vectors of a linear operator are special vectors that are not rotated by the operator, eigen-vectors can only be scaled. This requirement can be expressed using a simple equation:

$$\boxed{\widehat{L}\,\vec{a} = \alpha\vec{a}\,,}$$

or, if there exists another eigen-vector $\vec{b}$ different from $\vec{a}$:

$$\widehat{L}\,\vec{b} = \beta\vec{b}\,.$$

Here the numbers α and β specify the scaling coefficients. The equations given above are called *eigen-problem equations*. The vectors $\vec{a}$ and $\vec{b}$ are called *eigen-vector*, the scaling factors α and β are called *eigen-values*. In general, $\alpha \neq \beta$, as illustrated in the Figure 4.8(b).

Not all linear operators have eigen-vectors. A simple example is the

[5] Two vectors and their reversed images.

operator of rotation by a finite angle:

$$\widehat{R}_\theta \vec{a} \neq \alpha \vec{a}\,.$$

Simply speaking, it is impossible to rotate a vector without changing its direction.

All "well-behaved" linear operators[6] (excluding rotation operators like $\widehat{R}_\theta$) have the same number of eigen-vectors as the number of basis vectors. Any "well-behaved" linear operator that operates on vectors in a plane has two eigen-vectors, as illustrated in the Figure 4.8(b). Operators acting on vectors in three dimensions may have up to three eigen-vectors.

It is possible to find linear operators that have fewer eigen-vectors than the number of basis vectors. Such operators are called *degenerate operators*. They are important and we will consider them next for the special case of two dimensions.

4.5 Degenerate Linear Operators

In some special cases, a linear operator $\widehat{L}$, when applied to the basis arrows $\vec{e}_1$ and $\vec{e}_2$ can produce parallel arrows:

$$(\widehat{L}\,\vec{e}_1) \parallel (\widehat{L}\,\vec{e}_2)$$

or

$$\widehat{L}\,\vec{e}_1 = \lambda(\widehat{L}\,\vec{e}_2)$$

for some number λ.

Using components notation, this condition can be written as follows:

$$L_{11}\vec{e}_1 + L_{12}\vec{e}_2 = \lambda L_{21}\vec{e}_1 + \lambda L_{22}\vec{e}_2\,.$$

The vector on the left-hand side is the same vector as on the right-hand side if they have the same components in the given basis. Therefore, we must equate corresponding components:

$$L_{11} = \lambda L_{21}$$

and

$$L_{12} = \lambda L_{22}\,.$$

[6] We will encounter ill-behaved linear operators in the next section.

Cross-multiplying these equations, we obtain

$$\lambda L_{11} L_{22} = \lambda L_{12} L_{21} \,,$$

from which follows

$$L_{12} L_{21} - L_{11} L_{22} = 0 \,.$$

Linear operator satisfying this condition is called *degenerate* for the reasons explained below.

Exercise 4.3
Show that a degenerate linear operator $\widehat{L}$ "collapses" all vectors onto the same line, i.e. all $(\widehat{L}\,\vec{a})$ have the same direction.

Later in Section 5.7 we will encounter a whole class of useful degenerate linear operators, called *projectors*, whose job is to project any vector onto a specified direction.

4.5.1 Determinant of a Linear Operator

The quantity

$$\boxed{D = L_{12} L_{21} - L_{11} L_{22}}$$

computed from the components of a linear operator is called its *determinant*. Determinant is one of important characteristics of a linear operator, analogous to how the length of a vector is its important characteristic.

Determinant has a clear geometric meaning which can be seen from the action of the operator $\widehat{L}$ on orthonormal basis, as illustrated in the Figure 4.9. Denoting $L_{11} = L_1$, $L_{12} = L_2$, $L_{21} = L_3$, and $L_{22} = L_4$, we can calculate the area of the parallelogram built from the vectors $(\widehat{L}\,\vec{u}_1)$ and $(\widehat{L}\,\vec{u}_2)$ as follows:

$$A = (L_1 + L_3)(L_2 + L_4) - 2(L_1 L_2/2 + L_3 L_4/2 + L_2 L_3) \,.$$

It is simply the area of the rectangle $O126$ minus the areas of two rectangles with sides L_2 and L_3, together with two pairs of triangles.

After simplification the last equation reduces to

$$A = L_1 L_4 - L_2 L_3 = L_{11} L_{22} - L_{12} L_{21} \,.$$

Thus, the determinant of a linear operator L equals to the ratio of the area of the parallelogram built from the vectors $(\widehat{L}\,\vec{u}_1)$ and $(\widehat{L}\,\vec{u}_2)$

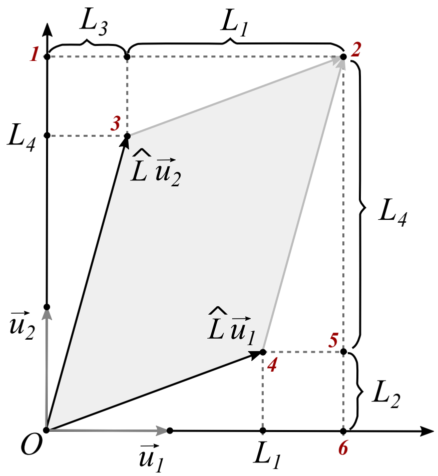


Fig. 4.9: A linear operator $\widehat{L}$ transforms basis vectors $\vec{u}_i$ into $\vec{v}_i = \widehat{L}\,\vec{u}_i$. For non-degenerate operator the vectors $\vec{v}_i$ form sides of a parallelogram with non-zero area $A = L_{12}L_{21} - L_{11}L_{22}$.

to the area of the parallelogram built from the unit vectors $\vec{u}_1$ and $\vec{u}_2$. Obviously, for a degenerate linear operator, this ratio is zero, since all arrows "collapse" onto a single direction.

> ***Determinant***
> For arrow-vectors in three dimensions and higher, we also have linear operators. The meaning of determinant in these higher-dimensional cases remains similar: Determinant expresses the ratio of *volumes* built from basis vectors $\vec{e}_1, \vec{e}_2 \ldots, \vec{e}_n$, and from their "transformed" versions $\widehat{L}\vec{e}_1, \widehat{L}\vec{e}_2 \ldots, \widehat{L}\vec{e}_n$.

4.6 Using Operator Components

It is possible, and often convenient, to work with a linear operator using only its components, without referring to arrows or some other graphical representation.

The action of a linear operators $\widehat{L}$ on a vector $\vec{a}$ can be written in terms of components:

$$\widehat{L}\,\vec{a} = \widehat{L}\,(a_i \vec{e}_i) = a_i\,(\widehat{L}\,\vec{e}_i) = a_i (L_{ij}\,\vec{e}_j).$$

On the other hand

$$\widehat{L}\,\vec{a} = \vec{b} = b_j \vec{e}_j\,.$$

Comparing the last two expressions, we observe that the action of a linear operator can be written entirely in components:

$$\boxed{a_i L_{ij} = b_j\,.}$$

Exercise 4.4
Show that the same relation holds in any other basis. Namely, prove that

$$a'_i L'_{ij} = b'_j\,.$$

The relation between components of the operator $\widehat{L}$ and vectors $\vec{a}$ and $\vec{b}$ can be written differently:

$$b_j = L_{ij} a_i \;\;\to\;\; b_j = L^T_{ji} a_i \;\;\to\;\; b_i = L^T_{ij} a_j\,.$$

Here we used the operation of transposition (introduced in the subsection **??** on page **??**) and index renaming to make summation over the second index of L^T_{ij}.

4.7 Components Transformation

This section has the highest density of algebraic manipulations in the whole book. However, the manipulations are rather trivial, they consist only in multiplications and additions. The challenge for the reader is to stay focused and follow the derivations closely because the final result is very important. The symbolic operations in this section are typical for linear algebra. Without good notation such calculations may become too tedious. This is why we will first demonstrate how to use Einstein's summation rule to quickly get the desired result. Only after that we will derive the same result again, showing all steps in details. Now let's get the job done!

At the first look at components, the difference between vectors and linear operators is in the number of component indices: one for vectors (a_i), and two for linear operators (L_{ij}). To further compare vectors and linear operators, we can study how the components of the latter transform between different bases.

Power of Index Notation

We first demonstrate the power of Einstein's summation rule and index notation and find how the transformation relations can be quickly deduced.

We start with the expansion of an operator in the "new" (primed) basis:

$$\widehat{L}\,\vec{e}\,'_i = L'_{il}\,\vec{e}\,'_l\,. \tag{4.2}$$

Next, we use the linearity of $\widehat{L}$ and the relation between "new" and "old" basis vectors to fully expand the left-hand side:

$$\vec{e}\,'_i = E_{ij}\,\vec{e}_j\,, \quad \vec{e}_k = E'_{kl}\,\vec{e}\,'_l$$

$$\widehat{L}\,\vec{e}\,'_i = \widehat{L}\,(E_{ij}\,\vec{e}_j) = E_{ij}(\widehat{L}\,\vec{e}_j) = E_{ij}L_{jk}\,\vec{e}_k = E_{ij}L_{jk}E'_{kl}\,\vec{e}\,'_l$$

Comparing this to the right-hand side of the equation (4.2), we arrive at

$$\boxed{L'_{il} = E_{ij}L_{jk}E'_{kl}\,.}$$

Now we will repeat the steps in details.

The components (L'_{ij}) of a linear operator $\widehat{L}$ in the "new" (primed) basis are determined in the same way as for the "old" basis:

$$\widehat{L}\,\vec{e}\,'_1 = L'_{11}\,\vec{e}\,'_1 + L'_{12}\,\vec{e}\,'_2\,,$$

$$\widehat{L}\,\vec{e}\,'_2 = L'_{21}\,\vec{e}\,'_1 + L'_{22}\,\vec{e}\,'_2\,.$$

Next, in the left-hand sides of these equations we replace $\vec{e}\,'_i$ with their expansion in the "old" basis and use the linearity of the operator $\widehat{L}$:

$$\widehat{L}\,\vec{e}\,'_1 = \widehat{L}\,(E_{11}\,\vec{e}_1 + E_{12}\,\vec{e}_2) = E_{11}(\widehat{L}\,\vec{e}_1) + E_{12}(\widehat{L}\,\vec{e}_2)$$

and

$$\widehat{L}\,\vec{e}\,'_2 = \widehat{L}\,(E_{21}\vec{e}_1 + E_{22}\vec{e}_2) = E_{21}(\widehat{L}\,\vec{e}_1) + E_{22}(\widehat{L}\,\vec{e}_2)\,.$$

The action of $\widehat{L}$ on the "old" basis is determined by the components L_{ij}:

$$\widehat{L}\,\vec{e}_i = L_{ij}\vec{e}_j\,.$$

Using these relations, we further transform

$$\widehat{L}\,\vec{e}\,'_1 = E_{11}(L_{11}\vec{e}_1 + L_{12}\vec{e}_2) + E_{12}(L_{21}\vec{e}_1 + L_{22}\vec{e}_2)$$

and

$$\widehat{L}\,\vec{e}\,'_2 = E_{21}(L_{11}\vec{e}_1 + L_{12}\vec{e}_2) + E_{22}(L_{21}\vec{e}_1 + L_{22}\vec{e}_2)\,.$$

Opening the parentheses and grouping the terms with identical basis vectors, we get

$$\widehat{L}\,\vec{e}\,'_1 = \mu_1\vec{e}_1 + \mu_2\vec{e}_2$$

and

$$\widehat{L}\,\vec{e}\,'_2 = \nu_1\vec{e}_1 + \nu_2\vec{e}_2\,,$$

where, in order to avoid clutter, we introduced notation

$$\mu_1 = E_{11}L_{11} + E_{12}L_{21}\,, \quad \mu_2 = E_{11}L_{12} + E_{12}L_{22}$$

and

$$\nu_1 = E_{21}L_{11} + E_{22}L_{21}\,, \quad \nu_2 = E_{21}L_{12} + E_{22}L_{22}\,.$$

Finally, we expand the "old" basis vectors in terms of the "new" ones, to arrive at

$$\widehat{L}\,\vec{e}\,'_1 = \mu_1(E'_{11}\vec{e}\,'_1 + E'_{12}\vec{e}\,'_2) + \mu_2(E'_{21}\vec{e}\,'_1 + E'_{22}\vec{e}\,'_2)$$

and

$$\widehat{L}\,\vec{e}\,'_2 = \nu_1(E'_{11}\vec{e}\,'_1 + E'_{12}\vec{e}\,'_2) + \nu_2(E'_{21}\vec{e}\,'_1 + E'_{22}\vec{e}\,'_2)\,.$$

Opening the parentheses and grouping the terms with identical basis vectors results in

$$\widehat{L}\,\vec{e}\,'_1 = (\mu_1E'_{11} + \mu_2E'_{21})\vec{e}\,'_1 + (\mu_1E'_{12} + \mu_2E'_{22})\vec{e}\,'_2$$

and

$$\widehat{L}\,\vec{e}\,'_2 = (\nu_1E'_{11} + \nu_2E'_{21})\vec{e}\,'_1 + (\nu_1E'_{12} + \nu_2E'_{22})\vec{e}\,'_2\,.$$

From the last two equations, we can read the components L'_{ij} of the operator $\widehat{L}$ in terms of its components L_{mn}:

$$L'_{11} = \mu_1 E'_{11} + \mu_2 E'_{21}\,, \quad L'_{12} = \mu_1 E'_{12} + \mu_2 E'_{22}\,,$$

$$L'_{21} = \nu_1 E'_{11} + \nu_2 E'_{21}\,, \quad L'_{22} = \nu_1 E'_{12} + \nu_2 E'_{22}\,.$$

To keep the formulas manageable, we will use the summation convention and first write

$$L'_{11} = \mu_i E'_{i1}\,, \quad L'_{12} = \mu_i E'_{i2}\,,$$

$$L'_{21} = \nu_i E'_{i1}\,, \quad L'_{22} = \nu_i E'_{i2}\,.$$

Second, observing the matching indices, we can shorten these relations even further:

$$L'_{1j} = \mu_i E'_{ij}\,,$$

$$L'_{2j} = \nu_i E'_{ij}\,.$$

Third, having looked at the expressions for μ_i and ν_i we can shorten them to the following

$$\mu_i = E_{1k} L_{ki} \text{ and } \nu_i = E_{2k} L_{ki}\,.$$

Plugging these expressions into the formulas for L'_{ij}, we obtain

$$L'_{1j} = E_{1k} L_{ki} E'_{ij} \text{ and } L'_{2j} = E_{2k} L_{ki} E'_{ij}\,.$$

Finally, comparing the indices, we arrive at the most compact form of relations:

$$L'_{mj} = E_{mk} L_{ki} E'_{ij}\,. \tag{4.3}$$

Comparing this to the transformation of components of a contravariant vector:

$$a'_j = a_i E'_{ij},$$

we see both similarities and differences.

The similarity is seen in the transformation rule for the second index of the operator $\widehat{L}$: It transforms according to the *contravariant* rule. The difference is seen in the transformation rule for the first index which transforms according to the *covariant* rule.

We conclude that linear operators $\widehat{L}$ which map vector into vectors:

$$\vec{a} \xrightarrow{\widehat{L}} \vec{b}$$

have a "mixed nature": They combine behavior of covariant and contravariant vectors, as seen from the transformation of their components. Soon we will encounter linear operators whose components transform like two covariant vectors, or like two contravariant vectors (see *tensor product* in Section 5.8).

Objects and Components

Notions of vectors and operators are *independent of any basis and components*. Components simply *represent* vectors or operators in a given basis. Neither vectors nor operators are reduced to their components, like a building is not reduced to its projections on a piece of paper.

Components are simply computational tool; but their transformation properties tell something about the corresponding mathematical object.

Exercise 4.5

The area of the parallelogram built on unit vectors is 1. The area of parallelogram built on the vectors $(\widehat{L}\vec{u}_1)$ and $(\widehat{L}\vec{u}_2)$ does not depend on the basis used to express the components of the operator $\widehat{L}$. Therefore, the determinant should be a quantity independent of the basis.

Another characteristic of an operator which is independent of the basis is called *trace*. It is defined as the sum

$$\boxed{tr\,\widehat{L} = L_{11} + L_{22}\,.}$$

Prove that trace is independent of the choice of a specific basis. That is, show that, given two sets of basis vectors $\{\vec{e}_i\}$ and $\{\vec{e}\,'_i\}$, the following equality holds

$$L_{11} + L_{22} = L'_{11} + L'_{22}\,.$$

4.8 First Notion of Tensor

Although a more satisfactory definition of tensors will be deduced later, we can now recognize at least one type of tensor. This type of tensor is represented by linear operators considered above.

Tensors are mathematical objects which, like numbers and vectors, can be added pairwise, multiplied by numbers, and, very importantly, *whose components transform in a certain way.* The components of linear operators of the type $\widehat{L}\,\vec{a} = \vec{b}$ transform according to the formula

$$\boxed{L'_{ij} = E_{im} L_{mn} E'_{nj}\,.}$$

Later in Section 5.8 we will encounter tensors with other transformation rules. However, those rules will be analogous to the expression given above. We will also learn what it means to add two tensors, and even what it means to "multiply" a pair of tensors.

Finally, another important aspect of operators and tensors can be seen when we write

$$\vec{b} = \widehat{L}\,\vec{a} \qquad \text{or} \qquad b_j = a_i L_{ij}\,.$$

These formulas express a *linear relation* between two vectors, and the operator $\widehat{L}$ (or tensor) encodes that relation. Linear relations between vector quantities are very important in physics. Examples of this are given in the last chapter.

Chapter Highlights

- *Operators extends the idea of functions.*
- *Numeric functions (e.g.,* $\sin x$*) act on numbers and yield other numbers. Operators may act on vectors to yield other vectors or numbers.*
- *Linear operators represent the simplest and yet powerfull class of operators on vectors.*
- *Linear operators can be represented graphically or symbolically.*
- *In a given basis, every linear operator can be specified using components. This is similar how a vector is represented via its components.*
- *When basis is changed, the components of a linear operator transform in a*

very specific way, called transformation rule or law.

- *Mathematical expression involving linear operators can be written using components (e.g., L_{ij}) or they can be written in more abstract operator form using operator notation $\widehat{L}$.*
- *Degenerate linear operators are special subset of linear operators that collapse two or more basis vectors onto the same line.*
- *There are two characteristics of linear operators that do not change when new basis is chosen: Determinant and trace.*
- *Linear operators represent the simplest type of tensors, called tensors of the second rank – one tier above vectors.*
- *Linear operators and tensors are used to express linear relations between different vector quantities.*

5. Quantum Physics

THE first type of operators – and corresponding tensors – that we encountered has a simple type:

$$\widehat{L}\,\vec{a} = \vec{b}\,.$$

It is a linear unary function mapping vectors into vectors.

As the next step, we will expand our toolset and study operators that take two vectors as their input arguments. Their action on the arguments can be symbolically written as follows:

$$\widehat{L}\,\vec{a}\,\vec{b} = x$$

if the results is a number, or

$$\widehat{L}\,\vec{a}\,\vec{b} = \vec{c}$$

if the results is another vector. We will start with the former case.

5.1 Dol-Operator and Scalar Product

Given a pair of vectors, for example arrows in a plane, we can *quantify their mutual orientation*. In other words, given two vectors $\vec{a}$ and $\vec{b}$, we can calculate a number that measures, for instance, their *degree of overlap* (or *alignment*) which tells how much in common the vectors have with regard to their directions and lengths.

We are looking for a binary operator $\widehat{\sigma}$ that yields a number based on two vectors:

$$\widehat{\sigma}\,\vec{a}\,\vec{b} = x\,.$$

We will call this operator $\widehat{\sigma}$ *dol*-operator[1], based on the key letters of the phrase "degree of overlap".

There are many methods to quantify mutual orientation of a pair of vectors. One simple way is to measure the angle between them:

$$\angle\,\vec{a}\,\vec{b} = \theta\,.$$

However, we will limit our search for candidates to linear operators or – in the case of binary operator $\widehat{\sigma}$ – to *bilinear operators*.

Exercise 5.1
Prove that the operator $\angle$ is not linear.

If the dol operator $\widehat{\sigma}$ is linear in each of its arguments, we must have

$$\widehat{\sigma}\,(2\vec{a})\,\vec{b} = \widehat{\sigma}\,(\vec{a}+\vec{a})\,\vec{b} = \widehat{\sigma}\,\vec{a}\,\vec{b} + \widehat{\sigma}\,\vec{a}\,\vec{b} = 2(\widehat{\sigma}\,\vec{a}\,\vec{b})\,.$$

In general, we require

$$\widehat{\sigma}\,(\alpha\vec{a})\,\vec{b} = \alpha(\widehat{\sigma}\,\vec{a}\,\vec{b})\,.$$

In addition, we will not be concerned with the order in which the vectors $\vec{a}$ and $\vec{b}$ are considered. Put differently, we consider $\vec{a}$ aligned to $\vec{b}$ to the same degree as $\vec{b}$ is aligned to $\vec{a}$. Mathematically this requirement is written as follows:

$$\widehat{\sigma}\,\vec{a}\,\vec{b} = \widehat{\sigma}\,\vec{b}\,\vec{a}\,.$$

Functions and operators with this property are called *symmetric* in their arguments.

Note
Since angles are measured as clockwise (negative) or counter-clockwise (positive), the operator $\angle$ is not symmetric. It is called *anti-symmetric*:

$$\angle\,\vec{a}\,\vec{b} = -(\angle\,\vec{b}\,\vec{a})\,.$$

[1] This is not a standard terminology.

Finally, we recognize that some pairs of vectors are not overlapping. It can be said that their degree of overlap is zero. Mutually orthogonal vectors are example of vectors with zero overlap, they kind of "have nothing in common." Thus, we expect

$$\widehat{\sigma}\,\vec{a}\,\vec{b} = 0$$

if $\vec{a} \perp \vec{b}$.

Now every vector can be written in the form

$$\vec{a} = a\vec{u}_a\,,$$

where a is the length of the vector, $\vec{u}_a$ is the unit-length vector pointing in the same direction as $\vec{a}$. Similarly for $\vec{b}$:

$$\vec{b} = b\vec{u}_b\,.$$

Since dol-operator $\widehat{\sigma}$ is linear in each argument, we can write

$$\widehat{\sigma}\,\vec{a}\,\vec{b} = ab(\widehat{\sigma}\,\vec{u}_a\,\vec{u}_b)\,.$$

The problem is thus reduced to quantifying mutual overlap of unit vectors $\vec{u}_a$ and $\vec{u}_b$. Although we can't use the angle between these vectors directly, we can use some function of the angle:

$$\widehat{\sigma}\,\vec{u}_a\,\vec{u}_b = f\,\theta\,.$$

Firstly, this function has to be periodic since changing the mutual angle by 2π results in the same mutual orientation. Secondly, for orthogonal vectors the function must return zero degree of overlap:

$$f\,(\pi/2) = 0\,.$$

After some search and reflection, a reasonable candidate function can be written in the following way:

$$f\,\theta = \cos\theta\,.$$

We arrived at the traditional operation of *scalar product* (numeric product) of two vectors. For scalar product a special *infix notation* is conventionally used:

$$\boxed{\widehat{\sigma}\,\vec{a}\,\vec{b} = \vec{a}\cdot\vec{b} = ab\cos\theta\,.}$$

What is the geometric meaning of scalar product? As shown in the Figure 5.1, one way to interpret the scalar product $\vec{a}\cdot\vec{b} = ab\cos\theta$ is to consider it as the area of a rectangle with sides a and $b\cos\theta$. A special care must be taken for cases when the mutual angle is greater than π.

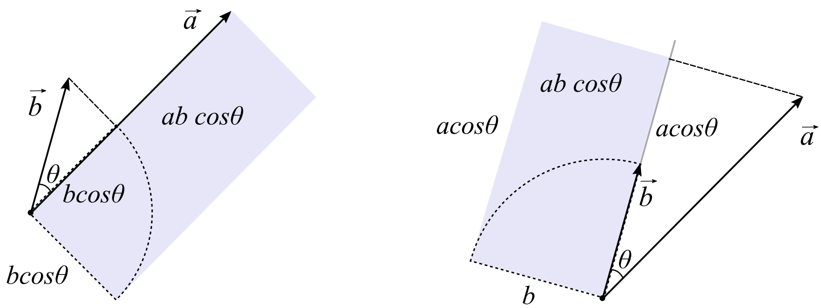


Fig. 5.1: Scalar product of two vectors can be given a simple geometric interpretation as the unsigned area of a rectangle with the sides a and $b\cos\theta$ or with the sides $a\cos\theta$ and b.

5.2 Scalar Product Properties

The expression for scalar product of two vectors

$$\vec{a}\cdot\vec{b} = ab\cos\theta$$

implies the *commutativity* of scalar product operation:

$$\vec{b}\cdot\vec{a} = ba\cos\theta = ab\cos\theta = \vec{a}\cdot\vec{b}\,.$$

> ***Non-commutativity***
> Commutativity is a familiar property, present for both addition and multiplication of numbers. It is sometimes accepted as a natural property of any multiplication-like operation. This view is limiting, however, and we will later learn how to multiply objects without commutativity:
>
> $$A \bowtie B \neq B \bowtie A\,.$$

Here we used an arbitrary infix symbol $\bowtie$ to denote non-commutative product of some objects A and B. What exactly those objects are and what their product might mean will be clear when we introduce them properly. Right now we want to highlight the non-commutativity as a valid property of many binary operations.

Besides commutativity scalar product has other useful properties. First, it is trivial to demonstrate that we can "pull out" constant scale factors of vectors:

$$(\alpha\vec{a})\cdot\vec{b} = \alpha(\vec{a}\cdot\vec{b})$$

and similarly

$$\vec{a}\cdot(\beta\vec{b}) = \beta(\vec{a}\cdot\vec{b})\,.$$

To arrive at the second important property of scalar product, recall that in the geometric interpretation of the scalar product we had expressions

$$b\cos\theta = b_{\parallel} \qquad \text{or} \qquad a\cos\theta = a_{\parallel}\,,$$

where $b_{\parallel}$ is the part of the vector $\vec{b}$ parallel to $\vec{a}$ ($a_{\parallel}$ is the part of the vector $\vec{a}$ parallel to $\vec{b}$.)

Now, if the vector $\vec{a}$ is "made from" two other vectors

$$\vec{a} = \vec{c} + \vec{d}\,,$$

then, as illustrated in the Figure (5.2),

$$a_{\parallel} = c_{\parallel} + d_{\parallel}\,,$$

where all terms represent parts of the respective vectors parallel to the vector $\vec{b}$. Of course, the angles between the vectors $\vec{a}$, $\vec{c}$, $\vec{d}$ and the vector $\vec{b}$ may all be different:

$$a_{\parallel} = a\cos\theta = c\cos\phi + d\cos\psi\,.$$

Given that

$$\vec{a}\cdot\vec{b} = ab\cos\theta\,, \quad \vec{c}\cdot\vec{b} = cb\cos\phi\,, \quad \vec{d}\cdot\vec{b} = db\cos\psi\,,$$

we arrive at the *distributive property* of scalar product:

$$(\vec{c} + \vec{d}) \cdot \vec{b} = \vec{a} \cdot \vec{b} = (\vec{c} \cdot \vec{b}) + (\vec{d} \cdot \vec{b}) \,.$$

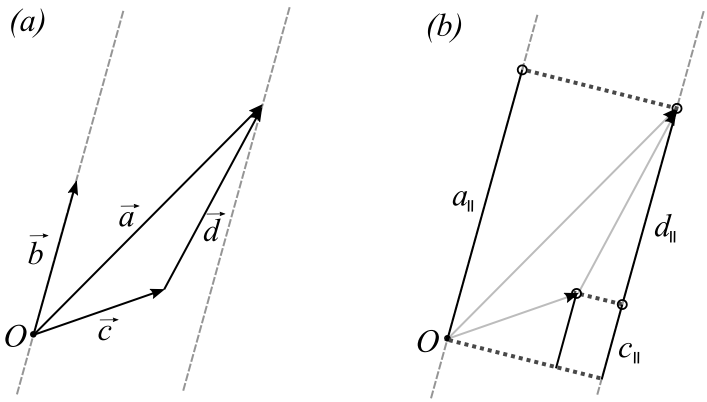


Fig. 5.2: Simple geometric construction illustrates the distributive property of scalar product: $(\vec{c} + \vec{d}) \cdot \vec{b} = (\vec{c} \cdot \vec{b}) + (\vec{d} \cdot \vec{b})$.

Putting it all together, we can express the properties of scalar product in a single expression:

$$(\alpha\vec{a} + \beta\vec{b}) \cdot \vec{c} = \vec{c} \cdot (\alpha\vec{a} + \beta\vec{b}) = \alpha(\vec{a} \cdot \vec{c}) + \beta(\vec{b} \cdot \vec{c}) \,.$$

Orthogonality

We might wonder whether we can define, instead of degree of alignment, some measure of orthogonality for a pair of vectors? Indeed, a reasonable candidate might be (using infix notation):

$$\vec{a} \vdash \vec{b} = ab \sin\theta \,.$$

While this is an acceptable definition, the expression on the right-hand side appears in a different, more powerful and useful, operation of *outer product* of two vectors. The concept of outer product is related to the concept of *tensor product* (see Section 5.8) but is beyond the scope of this book.

Using the properties of scalar product and expanding vectors in terms of the basis vectors, we can write scalar product in terms of vector components. First write

$$(a_1\vec{e}_1 + a_2\vec{e}_2)\cdot\vec{b} = a_1(\vec{b}\cdot\vec{e}_1) + a_2(\vec{b}\cdot\vec{e}_2)\,.$$

Next, expand $\vec{b}$ to find

$$(b_1\vec{e}_1 + b_2\vec{e}_2)\cdot\vec{e}_i = b_1(\vec{e}_1\cdot\vec{e}_i) + b_2(\vec{e}_2\cdot\vec{e}_i)\,.$$

Finally, combining the results of two previous steps, we arrive at the expression

$$\vec{a}\cdot\vec{b} = a_1b_1(\vec{e}_1\cdot\vec{e}_1) + a_1b_2(\vec{e}_2\cdot\vec{e}_1) + a_2b_1(\vec{e}_1\cdot\vec{e}_2) + a_2b_2(\vec{e}_2\cdot\vec{e}_2)\,.$$

For arbitrary basis vectors the scalar product is then given by

$$\boxed{\vec{a}\cdot\vec{b} = a_1b_1e_1^2 + (a_1b_2 + a_2b_1)e_1e_2\cos\theta + a_2b_2e_2^2\,,}$$

where θ is the angle between the basis vectors $\vec{e}_1$ and $\vec{e}_2$. For a special case of orthonormal basis the scalar product takes the simplest form:

$$\boxed{\vec{a}\cdot\vec{b} = a_1b_1 + a_2b_2 = a_ib_i\,.}$$

But in general we must know the values for all products $\vec{e}_i\cdot\vec{e}_j$. A special notation is used for these products:

$$\eta_{ij} = \vec{e}_i\cdot\vec{e}_j = \widehat{\sigma}\,\vec{e}_i\,\vec{e}_j\,.$$

This notation allows a more compact way of writing scalar product for general basis:

$$\boxed{\vec{a}\cdot\vec{b} = \eta_{ij}a_ib_j\,.}$$

Scalar Product Components

Let's take a look how the derivation of the last result can be done using index notation.

$$\vec{a} \cdot \vec{b} = (a_i \vec{e}_i) \cdot \vec{b} = a_i (\vec{e}_i \cdot \vec{b}) \,.$$

Expanding $\vec{b}$, we get

$$\vec{a} \cdot \vec{b} = a_i (\vec{e}_i \cdot [b_j \vec{e}_j]) = a_i (b_j [\vec{e}_i \cdot \vec{e}_j]) \,.$$

We thus showed that

$$\vec{a} \cdot \vec{b} = a_i b_j \vec{e}_i \cdot \vec{e}_j \,.$$

In the process we had to use twice the distributive property of the scalar product, as well as distributive property of number multiplication.

Exercise 5.2
In the index form of scalar product of two vectors

$$\vec{a} \cdot \vec{b} = (a_i b_j)(\vec{e}_i \cdot \vec{e}_j)$$

we observe the expression with two indices:

$$\beta_{ij} = a_i b_j \,.$$

Can β_{ij} represent the components of some linear operator $\widehat{\beta}$? If so, how does this operators act on vectors?

5.3 Inner Operations

The following fact is easy to take for granted and overlook: *Vectors are not used by themselves. They need numbers.* Without multiplying a vector by a number, we could not write the simplest expansion of a vector $\vec{a}$ in a basis:

$$\vec{a} = a_1 \vec{e}_1 + a_2 \vec{e}_2 \,.$$

Note that all operations we considered so far never took us outside of the combined realm of numbers $\mathbb{R}$ and vectors $\overset{\Rightarrow}{A}$. Indeed, a product of a number and a vector $\alpha\vec{a}$ uses one element of each space, and produces a vector. Similarly, a sum of two vectors $\vec{a} + \vec{b}$ takes two elements from $\vec{\vec{A}}$ and returns another element of $\vec{\vec{A}}$. Finally, the dol-operator $\widehat{\sigma}$ takes two elements from $\vec{\vec{A}}$ and returns an element of $\mathbb{R}$. These points are illustrated in the Figure **??**(a).

Another way to look at this is to notice that in the hierarchy of mathematical objects the operations we considered so far never took us up the ladder of *ranks*, as illustrated in the Figure **??**(b). At the lowest level we have *rank-0* elements – numbers. The first ladder corresponds to the *rank-1* elements – vectors. One step higher we have *rank-2* elements – tensors of the second rank, and so on.

All operations that *do not* result in the element of higher rank are called *inner operation*. For instance, scalar product of vectors is often called *inner product*[2]. The phrase *inner sum* is not used for the binary operator $(+)$.

5.4 Conjugate Objects

Starting from the dol-operator

$$\widehat{\sigma}\,\vec{a}\,\vec{b} = x$$

we can arrive at an important notion of *conjugate* objects. Conjugate objects, roughly speaking, are objects that are somehow related via a simple rule. We will study this notion using vectors.

For every vector $\vec{a}$ there exists a mathematical object, related to $\vec{a}$ via the dol-operator. To see this, we first need to revisit the idea of *partial application*, discussed in subsection 5.4.1. This time, however, we will extend the idea of partial application to binary operators.

5.4.1 Partial Application

Given a pair of vectors, the dol-operator yields a number:

$$\widehat{\sigma}\,\vec{a}\,\vec{b} = x\,.$$

What happens when we provide only one vector, leaving the second

[2]There also exists an *outer product*, which is related to *tensor product* discussed in Section 5.8.

input slot of $\widehat{\sigma}$ empty (the box here indicates a missing second argument):

$$\widehat{\sigma}\,\vec{a}\,\square\,?$$

This is called *partial application* of the operator $\widehat{\sigma}$. This construction has the behavior of a unary operator that maps any vector into a number:

$$\vec{c} \xrightarrow{\widehat{\sigma}\,\vec{a}} y\,.$$

> ***Linear Operator***
> The unary operator
> $$\widehat{L}_a = \widehat{\sigma}\,\vec{a}$$
> is a *linear operator*:
> $$\widehat{L}_a(\vec{b} + \vec{c}) = (\widehat{L}_a\,\vec{b}) + (\widehat{L}_a\,\vec{c})\,.$$

We will use a special notation for the partially applied operator:

$$\boxed{\overleftarrow{a} = \widehat{\sigma}\,\vec{a}\,.}$$

The action of a unary operator $\overleftarrow{a}$ on any vector is then naturally defined as

$$\overleftarrow{a}\;\vec{b} = \widehat{\sigma}\,\vec{a}\;\vec{b} = ab\cos\theta\,.$$

Notice that in the left-most expression the operator $\overleftarrow{a}$ and the argument vector $\vec{b}$ are separated by space, in agreement with the notation for application of functions and operator to their arguments.

All three possible "states" of the dol-operator $\widehat{\sigma}$ are illustrated in the Figure 5.3. The middle state of partial application corresponds to a linear operator built from the first input vector $\vec{a}$ and as denoted as $\overleftarrow{a}$. The operator $\overleftarrow{a}$ is called the *conjugate* to a vector $\vec{a}$. The conjugation is understood *relative to the binary operator* $\widehat{\sigma}$.

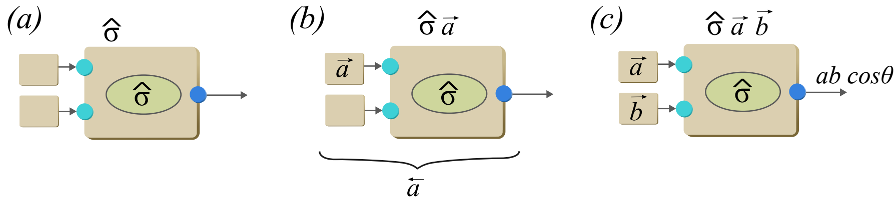


Fig. 5.3: (a) The dol-operator is a binary operator linear in each of its arguments (*bilinear*). (b) When only one argument is supplied, the operator becomes *partially applied* and is denoted using the *conjugate vector* notation $\overleftarrow{a}$. (c) Applied to two vector arguments (*fully applied*), dol-operator yields a number – scalar product of the vectors.

Quantum Notation

In quantum physics vectors and their duals are used to describe properties of quantum systems. Paul Dirac introduced a very powerful notation for these vectors, called *bra* and *ket* vectors. Ket vectors correspond to contravariant arrow-vectors and are denoted as follows:

$$\vec{a} \quad \longrightarrow \quad |a\rangle\,.$$

Bra vectors correspond to covariant vectors and are denoted as

$$\overleftarrow{a} \quad \longrightarrow \quad \langle a|\,.$$

Scalar product of bra and ket vectors is then written using *brackets*:

$$\overleftarrow{a}\,\vec{a} = \langle a|a\rangle\,.$$

5.5 Conjugate Vectors

The notation for the conjugate vectors is suggestive for a reason. The left-pointing arrow on top indicates that this is a vector. This might be surprising, since we just convinced ourselves that $\overleftarrow{a}$ is a linear operator. To convince ourselves that $\overleftarrow{a}$ is also a vector, we must check whether the conjugate of vectors possess defining properties of vectors. Let's do it.

To demonstrate that we can add two conjugate vectors – which are also linear operators – we must describe how to add operators. Operators are essentially functions, and we already understand how to add functions of a numerical arguments (go back to subsection **??** if you need a refresher.)

We can add two operators, $\overleftarrow{a}$ and $\overleftarrow{b}$, in a similar way:

$$\overleftarrow{a} + \overleftarrow{b} = \overleftarrow{c}\,,$$

where

$$\overleftarrow{c}\,\vec{d} = (\overleftarrow{a} + \overleftarrow{b})\,\vec{d} = (\overleftarrow{a}\,\vec{d}) + (\overleftarrow{b}\,\vec{d})\,.$$

As a sidenote, we point out once more that the addition operation "+" in $\overleftarrow{a} + \overleftarrow{b}$ is applied to a new type of mathematical objects – conjugate vectors. The addition operator in $\overleftarrow{a}\,\vec{d} + \overleftarrow{b}\,\vec{d}$ is applied to usual numbers.

It is easy to see how conjugate vectors can be multiplied by numbers:

$$\alpha\overleftarrow{a} = \overleftarrow{b}\,,$$

where

$$\overleftarrow{b}\,\vec{c} = \alpha(\overleftarrow{a}\,\vec{c})\,.$$

Conjugate Basis

If conjugate objects are vectors, they must have some basis. The basis for conjugate vectors can be taken by conjugating any basis from the arrow-vectors:

$$\boxed{\overleftarrow{e}_1 = \widehat{\sigma}\,\vec{e}_1\,, \qquad \overleftarrow{e}_2 = \widehat{\sigma}\,\vec{e}_2\,.}$$

The linearity of the operator $\widehat{\sigma}$ for both arguments results in the following relation

$$\overleftarrow{a} = \widehat{\sigma}\,\vec{a} = \widehat{\sigma}\,(a_1\vec{e}_1 + a_2\vec{e}_2) = a_1\overleftarrow{e}_1 + a_2\overleftarrow{e}_2\,.$$

In other words, *any conjugate vector can be expanded in terms of some basis conjugate vectors* $\{\overleftarrow{e}_i\}$.

Exercise 5.3

Derive the relationship

$$\overleftarrow{a} = a_1\overleftarrow{e}_1 + a_2\overleftarrow{e}_2$$

in more details, without leaving out steps.

Component Transformation

Finally, we must show that the components of any conjugate vector change properly when the (conjugate) basis is switched.

We start by writing the same operator $\overleftarrow{a}$ in different bases:

$$\overleftarrow{a} = a_1 \overleftarrow{e}_1 + a_2 \overleftarrow{e}_2 = a_1' \overleftarrow{e}_1' + a_2' \overleftarrow{e}_2' \,.$$

Here

$$\overleftarrow{e}_1' = \widehat{\sigma}\, \overrightarrow{e}_1' \,, \qquad \overleftarrow{e}_2' = \widehat{\sigma}\, \overrightarrow{e}_2' \,.$$

Expanding the primed basis in terms of the non-primed, and using the linearity of the operator $\widehat{\sigma}$ for all arguments, we get

$$\overleftarrow{e}_1' = \widehat{\sigma}\, \overrightarrow{e}_1' = E_{11} \overleftarrow{e}_1 + E_{12} \overleftarrow{e}_2 \,,$$

$$\overleftarrow{e}_2' = \widehat{\sigma}\, \overrightarrow{e}_2' = E_{21} \overleftarrow{e}_1 + E_{22} \overleftarrow{e}_2 \,.$$

Plugging these equations into the expansion of $\overleftarrow{a}$, after grouping the terms, we arrive at

$$\overleftarrow{a} = a_1 \overleftarrow{e}_1 + a_2 \overleftarrow{e}_2 = (a_1' E_{11} + a_2' E_{21}) \overleftarrow{e}_1 + (a_1' E_{12} + a_2' E_{22}) \overleftarrow{e}_2 \,.$$

Comparing the coefficients in front of the basis vectors, we conclude that

$$a_1 = a_i' E_{i1} \,,$$

$$a_2 = a_i' E_{i2} \,.$$

In a more compact form:

$$\boxed{a_j = a_i' E_{ij} \,.}$$

This is the same form we obtained before for arrow-vectors, with the only (unessential) difference of notation – using E instead of L to denote the relation between the “old” and “new” bases.

The conclusion is as follows: *The conjugate vectors have completely analogous properties as the arrow-vectors, when referred to their own conjugate bases.*

We thus showed that vectors in plane have “conjugate image” – a set of linear operators which also behave like vectors. These conjugate vectors belong to the special vector space called *conjugate vector space* or

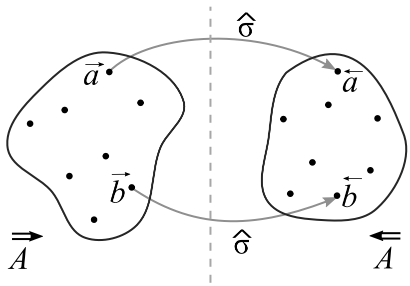


Fig. 5.4: Conjugate vectors like $\overleftarrow{a}$ form a conjugate space which we denote $\overset{\Leftarrow}{A}$. It is *conjugate* or *dual* to the "usual" vectors space $\overset{\Rightarrow}{A}$.

dual vector space. The "usual" vector space is denoted as $\overset{\Rightarrow}{A}$ and its dual companion is denoted as $\overset{\Leftarrow}{A}$. The relationship between the "original" and the *conjugate/dual* space is illustrated in the Figure 5.4.

5.6 Operators Are Also Vectors

Starting with the binary bilinear dol-operator $\widehat{\sigma}$, we obtained unary linear operators using partial application

$$\overrightarrow{a} \quad \Longrightarrow \quad \overleftarrow{a} = \widehat{\sigma}\,\overrightarrow{a}\,.$$

We then showed that all such linear operators $\overleftarrow{a}$ behave like vectors. They can be multiplied by a number, added, written in terms of some basis, and have components that transform according to a special rule:

$$\overleftarrow{a} = a_i \overleftarrow{e}_i = a'_j \overleftarrow{e}'_j$$

where

$$\boxed{a_i = a'_j E_{ji}\,.}$$

We discovered a natural connection (*duality*) between every arrow-vector and its conjugate vector:

$$\overrightarrow{a} \overset{\widehat{\sigma}}{\longleftrightarrow} \overleftarrow{a}\,.$$

We can start with unary linear operators, without specifying their origin, and show that they behave like vectors. We will do it now.

Imagine *all possible* unary linear operators that map vectors into numbers (see Figure 5.5). Let us denote some such operator as $\widehat{\Gamma}$:

$$\widehat{\Gamma}\,\overrightarrow{a} = x_a\,,$$

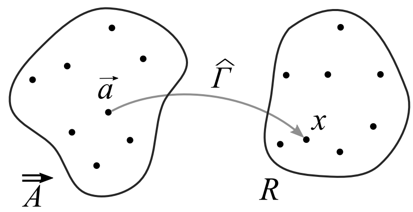


Fig. 5.5: Some linear operators map vectors into numbers: $\widehat{\Gamma}\,\vec{a} = x$. Such operators form a vector space of their own.

$$\widehat{\Gamma}\,(\alpha\vec{a}) = \alpha(\widehat{\Gamma}\,\vec{a}) = \alpha x_a\,.$$

For $\vec{c} = \vec{a} + \vec{b}$:

$$\widehat{\Gamma}\,\vec{c} = (\widehat{\Gamma}\,\vec{a}) + (\widehat{\Gamma}\,\vec{b}) = x_a + x_b = x_c\,.$$

All x_a, x_b, and x_c are real numbers.

We can demonstrate that operators like $\widehat{\Gamma}$ can be added:

$$\widehat{\Gamma} = \widehat{\Gamma}_1 + \widehat{\Gamma}_2$$

and they can be multiplied by numbers:

$$\widehat{\Gamma}_2 = \alpha\widehat{\Gamma}_1\,.$$

We can also find basis and expand an arbitrary operator in that basis:

$$\widehat{\Gamma} = \gamma_1\widehat{\Gamma}_1 + \gamma_2\widehat{\Gamma}_1 + \ldots + \gamma_n\widehat{\Gamma}_n = \gamma_i\widehat{\Gamma}_i$$

and establish the transformation rule for the components γ_i between different bases. In effect, we will demonstrate that unary linear operators of the same type as $\widehat{\Gamma}$ (mapping arrow-vectors into numbers) behave like vectors and therefore must be considered as such.

❗ ***Reminder***
When we say that an operator $\widehat{\Gamma}$ is given or known, we mean that we know how it acts on *any vector* $\vec{a}$:

$$\widehat{\Gamma}\,\vec{a} = x_a\,.$$

Addition

If we know two operators $\widehat{\Gamma}_1$ and $\widehat{\Gamma}_2$, then making sense of their *operator sum* $\widehat{\Gamma} = \widehat{\Gamma}_1 + \widehat{\Gamma}_2$ is easy:

$$\widehat{\Gamma}\,\vec{a} = (\widehat{\Gamma}_1\,\vec{a}) + (\widehat{\Gamma}_2\,\vec{a}) = x_a + y_a\,,$$

where $x_a = \widehat{\Gamma}_1\,\vec{a}$ and $y_a = \widehat{\Gamma}_2\,\vec{a}$.

Multiplication

If we know the operator $\widehat{\Gamma}_1$ then making sense of the product of that operator with any number $\widehat{\Gamma} = \alpha\widehat{\Gamma}_1$ is easy:

$$\widehat{\Gamma}\,\vec{a} = \alpha(\widehat{\Gamma}_1\,\vec{a}) = \alpha x_a\,,$$

where $x_a = \widehat{\Gamma}_1\,\vec{a}$.

Finding Basis

Using the linearity of the operators $\widehat{\Gamma}$, we can apply it to an arbitrary vector $\vec{a}$ as follows:

$$\widehat{\Gamma}\,\vec{a} = \widehat{\Gamma}\,(a_i\vec{e}_i) = a_i\,(\widehat{\Gamma}\vec{e}_i)\,.$$

The last expression states that to know the action of the operator $\widehat{\Gamma}$ on an arbitrary vector $\vec{a}$ it is sufficient to specify its action on all basis vectors $\vec{e}_i$. In other words, the operator $\widehat{\Gamma}$ is fully specified if we know a set of numbers

$$\boxed{\gamma_i = \widehat{\Gamma}\,\vec{e}_i\,.}$$

Notice how similar this last expression is to the definition of components L_{ij} of a linear operator $\widehat{L}$.

Operators and Vectors

A vector $\vec{a}$ is completely determined if we specify its components in a given basis:

$$\vec{a} = a_i\vec{e}_i\,.$$

A linear operator $\widehat{\Gamma}$ that maps vectors into numbers

$$\vec{a} \quad \xrightarrow{\widehat{\Gamma}} \quad x_a$$

is completely determined if we specify its action on basis vectors:

$$\gamma_i = \widehat{\Gamma}\,\vec{e}_i\,.$$

This makes the similarity between vectors and operators $\widehat{\Gamma}$ stronger.

Let's expand the input vector $\vec{a}$ in a different basis:

$$\vec{a} = a'_j \vec{e}\,'_j\,.$$

In this case the action of the linear operator $\widehat{\Gamma}$ will be

$$\widehat{\Gamma}\,(a'_j \vec{e}\,'_j) = a'_j\,(\widehat{\Gamma}\vec{e}\,'_j) = a'_j \gamma'_j\,.$$

Here $\gamma'_j = \widehat{\Gamma}\,\vec{e}\,'_j$.

The relation between the components a_i and a'_i of the contravariant vector $\vec{a}$ is known; the relation between the values γ_i and γ'_j is then easily found:

$$\gamma'_i = \widehat{\Gamma}\vec{e}\,'_i = \widehat{\Gamma}(E_{ij}\vec{e}_j) = E_{ij}\gamma_j\,,$$

and, in a similar way:

$$\gamma_i = \widehat{\Gamma}\vec{e}_i = \widehat{\Gamma}(E'_{ij}\vec{e}\,'_j) = E'_{ij}\gamma'_j\,.$$

These relations correspond to the *covariant vector*.

Covariant Vectors

The components of contravariant vectors allow us to "assemble" them from the "building blocks" – basis vectors:

$$\vec{a} = a_i \vec{e}_i\,.$$

The basis vectors and all other vectors constructed from them all *live in the same vector space*, which we denoted $\overset{\Rightarrow}{A}$ – the space of *contravariant vectors*.

Unary linear operators, including all conjugate vectors, belong to a different vector space – *conjugate* or *dual* to $\overset{\Rightarrow}{A}$. We denoted this vectors space as $\overset{\Leftarrow}{A}$. The Figure 5.4 illustrates this point.

This implies that it is *incorrect* to write

$$\widehat{\Gamma} = \gamma_i \vec{e}_i \,. \qquad (\textit{incorrect!})$$

Conjugate space $\overleftarrow{A}$ has its own basis (or bases).

It is important to understand that the coefficients γ_i and γ'_j refer to bases used for contravariant vectors (bases $\{\vec{e}_i\}$ and $\{\vec{e}\,'_i\}$). They can also refer to bases used for covariant vectors. Let's find some such basis related to the coefficients γ_i.

Every covariant vector $\widehat{\Gamma}$ is completely determined if we know its action on *all* basis contravariant vectors $\vec{e}_i$. Therefore, to specify some basis vectors for $\widehat{\Gamma}$ we should find certain covariant vectors that can be used as "building blocks" for $\widehat{\Gamma}$. As a matter of fact, we already encountered basis covariant vectors when we discussed dol-operator and conjugate vectors of $\vec{e}_i$. We now will define similar vectors without referring to dol operator $\widehat{\sigma}$. Specifically, the first basis vector $\widehat{\Gamma}_1$ for covariant vectors acts on $\vec{e}_i$ as follows:

$$\widehat{\Gamma}_1 \vec{e}_1 = 1 \tag{5.1}$$

$$\widehat{\Gamma}_1 \vec{e}_2 = 0 \tag{5.2}$$

$$\widehat{\Gamma}_1 \vec{e}_3 = 0 \tag{5.3}$$

$$\ldots \tag{5.4}$$

Thus, $\widehat{\Gamma}_1$ returns zero for all basis vectors except for $\vec{e}_1$. Similarly, we define the second basis covariant vector $\widehat{\Gamma}_2$ to return zero for all basis vectors except for $\vec{e}_2$, and so on for other basis covariant vectors. A compact way to express this idea uses index notation:

$$\boxed{\widehat{\Gamma}_i \vec{e}_j = \delta_{ij} \,.}$$

Here δ_{ij} is the Kronecker delta, introduced in Chapter **??** on page **??**.

Having defined this covariant basis, we can write now

$$\widehat{\Gamma} = \gamma_1 \widehat{\Gamma}_1 + \gamma_2 \widehat{\Gamma}_1 + \ldots \gamma_n \widehat{\Gamma}_n = \gamma_j \widehat{\Gamma}_j \,.$$

Clearly,

$$\widehat{\Gamma} \vec{e}_i = \gamma_i \,.$$

(To demonstrate this quickly, recall that $\widehat{\Gamma}\,\vec{e}_i = \gamma_j \widehat{\Gamma}_j\,\vec{e}_i = \gamma_j \delta_{ji} = \gamma_i$.) In other words, the coefficients γ_i are also *components* of the covariant vector $\widehat{\Gamma}$ in *covariant basis* $\{\widehat{\Gamma}_i\}$.

Geometric Representation

Arrows provide a simple geometric representation of *contravariant* vectors. Now that we encountered *covariant* vectors, it is natural to ask what geometric representation do covariant vectors have?

Unary linear operators map vectors into numbers. In certain sense, they "complete" vectors to a mathematical construction that can be unambiguously assigned a number. One such construction is the oriented area (for a plane) or a volume (for three and higher dimensions).

Let's use arrows in three-dimensional space as an example. What completes a given arrow-vector $\vec{a}$ to a volume? We can build a solid figure with a well-defined volume using the arrow as the side of a cylinder, and some two-dimensional area-element as its "complement". This area-element will correspond to a linear operator that maps the vector $\vec{a}$ into a number – the volume of the solid object built using the vector and the area-element, as shown in the Figure 5.6(a).

Covariant Vectors

Contravariant arrow-vectors $\vec{a}$ are oriented line segments regardless of whether they are in a plane, in three-dimensional space, or in spaces of higher dimensions.

Covariant vectors have different "structure" for spaces of different dimensions. In three dimensions, as we saw, they are oriented areas. In a plane, they will be oriented line segments similar to arrow-vectors. In four dimensional space covariant vectors will be oriented three-dimensional volumes. Thus, unlike contravariant arrow-vectors, covariant vectors are more difficult to visualize.

The question of geometric meaning of addition of two covariant vectors – and other operations on them – although interesting and useful, is outside of the scope of this book.

5.7 Projectors

In many problems it is useful to take a vector $\vec{b}$ and find the part of it which will be parallel to another vector $\vec{a}$, as illustrated in the Figure

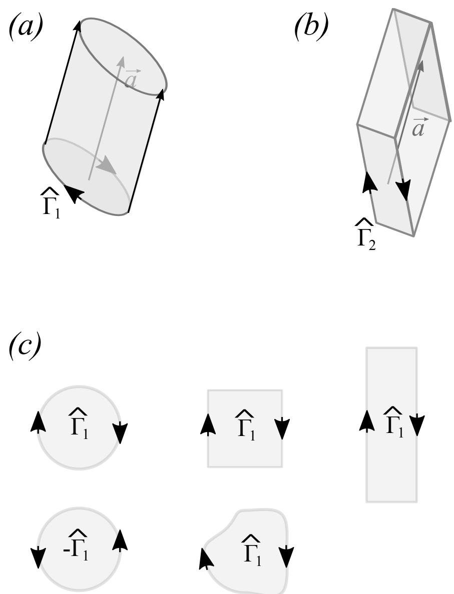


Fig. 5.6: (a) A covariant vector $\widehat{\Gamma}_1$ can be represented by an oriented piece of plane with certain area. Its action on a contravariant vector $\vec{a}$ results in a number – volume of a skewed cylinder built by moving the area along the vector. (b) A covariant vector $\widehat{\Gamma}_2$ has different orientation and magnitude from $\widehat{\Gamma}_1$. (c) The shape of the conjugate vector (an oriented piece of plane) does not matter, as long as its orientation and area stay the same.

5.7. This procedure can be described using the concept of a binary operator. This operator, when given a pair of vectors $\vec{a}$ and $\vec{b}$, returns the "component" of $\vec{b}$ oriented along $\vec{a}$ – a *projection* of $\vec{b}$ onto $\vec{a}$:

$$\widehat{P}\,\vec{a}\,\vec{b} = \vec{b}_{\parallel}\,,$$

where $\vec{a}$ is parallel to $\vec{b}_{\parallel}$.

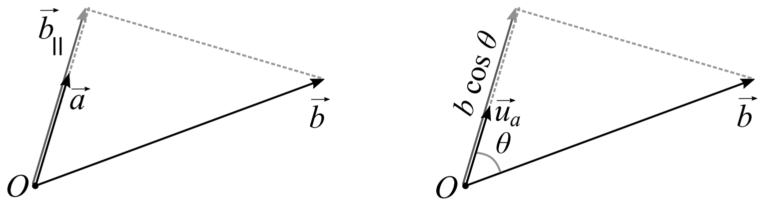


Fig. 5.7: A vector $\vec{b}$ has the "component" $\vec{b}_{\parallel}$ along a given vector $\vec{a}$. The length of this components is $b\cos\theta$ where θ is the angle between vectors $\vec{a}$ and $\vec{b}$.

From the Figure 5.7, it is clear that

$$\vec{b}_{\parallel} = \vec{u}_a b\cos\theta = \frac{\vec{a}}{a}\,b\cos\theta\,,$$

where $\vec{u}_a$ is a unit-length vector parallel to $\vec{a}$.

The right-hand side of the last expression can be written using dol-operator $\widehat{\sigma}$, or even better, using the conjugate vector notation:

$$\vec{b}_{\parallel} = \frac{\vec{a}}{a^2}\,ab\cos\theta = \frac{\vec{a}}{a^2}\,(\overleftarrow{a}\,\vec{b})\,.$$

For each vector $\vec{a}$ there exists corresponding projection operator that projects all other vectors onto the direction specified by $\vec{a}$.

Projector Notation

A projector operator that projects any vector $\vec{b}$ onto the direction specified by a vector $\vec{a}$ will be denoted as $\underline{\underline{\widehat{A}}}$:

$$\underline{\underline{\widehat{A}}}\,\vec{b} = \frac{\vec{a}}{a^2}\,(\overleftarrow{a}\,\vec{b})\,.$$

Such projector exists for any non-zero vector $\vec{a}$:

$$\vec{a} \longrightarrow \underline{\underline{\widehat{A}}}\,.$$

In this notation the same – but capitalized – letter is used for the projector as for the vector. In addition, the capitalized letter is doubly underlined to remind that we project onto the direction parallel to the specified vector.

Similarly, we will have other projectors

$$\vec{b} \longrightarrow \underline{\underline{\widehat{B}}}, \quad \vec{c} \longrightarrow \underline{\underline{\widehat{C}}}\,\ldots$$

Projector $\underline{\underline{\widehat{A}}}$ corresponding to the vector $\vec{a}$ acts on an arbitrary vector $\vec{b}$ as

$$\underline{\underline{\widehat{A}}}\,\vec{b} = \frac{\vec{a}}{a^2}\,(\overleftarrow{a}\,\vec{b})\,.$$

In the argument-free notation, this operator takes the following form:

$$\boxed{\underline{\underline{\widehat{A}}} = \frac{\vec{a}}{a^2}\,(\overleftarrow{a}) = \frac{\vec{a}\,\overleftarrow{a}}{a^2}\,.}$$

Note that the order of the vectors $\vec{a}$ and $\overleftarrow{a}$ in the last expression is very important because it has completely different meaning from the expression $\overleftarrow{a}\,\vec{a}$. Indeed, as we agreed, the expression

$$\overleftarrow{a}\,\vec{a} = \vec{a}\cdot\vec{a} = a^2$$

yields the length squared of the vector $\vec{a}$. In contrast, the expression

$$\vec{a}\,\overleftarrow{a}$$

works as an operator!

We obtained interesting and useful expression for an operator that accepts a vector $\vec{b}$ and produces another vector as the result. This expression involves some kind of "multiplication" of a vector $\vec{a}$ and its conjugate $\overleftarrow{a}$:

$$\vec{a}\,\overleftarrow{a}\,.$$

It is our first encounter with *tensor product*. We will learn more about this new type of multiplication below.

5.7.1 Projector Components

To find the components of a projector operator

$$\widehat{\underline{\underline{A}}} = \frac{\vec{a}\,\overleftarrow{a}}{a^2}$$

in a given basis, we apply it to the basis vectors:

$$\widehat{\underline{\underline{A}}}\,\vec{e}_i = \frac{\vec{a}}{a^2}\,\overleftarrow{a}\,\vec{e}_i = \frac{a_i}{a^2}\,\vec{a}.$$

Expanding the vector $\vec{a}$ in the same basis, we arrive at

$$\widehat{\underline{\underline{A}}}\,\vec{e}_i = \frac{a_i a_j}{a^2}\,\vec{e}_j = \underline{\underline{A}}_{ij}\,\vec{e}_j\,,$$

from which follows the expression for the components

$$\boxed{\underline{\underline{A}}_{ij} = \frac{a_i a_j}{a^2}\,.}$$

Exercise 5.4

Using the components of a projector

$$\underline{\underline{A}}_{ij} = \frac{a_i a_j}{a^2}\,,$$

calculate its determinant.

Symmetry of Projectors

From the expression for the components of a projector follows that

$$\underline{\underline{A}}_{ij} = \underline{\underline{A}}_{ji} \,,$$

which means that to fully specify its components, we need only 3 numbers (for vectors in a plane we need $\underline{\underline{A}}_{11}$, $\underline{\underline{A}}_{12}$, and $\underline{\underline{A}}_{22}$), as opposed to 4 numbers required to specify a general linear operator.

5.7.2 Composition of Projectors*

The idea of composing two functions - discussed in subsection **??** on page **??** - can be extended to linear operators. That is, some types of linear operators can be composed. Indeed, suppose we have a linear operator

$$\widehat{L}\,\vec{a} = \vec{b}$$

and

$$\widehat{M}\,\vec{b} = \vec{c}\,.$$

We can apply the operators $\widehat{L}$ and $\widehat{M}$ sequentially:

$$\widehat{M}\,(\widehat{L}\,\vec{a}) = \vec{c}\,.$$

This way we obtained a new operator $\widehat{K}$ which we call *composition of linear operators* $\widehat{L}$ and $\widehat{M}$. The same notation for composition of operators as for the composition of functions can be used:

$$\widehat{K} = \widehat{M} \circ \widehat{L}\,.$$

Now we can find the components of the operator $\widehat{K}$. On the one hand,

$$\widehat{K}\,\vec{e}_i = K_{ij}\,\vec{e}_j\,.$$

On the other,

$$\widehat{K}\,\vec{e}_i = \widehat{M}\,(\widehat{L}\,\vec{e}_i) = \widehat{M}\,(L_{iq}\,\vec{e}_q)\,.$$

Using the linearity of the operator $\widehat{M}$, we can write

$$\widehat{K}\,\vec{e}_i = L_{iq}\,(\widehat{M}\,\vec{e}_q) = L_{iq}(M_{qj}\,\vec{e}_j)\,.$$

We arrive at the following expression of the components of $\widehat{K}$:

$$K_{ij} = L_{iq} M_{qj} \,,$$

where the summation over q is implied according to Einstein's summation rule. Next let us apply this result to projectors.

Suppose we want to project a vector $\vec{c}$ first on the vector $\vec{a}$, and then project the result onto the vector $\vec{b}$. We can do it by sequential application of two projectors:

$$\underline{\underline{\widehat{A}}} = \frac{\overrightarrow{a}\overleftarrow{a}}{a^2} \quad \text{and} \quad \underline{\underline{\widehat{B}}} = \frac{\overrightarrow{b}\overleftarrow{b}}{b^2}$$

to the vector $\vec{c}$:

$$\underline{\underline{\widehat{B}}}\,(\underline{\underline{\widehat{A}}}\,\vec{c}) = (\underline{\underline{\widehat{B}}} \circ \underline{\underline{\widehat{A}}})\,\vec{c} \,.$$

Composing two linear operators we obtain another linear operator:

$$\widehat{L} = \underline{\underline{\widehat{B}}} \circ \underline{\underline{\widehat{A}}} \,.$$

The components of the product operator $\widehat{L}$ can be expressed in terms of the components of the factors – projectors $\underline{\underline{\widehat{A}}}$ and $\underline{\underline{\widehat{B}}}$:

$$L_{ij} = \underline{\underline{A}}_{ik}\underline{\underline{B}}_{kj} = \frac{a_k b_k}{a^2 b^2} a_i b_j \,.$$

The expression $a_k b_k$ is recognized as the scalar product of the vectors $\vec{a}$ and $\vec{b}$ in orthonormal basis, so the components L_{ij} are simply given by

$$L_{ij} = \lambda\, a_i b_j , \qquad \lambda = \frac{\vec{a} \cdot \vec{b}}{a^2 b^2} \,,$$

where λ is a scalar value – a number.

Note that $\widehat{L} = \underline{\underline{\widehat{B}}} \circ \underline{\underline{\widehat{A}}}$ is no longer a projector in the sense in which it was defined earlier. There is no vector $\vec{d}$ such that

$$L_{ij} = \frac{d_i d_j}{d^2} \,.$$

The following exercise explores this point.

Exercise 5.5
(a) Show that a projector

$$\underline{\underline{\widehat{A}}} = \frac{\vec{a}\,\overleftarrow{a}}{a^2}$$

has the following property:

$$\underline{\underline{\widehat{A}}} \circ \underline{\underline{\widehat{A}}} = \underline{\underline{\widehat{A}}}\,.$$

(b) Does the result of composition $\underline{\underline{\widehat{L}}} = \underline{\underline{\widehat{B}}} \circ \underline{\underline{\widehat{A}}}$ have this property?

Exercise 5.6
Consider the composition

$$\underline{\underline{\widehat{M}}} = \underline{\underline{\widehat{A}}} \circ \underline{\underline{\widehat{B}}}\,.$$

Find its components and compare them to the components of

$$\underline{\underline{\widehat{L}}} = \underline{\underline{\widehat{B}}} \circ \underline{\underline{\widehat{A}}}\,.$$

5.8 Tensor Product

We arrived at an extremely important idea that will allow "building" tensors of various kinds from simple "ingredients," such as vectors. To begin, let us take a closer look at a projector operator:

$$\underline{\underline{\widehat{A}}} = \frac{1}{a^2}\vec{a}\,\overleftarrow{a}\,, \quad \underline{\underline{A}}_{ij} = \frac{1}{a^2} a_i a_j\,.$$

From the first expression, written without the components, it is clear that the operator $\underline{\underline{\widehat{A}}}$ involves a contravariant vector $\vec{a}$ and its covariant conjugate $\overleftarrow{a}$. This distinction is absent in the second expression for the components of the operator $\underline{\underline{A}}_{ij}$, making the expression for components misleading. To fix this, a special notation for components is introduced. In this notation, the components of contravariant vectors are written using superscripts:

$$\vec{a} \quad \longrightarrow \quad a^i \,,$$

while the components of covariant vectors are written in the "usual" way, as subscripts:

$$\overleftarrow{a} \quad \longrightarrow \quad a_i \,.$$

With this in mind, the components of the projector $\underline{\underline{\widehat{A}}}$ are written in the following way:

$$\boxed{\underline{\underline{A}}^{i\,\bullet}{}_{\bullet\, j} = \frac{1}{a^2}\, a^i a_j \,,}$$

Now it should be clear that in the last expression vectors of different kinds are used: one contravariant, and the other covariant. The little grey circles surve visual purpose only, they help separate the first contravariant index from the second covariant one.

> ***Clash of Notation***
>
> The use of superscript to denote components of contravariant vectors leads to the clash of notations. For example, given
>
> $$a^2 \,,$$
>
> how should we understand it: length squared of a vector, or its second component?
>
> Surprisingly, this is not a serious problem at all, since the meaning of the superscript is usually clear from the context in which such expressions appears.

5.8.1 Tensor Product 1

In the expressions

$$\underline{\underline{\widehat{A}}} = \frac{1}{a^2}\, \vec{a}\,\overleftarrow{a} \,, \qquad \underline{\underline{A}}^{i\,\bullet}{}_{\bullet\, j} = \frac{1}{a^2}\, a^i a_j$$

the combination of vectors $\vec{a}\,\overleftarrow{a}$ and their component expression $a^i a_j$ represent a mathematical object – operator in this case – that is neither a vector, nor a number. Such "amalgamation" of two vectors into a tensor is called *tensor product*. Tensor product is a simple and versatile way to construct tensors.

Special notation for tensor product of two vectors exists:

$$\boxed{\vec{a}\,\overleftarrow{a} = \vec{a} \otimes \overleftarrow{a}\,.}$$

This separate notation may appear redundant, since we understand from the order of the vectors $\vec{a}$ and $\overleftarrow{a}$ that the expression on the left is *not a scalar product*. However, using the infix operator $\otimes$ is convenient because it allows writing other types of tensor products with ease and consistency.

5.8.2 Tensor Product 2

Using the tensor product notation, we can write

$$\overleftarrow{a} \otimes \vec{a} \quad \text{or, more generally,} \quad \overleftarrow{b} \otimes \vec{a}\,.$$

These expressions *represent tensors*, and not scalar products $\overleftarrow{b}\,\vec{a} = \vec{b} \cdot \vec{a}$. The components of a tensor $T = \overleftarrow{b} \otimes \vec{a}$ are as follows:

$$T_{i\bullet}^{\bullet j} = b_i a^j\,.$$

The important fact is reflected in the position of indices of the tensor T: it behaves like covariant vector in the first index, and as contravariant vector in the second index.

5.8.3 Tensor Product 3

With the help of the infix operator $\otimes$ we can write, without creating ambiguity, a tensor product of two contravariant vectors:

$$T = \vec{a} \otimes \vec{b}\,, \quad T^{ij} = a^i b^j\,.$$

The positions of the indices reflect the fact that this kind of tensor is contravariant in both of them.

A simplified notation is sometimes used:

$$\vec{a} \otimes \vec{b} = \vec{a}\,\vec{b}\,,$$

but it may lead to confusion, since the expression $\vec{a}\,\vec{b}$ is too similar to the scalar product $\vec{a} \cdot \vec{b}$, especially when using handwriting. We will avoid this simplified notation.

5.8.4 Tensor Product 4

The last kind of tensor that we can construct from vectors is given by the tensor product of two covariant vectors:

$$T = \overleftarrow{a} \otimes \overleftarrow{b}\,, \quad T_{ij} = a_i b_j\,.$$

Again, one may encounter expressions like $\overleftarrow{a}\,\overleftarrow{b}$, but we will prefer to use the infix operator $\otimes$.

Exercise 5.7
Write the transformation rules for tensors of all four kinds considered above.

5.9 Tensors Defined

We are in a good position to summarize our understanding of tensors. Before we do this, let's quickly review the path we took to reach this position.

Having defined contravariant and covariant vectors, we examined natural idea of operators – functions on vectors. We focused on an important class operators called linear operators.

We studied linear operators that map vectors to other vectors, like rotation, and, having examined the transformation of operator components, derived the first type of transformation (see Section 4.7). This type of transformation corresponds to the tensor of mixed kind: contravariant in the first index and covariant in the second index. Later we encountered many operators of this kind – projector operators (see Section 5.7.)

Projector operators are unary linear operators and are built on the bilinear dol-operator $\widehat{\sigma}$. This binary operator introduced to us the idea of conjugate space and unary linear operators mapping vectors into numbers.

From further analysis of projector operators, we arrived at the idea of tensor product and unlocked a key method of building tensors of various kinds. Using a pair of vectors, we listed four different kinds of tensors: covariant-covariant, covariant-contravariant, contravariant-covariant, and contravariant-contravariant. All these tensors can be viewed as operators acting on vectors, either covariant or contravariant, depending on the tensor type.

Having reviewed our steps, we can define tensors as follows:

Definition 5.1 ☞ *Tensors*
Tensors are *mathematical objects* with the following essential properties:

- Tensors can be combined (added) pairwise to yield another tensor.
- Tensors can be multiplied by real numbers to yield another tensor.
- Tensors can be represented via components, written relative to some basis.
- When basis changes, components of tensors transform in a very specific way, to ensure that the *tensor remains the same.*

This definition is deliberately analogous to the definition of vectors. In some sense tensors are *next tier vectors.* Tensors are mathematical objects following vectors in the ladder of abstraction and power, like vectors are mathematical objects following numbers in the ladder of abstraction and power.

5.9.1 Other Definitions

Let us revisit the definitions of tensors given in the introduction. The first one read:

Definition 5.2 ☞ *Tensors Definition 1*
Tensor on a vector space V over a field k is an element t of the vector space

$$T^{p,q}(V) = (\otimes^p V) \otimes (\otimes^q V^*),$$

where $V^* = \mathrm{Hom}(V, k)$ is the dual space of V.

In this definition the set of vectors V, which we denoted as $\overset{\Rightarrow}{A}$, is called *vector space.* Recall that vectors and operation with vectors require numbers. Numbers, which can be added, multiplied in a usual way, form what mathematicians call *field*. Therefore, all vectors taken together are technically called "vector space V over a field k."

Next, the use of the operator of tensor product $\otimes$ in the definition makes more sense, since it is the basic way to build tensors from vectors. Notice that two types of vectors are mentioned: one from the vector space V, and the other from its conjugate (dual) – vector space V^*; in our notation $V^* = \overset{\Leftarrow}{A}$. The conjugate vectors behave very much like vectors from the original vector space: they can be added, multiplied by numbers, expanded in their own basis, and so on. This close correspondence between vector space and its conjugate is called *homomorphism*[3] and denoted $\mathrm{Hom}(V, k)$.

Tensor of the kind $T^{p,q}(V) = (\otimes^p V) \otimes (\otimes^q V^*)$ has p contravariant indices and q covariant indices. We mostly dealt with tensors with the following types: $T^{1,1}$, $T^{0,2}$, $T^{2,0}$.

Finally, tensors are vectors because they can be added, multiplied by numbers, have components, and have vector-like transformation rules. Tensors of a given type, e.g. $T^{2,1}$, taken together, form a vector space; there is a separate vector space for each type of tensor.

The second defintion stated:

Definition 5.3 ***Tensors Definition 2***
An nth-rank tensor in m-dimensional space is a mathematical object that has n indices and m^n components and obeys certain transformation rules.

Here tensor is mentioned in connection with some m-dimensional space, which we recognize as the underlying vector space $\overset{\Rightarrow}{A}$. In our case the number of dimensions, given by the number of independent directions, equals 2.

The rank of a tensor is given by the number of indices, or the number of vectors that go into the tensor product. For example:

$$L = \overleftarrow{a} \otimes \overleftarrow{b} \longrightarrow L_{ij} \quad \text{is the tensor of the second rank,}$$

$$M = \overleftarrow{a} \otimes \overleftarrow{b} \otimes \overrightarrow{c} \longrightarrow M_{i\;j\;\bullet}^{\bullet\;\bullet\;k} \quad \text{is the tensor of the third rank.}$$

Since the value of each index runs from 1 to m, where m is the dimension of the vector space, tensors of the rank n should have m^n components in total. The essential property of tensor components is how they are transformed when the basis changes – the tensor transformation rule.

---

[3]From Greek *homos* (same) and *morphe* (shape).

The last definition, given in the introduction, looks as follows:

Definition 5.4 ☞ ***Tensors Definition 3***
Just as a *vector* is a mathematical quantity that describes translations in two- or three-dimensional space, a tensor is a mathematical quantity used to describe general transformations in n-dimensional space. Precisely, if the locations of points in n-dimensional space are given in one coordinate system by $(x^1, x^2, \ldots, x^n)$ and in a transformed coordinate system by $(y^1, y^2, \ldots, y^n)$ (it is convenient to use superscripts rather than subscripts), then a "rank 1 contravariant tensor" is a quantity T, with single components, that transforms according to the rule:

$$T^i_{new} = \sum_{r=1}^{n} \frac{\partial y^i}{\partial x^r} T^r \ .$$

We now know that tensors can be used to describe transformation of one vector into another (e.g. projectors). Similar ideas can be applied to vectors in spaces with higher dimensions. In this sense, tensors can describe general transformaions in n-dimensional space.

The essential property of tensor components is their tranformation rule. In this definition a transformation rule for a tensor of the first rank (a usual vector) is given:

$$T'^i = Z^i{}_r T^r \ ,$$

where the set of numbers $Z^i{}_r$ describes the relation between the "old" basis (coordinates x) and the "new" basis (coordinates y).

Chapter Highlights

- *Two vectors can be compared for similarity by calculating the "degree of overlap". The longer two vectors are and the closer their mutual direction – the greater the overlap is.*
- *Degree of overlap can be described by a binary linear operator $\widehat{\sigma}$. This operator is closely related to the concept of scalar product of two vectors.*

- *When scalar product (or, equivalently, degree of overlap) is defined for vectors, each vector receives a "special relative" – conjugate vector – that lives in different vector space, called conjugate or dual space.*
- *When the degree-of-overlap operator $\widehat{\sigma}$ is partially applied, the result is a unary linear operator that yields a number for every input vector. Importantly, such an operator is also a vector, albeit not an arrow-like vector.*
- *Unary linear operators that act on arrow-like vectors themselves form a space of vectors. The latter vectors are conjugate to the "usual" space of arrow-like vectors.*
- *If arrow-like vectors are contravariant vectors, then their conjugate counterparts are covariant vectors. Covariant vectors can be represented geometrically as oriented area elements (for three dimensional space).*
- *Projectors are unary linear operators that act on input arrows to yield another arrow that is parallel to a certain direction. Projectors are degenerate operators.*
- *Projectors can be written using the efficient tensor product notation.*
- *Tensor product is a simple and powerful way to build up tensors of any rank and kind from a number of covariant and contravariant vectors.*
- *Tensors are mathematical objects with many properties similar to vectors. However, their rank is higher and tensors can be used to express linear relations between vectors.*

6. Applications

We are now ready to appreciate how tensors are used in “real life”. In this chapter we will encounter examples of tensors that are used in mathematics, physics, and engineering. Before we get to the examples of tensors, one more helpful notation must be explained.

δ-Notation

When a quantity x changes by a tiny amount, we will denote the change using small Greek letter δ (delta) as follows:

$$\boxed{\delta x \text{ - tiny change of } x.}$$

For example, for the earth going around the sun in 365 days, one second elapsed on a clock can be considered a tiny change δt. When a drop of water falls into a nearly full bucket the mass of the latter changes by a tiny amount δm, and so on.

That’s all there is to δ-notation. We are not going into the realm of calculus, where mathematicians talk about infinitesimal quantities and limits; we will be simply using “tiny changes.” Now on to tensors.

6.1 Famous Tensors

We will study several examples of tensors that readers most likely encounter in geometry, physics, and engineering. The material in the previous chapters should be enough to prepare a reader to deal with tensors of any kind. However, we will limit considerations to simple tensors of lower ranks.

6.1.1 Metric Tensor

Metric tensor is used to determine distances between pairs of points in space. A distance between two points is equal to the length of a vector connecting them, as shown in the Figure ??. For a vector

$$\vec{d} = \vec{b} - \vec{a}\,, \quad d^i = b^i - a^i$$

its length squared is given by the scalar product

$$d^2 = \vec{d} \cdot \vec{d} = \widehat{\sigma}\,\vec{d}\,\vec{d}\,.$$

Using components in arbitrary basis (not orthonormal), the length squared is written as

$$d^2 = \widehat{\sigma}\,(d^i \vec{e}_i)\,(d^j \vec{e}_j) = d^i d^j\,(\widehat{\sigma}\,\vec{e}_i\,\vec{e}_j)\,.$$

The set of values

$$\eta_{ij} = \widehat{\sigma}\,\vec{e}_i\,\vec{e}_j$$

corresponds to the components of a special tensor – *metric tensor*. The transformation rule of these components is easily found by expanding "old" basis vectors in terms of the "new" (primed) basis, and using the linearity of $\widehat{\sigma}$:

$$\eta'_{ij} = \widehat{\sigma}\,\vec{e}\,'_i\,\vec{e}\,'_j = \widehat{\sigma}\,(E_{im}\vec{e}_m)\,(E_{jn}\vec{e}_n) = E_{im}E_{jn}\eta_{mn}\,.$$

This is the transformation rule of a covariant-covariant tensor of the second rank. The metric tensor has to be of this kind since it maps a contravariant-contravariant tensor

$$\vec{d} \otimes \vec{d}\,, \quad d^i d^j$$

into a scalar. Each index of the metric tensor must transform in a way that "compensates" the contravariant transformation of $\vec{d}$ in the tensor product $\vec{d} \otimes \vec{d}$. This is analogous to how a covariant vector $\overleftarrow{b}$ maps a contravariant vector $\vec{a}$ into a number:

$$\overleftarrow{b}\,\vec{a} \longrightarrow x$$

$$\eta\,(\vec{a} \otimes \vec{a}) \longrightarrow y\,.$$

Exercise 6.1
Although the primary use of the metric tensor is to calculate distances between a pair of points connected by a vector $\vec{d}$, it can be applied to any pair of contravariant vectors:

$$\eta\left(\vec{a} \otimes \vec{b}\right).$$

a) What is the meaning of this operation? b) What are the components of the metric tensor in orthonormal basis?

In most elementary problems of geometry and physics, a coordinate system in a plane is Cartesian and basis vectors are orthonormal. As the result, the components of the metric tensor are trivial – zeros and ones – and are the same everywhere in the plane.

When non-Cartesian coordinates are used in a plane, e.g. polar coordinates shown in the Figure **??**, the basis vectors are aligned with the coordinate grid[1] and have different orientation in different points. In this case the components of the metric tensor η_{ij} change from point to point to ensure that the lengths of the vector $\vec{d}$

$$d^2 = \eta_{ij} d^i d^j$$

remains the same.

Moreover, for surfaces more sophisticated than a plane (e.g., sphere, paraboloid, saddle-like surface, and myriad of others), it is impossible to use coordinate system and basis such that the metric tensor is constant. The components of the metric tensor will vary across the surface to reflect real, and not a merely "coordinate induced", difference of a surface from a plane. In other words, variation of the components of the metric tensor indicate that the surface is *curved*.

As an example, consider two dimensional surface of a sphere, shown in the Figure **??**. Each point on the surface can be located using a pair of coordinates – the angle $\theta = x^1$ complimentary to the latitude, and the longitude angle $\phi = x^2$. If a pair of close points on the sphere have

[1]In principle, it is possible to have basis vectors "decoupled" from the coordinate system, but this is not very convenient.

coordinates

$$\begin{array}{ccc} 1 & \longrightarrow & (\theta, \phi), \\ 2 & \longrightarrow & (\theta + \delta\theta, \phi + \delta\phi), \end{array}$$

then the distance squared between these points is given by

$$d^2 = R^2(\delta\theta)^2 + R^2 \sin^2\theta(\delta\phi)^2.$$

This formula is obtained by applying Pythagoras theorem to the tiny right triangle with the sides indicated using arrows in the Figure **??**. The length of the side resulting from the change of the coordinate ϕ is $r\delta\phi = R\sin\theta\delta\phi$; the length of the side resulting from the change of the coordinate θ equals $R\delta\theta$.

Using the uniform notation for coordinates, the distance squared is written as

$$d^2 = R^2(\delta x^1)^2 + R^2 \sin^2 x^1 (\delta x^2)^2 \,.$$

Comparing this to the Cartesian expression $d^2 = (\delta x^1)^2 + (\delta x^2)^2$, we can see that not all components of the metric tensor in the spherical coordinate basis are constant. Namely, the component $\eta_{22} = R^2 \sin^2 x^1$ depends on the coordinate $x^1 = \theta$.

6.1.2 Anisotropy Tensor

Anisotropy tensor is a general term for various tensors used in physics to describe properties of materials like crystalline solids. Many physical properties – mechanical, optical, electronic, thermal – describe the response of material to the external "forces" or perturbations. Mathematical description of such responses requires tensors.

To understand the general idea, let us consider a simple situation, depicted in the Figure **??**. Suppose that a tree bends in the wind, so that when the wind blows in the direction of the x axis, the displacement of the tree-top is also along the x axis, with the magnitude proportional to the magnitude of the wind's velocity $\vec{v}$:

$$\vec{d} = d_x \vec{u}_1 = A v_x \vec{u}_1 = A\vec{v} \,.$$

Next, suppose that when the wind blows along the y axis, the tree-top is also displaced in the direction of the y-axis:

$$\vec{d} = d_y\vec{u}_2 = Bv_y\vec{u}_2 = B\vec{v}\,.$$

In both cases the displacement vector $\vec{d}$ is parallel to the vector of wind's velocity.

For a general direction of the wind, the magnitude of the tree-top displacement will be proportional to the magnitude of the wind's velocity, but *the direction of the displacement will differ from the direction of the wind*:

$$d \propto v\,, \quad \vec{d} \nparallel \vec{v}\,.$$

Indeed, for a wind vector

$$\vec{v} = v_x\vec{u}_1 + v_y\vec{u}_2 = \vec{v}_1 + \vec{v}_2\,,$$

the "response" of the tree-top will be different for different components of the wind vector:

$$\vec{d} = \vec{d}_1 + \vec{d}_2 = A\vec{v}_1 + B\vec{v}_2 = Av\cos\theta\vec{u}_1 + Bv\sin\theta\vec{u}_2 = d_x\vec{u}_1 + d_x\vec{u}_2\,.$$

Clearly,

$$\frac{d_y}{d_x} = \frac{B\sin\theta}{A\cos\theta} \neq \tan\theta\,.$$

Thus, although the displacement magnitude is still proportional to the magnitude of the wind's velocity, the direction of the displacement no longer coincides with the direction of the wind. This fact can be expressed using tensor notation:

$$d^i = T^i{}_j v^j\,. \tag{6.1}$$

In the special coordinates, considered at the beginning of this problem, the components of the tensor T are simple:

$$T^1{}_1 = A\;, T^1{}_2 = 0\;, T^2{}_1 = 0\;, T^2{}_2 = B\;.$$

In all other coordinate systems and bases, the components of the "response tensor" T can be found using the transformation rule for the contravariant-covariant tensor of the second rank.

Expressions similar to the equation (6.1) can be written for a variety of physical phenomena. We will consider a couple of examples next.

Mechanics: Stress Tensor

Mechanics of elastic media uses many tensor tools. One of the basic tensors is the *stress tensor.*[2] This tensor describes the distribution of mechanical stress inside a deformed elastic body, as illustrated in the Figure **??**.

An elastic ball, shown in the Figure **??**(a), can be squeezed by external forces, resulting in the change of shape (*deformation*) and the appearance of a mechanical stress inside the ball; see Figure **??**(b). In general, the induced stress will change from point to point inside the deformed ball. To describe the stress in a given point P, we can imagine that a small part of the body is removed, leaving a tiny square-shaped hole. If nothing is done, the empty part of the ball around the point P will not remain square, due to the "forces" acting within the body and at the boundary of the hole. To keep the hole square, we must compensate the forces due to mechanical stress and apply the balancing forces $\vec{f}_1$, $\vec{f}_2$, $\vec{f}_3$, $\vec{f}_4$ to each side of the square[3]. Only two such forces are shown in the Figure **??**(b) for simplicity.

The direction and magnitude of a force needed for a given side can be found as follows: First, find the unit-length vector $\vec{e}_i$ perpendicular to the side. Second, calculate the force using the Cauchy stress tensor:

$$\vec{f}_i = \widehat{\sigma}\,\vec{e}_i \,.$$

The traditional notation for mechanical stress tensors is Greek letter sigma $-\sigma$. It should not be confused with out notation for dol-operator defined earlier (see Section 5.1).

It is easy to understand why the "balancing forces" are not, in general, pushing perpendicular to the sides of the square. The idea is illustrated in the Figure **??**(c). An elastic square (cube) can be deformed in two basic ways: 1) A square can be squeezed by forces perpendicular to the sides (*normal stress*); 2) A square can be deformed into a parallelogram by forces parallel to the sides (*shear stress*). Both types of stress can exist at the same time, resulting from forces directed at an arbitrary angle relative to the vector $\vec{e}_i$ perpendicular to the sides.

Electronics: Mobility Tensor

[2]Also known as *Cauchy stress tensor* or *true stress tensor.*

[3]For a three dimensional ball the shape of the hole will be a cube, and the number of forces will be 8 – one for each face of the cube.

Many materials conduct electric current. An important characteristic of such a material is its *resistance*. When a voltage V is applied to a piece of conducting material, the current I will flow between the terminals, as shown in the Figure **??**.

The basic law that relates the voltage V and the current I between the terminals is *Ohm's law*:

$$V = IR\,.$$

Here R is the electric resistance of a given piece of material.

For the same material, currents flowing in different directions may experience different resistances even for the same geometrical shape. In the example shown in the Figure **??**(a,b), for currents flowing horizontally and vertically through a square we can write

$$V = I_1 R_1 \text{ and } V = I_2 R_2\,.$$

We can take another view on the movement of electric charge through the material if we rewrite Ohm's law as follows:

$$I = GV\,.$$

Here instead of resistance we use an equally useful physical parameter called *conductance* G. In certain sense, conductance is more fundamental since it is closely related to basic physical laws that govern the motion of electric charges.

Electric current is the flow of a large number of charge carriers, such as electrons or ions, shown as red dots in the Figure **??**(a,b). The current I is proportional to the average speed u of the carriers through the material:

$$I \propto u\,.$$

The carriers, in their turn, move because there is an electric field E between the terminals due to applied voltage V. The average speed u of charge carriers is often simply proportional to the electric field:

$$u = \mu E\,,$$

where the coefficient μ is called *mobility* of charge carriers.

Now for anisotropic materials, the relation between the average velocity $\vec{u}$ of charge carriers and the applied electric field $\vec{E}$ can be written using the concept of *mobility tensor*:

$$\vec{u} = \widehat{\mu}\,\vec{E}\,.$$

The mobility tensor expresses how easy it is to make electrons move in a given direction by applying an external electric field $\vec{E}$.

Let us summarize: Applying voltage between terminals creates an electric field $\vec{E}$ in a given direction. The electric field is proportional to the voltage between the terminals: $E = V/d$. The electric field leads to the "mass migration" of charge carriers with the average speed

$$u = \mu E = \mu V/d\,.$$

This type of motion is called electric current:

$$I \propto u \quad\longrightarrow\quad I \propto \frac{\mu}{d} V\,.$$

From the last expression we can see how the relationship $I = GV$ or Ohm's law $V = IR$ appear. Furthermore, because the mobility $\widehat{\mu}$ is in general a tensor, the measured resistance of a given piece of material may be different for different direction of applied voltage drop V.

Anisotropy Tensors in Physics

Besides two examples of tensors (stress and mobility) given above, there are many other tensors used in physics. Some tensors are similar to stress and mobility tensors in the sense that they express linear relationship between "action" ($\vec{a}$) and "response" ($\vec{r}$) vectors

$$\vec{r} = \widehat{t}\,\vec{a}\,.$$

But more advanced tensors are also used to express linear relationships between more simple tensors and vectors. For example, in certain materials mechanical stress can lead to separation of electric charges and thus create voltage drop between different points of the body. This phenomenon is known as *piezoelectric effect*. Now if we characterize induced charge separation using a vector $\vec{p} = p^i$, then we can write using index notation

$$p^i = d^{ijk}\sigma_{jk} \,,$$

where σ_{jk} are the components of the stress tensor described above, and d^{ijk} – piezoelectric tensor of the third rank (three indices!)

For the reader interested in more examples and details, the book *"Physical Properties of Crystals: Their Representation by Tensors and Matrices"* by J. F. Ney is highly recommended.

As the last example of tensors in physics, we will consider a more fundamental case from field theory.

6.1.3 Electromagnetic Tensor

In applied physics and engineering one works with electric and magnetic fields that are described using two *different* physical vector quantities: $\vec{E}$ – for electric field strength, and $\vec{B}$ – for magnetic field strength.

When a charged particle, say an electron, is placed in electric field, the latter acts on that particle with force proportional to the field strength:

$$F_e = qE \,,$$

where q is the charge of the particle, F_e denotes the force due to the electric field $\vec{E}$.

When the same charged particle is *moving* in a magnetic field, the latter acts on the particle with the force proportional to the magnetic field strength:

$$F_m = qvB \,,$$

where v is the speed of the charge particle, F_m denotes the force due to the magnetic field $\vec{B}$. The difference between the effects of electric and magnetic fields on a charged particle is illustrated in the Figure **??**.

The distinction between electric and magnetic fields is technical and *not fundamental*. Figuratively speaking, electric field differs from magnetic field to the same degree as rest differs from uniform motion. Electric and magnetic fields are different aspects of the same physical entity – *electromagnetic field*.

From the expressions for the electric and magnetic forces F_e and F_m we can see that physical quantities E and B have different units of measurement – the fact which upsets some physicists. They note: If electric and magnetic fiels are different aspects of the *same physics object*, they must be measured using the same units, similar to how we measure height and width of a building using the same units of length.

The way to fix the issue with different units for electric and magnetic fields, is to change the way we measure...*velocity*! Nature provides us with a special standard of speed – the speed of light in vacuum, denoted as c. The speed of light in vacuum is a “nature-made” absolute quantity, in contrast to such human-made standards as units of length (meter) or time (second). This is why in fundamental physical theories, including the theory of electromagnetic field, it is wise to specify all speeds as fractions of the speed of light.

Thus, in physical formulas, instead of writing v as meters per second, we should use a “normalized” quantity:

$$\bar{v} = v/c \quad \rightarrow \quad v = \bar{v}c\,.$$

Once we apply this approach to the expression of the magnetic force, we obtain

$$F_m = qvB = q\bar{v}(cB)\,.$$

Now we can see that the quantities E and cB have the same physical meaning – *the force per unit charge*. It is these physical quantities that should be used to describe different aspects fo the same electromagnetic field. We will denote them as follows:

$$\vec{E} = \vec{\mathcal{E}} = \mathcal{E}^i\,, \quad c\vec{B} = \vec{\mathcal{B}} = \mathcal{B}^j\,.$$

Electromagnetic Tensor

A deep and beautiful discovery of the theory of electromagnetic field can now be stated: Electric and magnetic fields $\mathcal{E}^i$ and $\mathcal{B}^j$ are not separate *vector* quantities, they are, in fact, represent certain *components of a tensor* that describes electromagnetic field.

This tensor is conventionally written as $F^{\mu\nu}$ (F here stands for *field*, not force!) $F^{\mu\nu}$ is a second rank tensor. The indices μ and ν run from 0 to 3.

The relationships between the “usual” electric field and the electromagnetic tensor are given by

$$\mathcal{E}^i = F^{i0}\,, \quad i = 1, 2, 3\,.$$

The relationships between the "usual" magnetic field and the electromagnetic tensor can be written in the following way:

$$\mathcal{B}^1 = F^{32}\,, \mathcal{B}^2 = F^{13}\,, \mathcal{B}^3 = F^{21}\,.$$

A convenient way to write all components of a second rank tensor is to use table-like structure called *matrix*:

$$F^{\mu\nu} = \begin{pmatrix} F^{00} & F^{01} & F^{02} & F^{03} \\ F^{10} & F^{11} & F^{12} & F^{13} \\ F^{20} & F^{21} & F^{22} & F^{23} \\ F^{30} & F^{31} & F^{32} & F^{33} \end{pmatrix}.$$

In the matrix, the first index μ of $F^{\mu\nu}$ corresponds to the row, while the second index ν corresponds to the column. Both rows and columns are enumerated from 0 to 3.

Using matrix form, we can write the electromagnetic tensor in terms of the electric and magnetic fields:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -\mathcal{E}^1 & -\mathcal{E}^2 & -\mathcal{E}^3 \\ \mathcal{E}^1 & 0 & -\mathcal{B}^3 & \mathcal{B}^2 \\ \mathcal{E}^2 & \mathcal{B}^3 & 0 & -\mathcal{B}^1 \\ \mathcal{E}^3 & -\mathcal{B}^2 & \mathcal{B}^1 & 0 \end{pmatrix}.$$

The last expression makes apparent two features of electromagnetic tensor components. First, all diagonal elements vanish:

$$F^{00} = F^{11} = F^{22} = F^{33} = 0\,.$$

Second,

$$F^{\mu\nu} = -F^{\nu\mu}\,,$$

a property known as *antisymmetry*. This property requires all diagonal elements to be equal zero.

Electromagnetic Tensor Components

The first kind of tensor of the second rank that we encountered was a linear operator $\widehat{L}$. The components of any linear operator are given relative to some basis and the components specify how the operator transforms basis vectors:

$$\widehat{L}\,\vec{e}_i = L_{ij}\vec{e}_j \,.$$

What is the basis used to express components of electromagnetic tensor $F^{\mu\nu}$?

Electromagnetic tensor is a physical operator which is used to express the action of electromagnetic field on a moving charged particle. Components of electromagnetic tensor connect special versions of velocity (v_ν) and force (f^μ) acting on a charged particle:

$$f^\mu = qF^{\mu\nu}v_\nu \,.$$

Without going into details, we will note that in the left-hand side of this equation we have a four-component force, and on the right-hand side we have both electric and magnetic effects combined in a single tensor.

Chapter Highlights

- *Tensors find application in various areas of science and math.*
- *Geometrical properties of surfaces and spaces can be described using metric tensor.*
- *Physical properties of solids are often anisotropic – depend on the direction of applied "force". Such properties are best described by various tensors: stress tensor, mobility tensor, piezoelectric tensor, and others.*
- *At the fundamental level electric and magnetic fields are united in a single physical object – electromagnetic field. Electromagnetic field is described by an antisymmetric tensor of the second rank.*
- *The concept of linear operators, and in particular of the rotation operator $\widehat{J}$, can be used to extend the numbers from a number line to the number plane*

and arrive at complex numbers (or compound numbers, as we called them).

- *Operators and compound numbers are used in many physical theories, and play an especially important role in Hamiltonian dynamics and quantum mechanics.*

7. Implications

Exercise 1.1

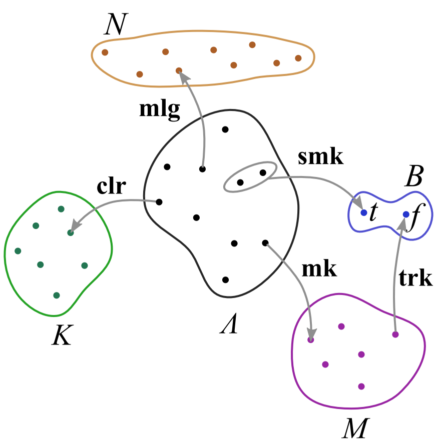


Fig. 7.1: The set M contains all possible makes of cars: Ford, Toyota, etc.

The diagram in the Figure 9.1 shows the set M – the set of all possible makes of cars. A mapping **trk** returns $true$ if a given car maker produces trucks.

Exercise 2.1

Any binary function can be viewed as a unary function if two inputs are replaced by a single input of a *pair of numbers*. Similarly for a function with two outputs. This idea is illustrated in the Figure 9.2(a): The function **swp** is viewed as a unary function which swaps the numbers in an *ordered pair*:

$$\mathbf{swp}\ (n, m) = (m, n)\,.$$

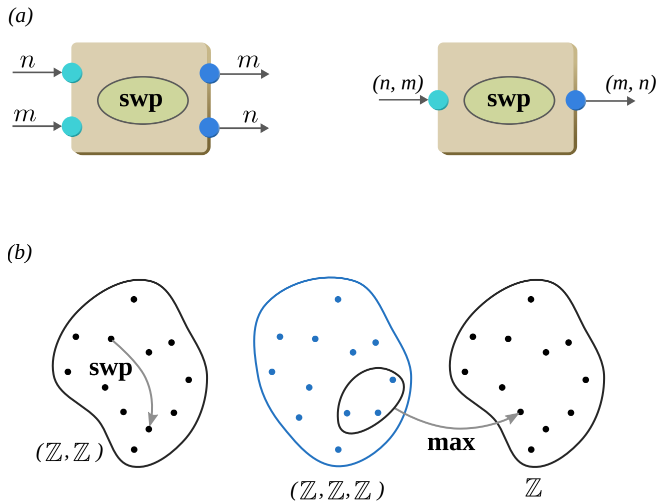


Fig. 7.2: (a) Two inputs (outputs) of a function can be replaced with a single input of a ***pair*** of numbers, turning a binary function into a unary one. (b) That.

Given the set $\mathbb{Z}$ of whole numbers, we can create the set of all possible *ordered pairs* (n, m). This set can be denoted as follows:

$$(\mathbb{Z}, \mathbb{Z}) \;\text{ or }\; \mathbb{Z} \times \mathbb{Z}\,.$$

The latter notation is standard in mathematics, but the former way of writing is also acceptable. We can similarly denote the set of all *ordered triples*:

$$(\mathbb{Z}, \mathbb{Z}, \mathbb{Z}) \;\text{ or }\; \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}\,.$$

With the notation introduced above, the action of functions with multiple inputs or outputs can be depicted on the level of sets. The Figure 9.2(b) shows how this works for the functions **swp** and **max** .

Exercise 2.2

Consider a binary function that accepts a pair of natural numbers and returns the third natural number in the following way:

$$\mathbf{rep}\;\; 3\,2 = 33 \qquad \mathbf{rep}\;\; 1\,4 = 1111\,.$$

Thus, the output is a natural number with identical digits given by the first number, repeated a number of times specified by the second number.

Infix variant of this operation can be written, rather arbitrarily, like this:

$$\mathbf{rep}\ n\,m = n \succ m = nnn\ldots n\,.$$

Exercise ??

A linear function f must satisfy the linearity condition

$$f\,(a * n) = a * (f\,n)\,.$$

For $a = 0$ we must have

$$f\,(0 * n) = 0 * (f\,n)\,,$$

or, equivalently

$$f\,0 = 0\,.$$

Also, for $a = m$ and $n = 1$ we must have

$$f\,(m * 1) = m * (f\,1)\,,$$

from which follows

$$f\,m = m(f\,1)\,.$$

Exercise ??

The schematics in the Figure 8.3(a) and (b) demonstrate the two linearity requirements.

Exercise ??

Using Einstein's summation rule, a polynomial of degree n can be written as follows:

$$P_n\,x = a_i x^i\,, \quad i = 0, 1, 2\ldots, n\,.$$

Exercise ??

The expression

$$b_i y_i\,, \quad i = 1, 2, 3, 4$$

represents the sum of four terms:

$$S = b_1 y_1 + b_2 y_2 + b_3 y_3 + b_4 y_4\,.$$

Similarly,

$$b_j y_j\,, \quad j = 1, 2, 3, 4$$

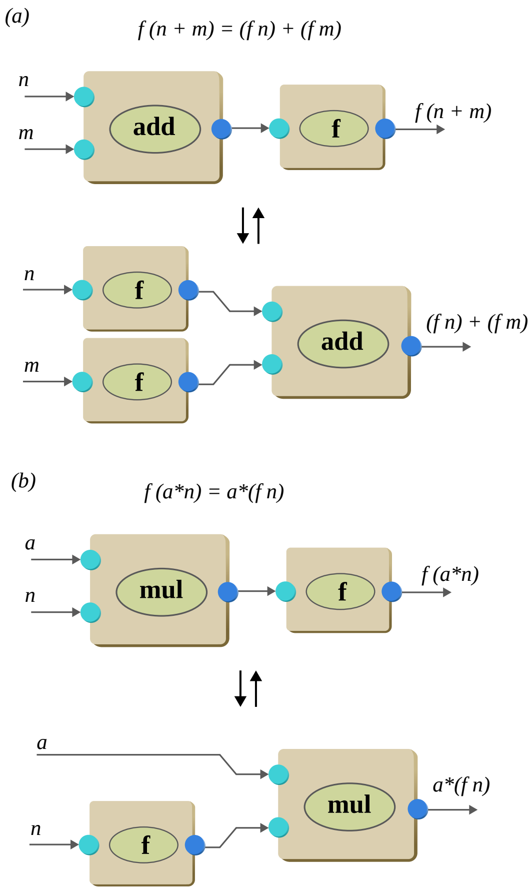


Fig. 7.3: The linearity conditions can be represented schematically with different relative configurations (order) of the "boxes".

stands for the same sum S, just as the expression

$$b_k y_k = b_1 y_1 + b_2 y_2 + b_3 y_3 + b_4 y_4 \,.$$

Exercise ??

In the expression

$$(a_i x_i)(a_j x_j)$$

both parentheses contain the identical sum:

$$a_i x_i = a_j x_j = a_1 x_1 + a_2 x_2 \,.$$

Opening the parentheses, we obtain

$$(a_i x_i)(a_j x_j) = a_1^2 x_1^2 + a_2^2 x_2^2 + 2 a_1 a_2 x_1 x_2 \,.$$

In contrast, the expression $a_i^2 x_i^2$ stands for

$$a_i^2 x_i^2 = a_1^2 x_1^2 + a_2^2 x_2^2 \,.$$

Clearly,

$$a_i^2 x_i^2 \neq (a_i x_i)^2 \,.$$

Exercise ??

The left-hand side of the expression

$$(a_i x_i)^2 = \frac{b_j y_j}{c_k c_k}$$

can be written as

$$(a_i x_i)^2 = \left(\sum_{i=1}^{i=N} a_i x_i \right)^2$$

The right-hand side takes the form:

$$\frac{b_j y_j}{c_k c_k} = \frac{\sum\limits_{j=1}^{j=N} b_j y_j}{\sum\limits_{k=1}^{k=N} c_k c_k} \,.$$

Therefore, the original equality can be re-written using the traditional summation sign:

$$\left(\sum_{i=1}^{i=N} a_i x_i\right)^2 = \frac{\sum\limits_{j=1}^{j=N} b_j y_j}{\sum\limits_{k=1}^{k=N} c_k c_k} .$$

Already one can see the advantage of Einstein's summation rule.

Exercise ??

The expression

$$\delta_{1i} a_i \,, \quad i = 1, 2, 3, 4$$

represents the sum

$$\delta_{11} a_1 + \delta_{12} a_2 + \delta_{13} a_3 + \delta_{14} a_4 \,.$$

The only non-zero term corresponds to $\delta_{11} = 1$, therefore

$$\delta_{1i} a_i = a_1 \,.$$

Similarly, we have

$$\delta_{3k} a_k \,, \quad k = 1, 2, 3, 4$$

representing

$$\delta_{31} a_1 + \delta_{32} a_2 + \delta_{33} a_3 + \delta_{34} a_4 = \delta_{33} a_3 \,.$$

Consequently,

$$\delta_{3k} a_k = a_3 \,.$$

Next,

$$\epsilon_{1j} a_j = \epsilon_{11} a_1 + \epsilon_{12} a_2 + \epsilon_{13} a_3 + \epsilon_{14} a_4$$

can be simplified to

$$\epsilon_{1j} a_j = a_2 + a_3 + a_4$$

using the definition of ϵ_{ij}.

Finally, the expression

$$\epsilon_{3j} a_j = \epsilon_{31} a_1 + \epsilon_{32} a_2 + \epsilon_{33} a_3 + \epsilon_{34} a_4$$

is reduced to

$$\epsilon_{3j}a_j = a_4 - a_1 - a_2 \,.$$

Exercise ??

The sum

$$a_i + a_j$$

can be rewritten using the facts $a_j = \delta_{ji}a_i$ and $\delta_{ij} = \delta_{ji}$:

$$a_i + a_j = a_i + \delta_{ij}a_i = (1 + \delta_{ij})a_i = (1 + \delta_{ji})a_i \,.$$

Exercise ??

(a) The expression $\delta_{ij}a_ib_j$ can be simplified using the fact $\delta_{ij}b_j = b_i$:

$$\delta_{ij}a_ib_j = a_ib_i = a_1b_1 + a_2b_2 \,.$$

(b) Fully writing out $\epsilon_{ij}a_ib_j$ results in

$$\epsilon_{11}a_1b_1 + \epsilon_{12}a_1b_2 + \epsilon_{21}a_2b_1 + \epsilon_{22}a_2b_2 \,.$$

From the definition of ϵ_{ij} follows that only the terms with $i \neq j$ survive:

$$\epsilon_{ij}a_ib_j = a_1b_2 - a_2b_1 \,.$$

Exercise ??

(a) Firstly, we can recall that when δ_{ij} is summed with a vector a_j it simply "renames" the index that is being used for summation:

$$\delta_{ij}a_j = a_i \,.$$

Using this property, we immediately get

$$\delta_{ij}\delta_{jk} = \delta_{ik} \,.$$

Another – and longer – way to get this result is to write out the summation fully:

$$\delta_{ij}\delta_{jk} = \delta_{i1}\delta_{1k} + \delta_{i2}\delta_{2k} + \ldots + \delta_{in}\delta_{nk} \,.$$

If $i \neq k$, all terms on the right are zero. Indeed, $\delta_{i1}\delta_{1k}$ is zero unless $i = 1$ and $k = 1$; similarly, $\delta_{i2}\delta_{2k}$ is zero unless $i = 2$ and $k = 2$ and so on. Therefore, the only non-zero value for $\delta_{ij}\delta_{jk}$ is when $i = k$. Let $i = k = m$, then in the sum

$$\delta_{m1}\delta_{1m} + \delta_{m2}\delta_{2m} + \ldots + \delta_{mm}\delta_{mm} + \ldots + \delta_{mn}\delta_{nm} \,.$$

there is only one non-zero term, namely

$$\delta_{mm}\delta_{mm} = 1 \cdot 1 = 1 \,.$$

Summarizing the above arguments, we conclude that

$$\delta_{ij}\delta_{jk} = 1 \text{ if } i = k \text{ and } 0 \text{ otherwise } \,.$$

This is equivalent to the expression

$$\delta_{ij}\delta_{jk} = \delta_{ik} \,.$$

(b) The expression $\epsilon_{ij}\epsilon_{jk}$, when fully expanded as a sum, takes the form

$$\epsilon_{ij}\epsilon_{jk} = \epsilon_{i1}\epsilon_{1k} + \epsilon_{i2}\epsilon_{2k} \,.$$

If $i = k = 1$, the sum is reduced to

$$\epsilon_{1j}\epsilon_{j1} = \epsilon_{11}\epsilon_{11} + \epsilon_{12}\epsilon_{21} = 1 \cdot (-1) = -1 \,.$$

Similarly, for $i = k = 2$ we get

$$\epsilon_{2j}\epsilon_{j2} = \epsilon_{21}\epsilon_{12} + \epsilon_{22}\epsilon_{22} = (-1) \cdot 1 = -1 \,.$$

On the other hand, if $i = 1$ and $k = 2$, we obtain

$$\epsilon_{1j}\epsilon_{j2} = \epsilon_{11}\epsilon_{12} + \epsilon_{12}\epsilon_{22} = 0 \,.$$

Same for $i = 2$ and $k = 1$:

$$\epsilon_{2j}\epsilon_{j1} = \epsilon_{21}\epsilon_{11} + \epsilon_{22}\epsilon_{21} = 0 \,.$$

We thus manually checked all cases and showed that

$$\epsilon_{ij}\epsilon_{jk} = -\delta_{ik} \quad i, j, k = 1, 2 \,.$$

Exercise 2.3

Let us denote:

$$x = \epsilon_{ij}a_i a_j \,.$$

Since we can rename the summation indices, we can write

$$\epsilon_{ij} a_i a_j = \epsilon_{ik} a_i a_k = \epsilon_{jk} a_j a_k = \epsilon_{ji} a_j a_i \,.$$

Now we have $\epsilon_{ji} = -\epsilon_{ij}$ and this leads to

$$\epsilon_{ji} a_j a_i = -\epsilon_{ij} a_i a_j \,.$$

We thus showed that $x = -x$ and therefore $x = 0$.

Exercise ??

An expansion of an arbitrary vector $\vec{a}$ in terms of the basis vectors is given by

$$\vec{a} = a_1 \vec{e}_1 + a_2 \vec{e}_2 + a_3 \vec{e}_3 + \ldots a_n \vec{e}_n \,.$$

This can be compactly written using Einstein's summation rule:

$$\vec{a} = a_i \vec{e}_i \quad i = 1, 2, \ldots, n \,.$$

If the number of basis vectors is known and fixed, as is usually the case, we can omit the range of the summation index and simply write

$$\vec{a} = a_i \vec{e}_i \,.$$

Exercise ??

The expression

$$\vec{e}\,'_1 = E_{11} \vec{e}_1 + E_{12} \vec{e}_2 \,,$$

can be written using Einstein's summation rule as follows:

$$\vec{e}\,'_1 = E_{1j} \vec{e}_j \,.$$

Similarly for the second basis vector:

$$\vec{e}\,'_2 = E_{2j} \vec{e}_j \,.$$

Combining both results, we obtain

$$\vec{e}\,'_i = E_{ij} \vec{e}_j \,.$$

Exercise ??

(a) Writing the expansion of the "new" basis as follows:

$$\vec{e}'_1 = \mu\vec{e}_1 + 0\vec{e}_2\,,$$

$$\vec{e}'_2 = 0\vec{e}_1 + \nu\vec{e}_2\,,$$

we can immediately find the components E_{ij}:

$$E_{11} = \mu,\ E_{12} = 0,\ E_{21} = 0,\ E_{22} = \nu\,.$$

We note that the "new" basis vectors are simply scaled version of the "old" ones: $\vec{e}'_i$ is parallel to $\vec{e}_i$ but may have different length (if $\mu, \nu \neq 1$).
(b) The simple relations between the "new" and "old" basis vectors allow us to find

$$\vec{e}_1 = \vec{e}'_1/\mu$$

and

$$\vec{e}_2 = \vec{e}'_2/\nu\,.$$

If the vector $\vec{a}$ is expanded using the "old" basis:

$$\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2\,,$$

then we can write

$$\vec{a} = (a'_1/\mu)\vec{e}'_1 + (a'_2/\nu)\vec{e}'_2\,,$$

and immediately find

$$a'_1 = a_1/\mu\,, \quad a'_2 = a_2/\nu\,.$$

Therefore, when the "new" basis vectors are scaled by factors μ and ν, the corresponding "new" components of the vectors are scaled by $1/\mu$ and $1/\nu$ – *in the opposite direction*, to counter the effect of basis variation. The arrow-like vectors are thus called *contravariant vectors.*

Exercise ??

The compact expression

$$E'_{ij}E_{jk}$$

for $i = 1$ and $k = 2$ can be expanded into a sum:

$$E'_{1j}E_{j2} = E'_{11}E_{12} + E'_{12}E_{22} \,.$$

Exercise ??

The system of four equations

$$\begin{aligned} aw + cx &= 1\,, && (7.1) \\ bw + dx &= 0\,, && (7.2) \\ ay + cz &= 0\,, && (7.3) \\ by + dz &= 1 && (7.4) \end{aligned}$$

can be solved by noticing that the first two equations do not involve the unknowns from the second pair of equations, and vice versa.

From the equation

$$bw + dx = 0$$

we first find $w = -dx/b$ and substitute in into the first equation:

$$-adx/b + cx = 1\,,$$

from which we easily find

$$x = \frac{b}{cb - ad} = -\frac{b}{\Delta}\,,$$

where we introduced the notation $\Delta = ad - bc$. Then

$$w = -\frac{dx}{b} = \frac{d}{\Delta}\,.$$

The second pair of equations can be solved similarly. First, we get

$$z = -\frac{ay}{c}\,,$$

and substitute it into the last of four equations:

$$by - \frac{ady}{c} = 1\,.$$

From the last expression follows

$$y = -\frac{c}{\Delta}\,.$$

Consequently,

$$z = \frac{a}{\Delta}\,.$$

Exercise ??

Firstly, we start with the compact expression

$$E_{ij}E'_{jk} = \delta_{ik}$$

and write it out fully for all four combinations of the indices i and k:

$$\begin{aligned} E_{11}E'_{11} + E_{12}E'_{21} &= 1\,, \\ E_{11}E'_{12} + E_{12}E'_{22} &= 0\,, \\ E_{21}E'_{11} + E_{22}E'_{21} &= 0\,, \\ E_{21}E'_{12} + E_{22}E'_{22} &= 1\,. \end{aligned}$$

Secondly, using the notation

$$E_{11} = a\,, \quad E_{12} = b\,, \quad E_{21} = c\,, \quad E_{22} = d\,,$$

and

$$E'_{11} = w\,, \quad E'_{12} = x\,, \quad E'_{21} = w\,, \quad E'_{22} = z\,,$$

we arrive at the four equations which we can group into two pairs of equations, each pair involves only two unknowns. The first pair is

$$\begin{aligned} aw + by &= 1\,, \\ cw + dy &= 0\,; \end{aligned}$$

the second pair:

$$\begin{aligned} cx + dz &= 1\,, \\ ax + bz &= 0\,. \end{aligned}$$

The first pair is easily solved when we find

$$w = -\frac{dy}{c}\,,$$

and substitute it into the first equation of the first pair:

$$-\frac{ady}{c} + by = 1\,,$$

from which follows:

$$y = -\frac{c}{\Delta} \qquad \Delta = ad - bc\,.$$

Immediately we get

$$w = \frac{d}{\Delta}\,.$$

Similarly, we first find

$$z = -\frac{ax}{b}\,,$$

and substitute into the first equation of the second pair:

$$cx - \frac{adx}{b} = 1\,.$$

Solving for x, we get

$$x = \frac{b}{\Delta}\,,$$

and therefore

$$z = \frac{a}{\Delta}\,.$$

We conclude that although two conditions $E'_{ij}E_{jk} = \delta_{ik}$ and $E_{ij}E'_{jk} = \delta_{ik}$ result in slightly different equations, they put the same constraints on the relations between the coefficients $E_{ij}\,(a, b, c, d)$ and $E'_{nm}\,(w, x, y, z)$.

Exercise 4.1

The equation of a circle with the radius R can be written using Cartesian coordinates:

$$x^2 + y^2 = R^2\,.$$

The transformation

$$b_1 = a_1 + a_2\,, \qquad b_2 = a_1 * a_2$$

moves every point (x, y) into a new point (x', y') related by the same equations:

$$x' = x + y\,, \qquad y' = xy\,.$$

Squaring x', we get

$$(x')^2 = x^2 + y^2 + 2xy = R^2 + 2y'\,.$$

Therefore, the components of the transformed vector are related as follows:

$$y' = (x')^2/2 - R^2/2 \quad \Leftrightarrow \quad b_2 = b_1^2/2 - R^2/2\,.$$

Exercise 4.2

The operator of normalization $\widehat{N}$ fails to satisfy the first linearity condition because

$$\widehat{N}\,(\alpha\vec{a}) \neq \alpha(\widehat{N}\,\vec{a})\,.$$

Indeed, the left-hand side must be a unit vector in the direction of $\alpha\vec{a}$, which is the same as the direction of $\vec{a}$:

$$\widehat{N}\,(\alpha\vec{a}) = \vec{u}_a = \widehat{N}\,\vec{a}\,.$$

In addition, the operator $\widehat{N}$ does not satisfy the second linearity condition:

$$\widehat{N}\,(\vec{a} + \vec{b}) = (\widehat{N}\,\vec{a}) + (\widehat{N}\,\vec{b})\,.$$

Take, for instance, $\vec{a} = \vec{e}_1$ and $\vec{b} = 1000\vec{e}_2$. The sum-vector $\vec{a} + \vec{b}$ will be pointing almost along the second basis vector $\vec{e}_2$, therefore

$$\widehat{N}\,(\vec{e}_1 + 1000\vec{e}_2)$$

will be a unit vector *almost parallel* to $\vec{e}_2$. However, the vector

$$(\widehat{N}\,\vec{e}_1) + (\widehat{N}\,[1000\vec{e}_2]) = \vec{e}_1 + \vec{e}_2$$

will go diagonally between $\vec{e}_1$ and $\vec{e}_2$.

Exercise 4.3

The condition for degeneracy of a linear operator can be written as follows:

$$\widehat{L}\,\vec{e}_1 = \lambda\left(\widehat{L}\,\vec{e}_2\right)\,.$$

For simplicity, let us denote $\vec{v} = \widehat{L}\,\vec{e}_2$. Then we have

$$\widehat{L}\,\vec{e}_1 = \lambda\vec{v}\,.$$

For an arbitrary vector $\vec{a}$ we can find the action of the degenerate linear operator $\widehat{L}$

$$\widehat{L}\,\vec{a} = \widehat{L}\,(a_1\vec{e}_1 + a_2\vec{e}_2) = a_1(\widehat{L}\,\vec{e}_1) + a_2(\widehat{L}\,\vec{e}_2)\,.$$

Now we can use the degeneracy condition to find

$$\widehat{L}\,\vec{a} = a_1\lambda\vec{v} + a_2\vec{v} = (a_1\lambda + a_2)\vec{v} = \alpha\vec{v}\,.$$

Thus, any vector is mapped into a vector parallel to $\vec{v}$. In other words, degenerate linear operator "collapses" all vectors onto a single line.

Exercise 4.4

The relation

$$\widehat{L}\,\vec{a} = \vec{b}$$

can be written using components relative to the "new" basis $\{\vec{e}\,'_i\}$. Expanding the vector $\vec{a}$, we get:

$$\widehat{L}\,(a'_i\vec{e}\,'_i) = a'_i(\widehat{L}\,\vec{e}\,'_i)\,.$$

Now the components of the operator $\widehat{L}$ relative to the "new" basis are defined similarly to the components relative to the "old" basis:

$$\widehat{L}\,\vec{e}\,'_i = L'_{ij}\vec{e}\,'_j\,.$$

Combining the last two expressions, we obtain

$$\widehat{L}\,\vec{a} = (a'_iL'_{ij})\vec{e}\,'_j\,,$$

where we indicated that the summation with the components of the vector $\vec{a}$ happens first. Comparing this result with the expansion of the vector $\vec{b}$ relative to the "new" basis

$$\vec{b} = b'_j\vec{e}\,'_j\,,$$

we can see that the following relation holds

$$a'_iL'_{ij} = b'_j\,.$$

Exercise 4.5

To evaluate the right-hand side of the expression

$$L_{11} + L_{22} = L'_{11} + L'_{22}\,,$$

we need to recall to rule of transformation of components of the linear operator:

$$L'_{mj} = E_{mk} L_{ki} E'_{ij}\,.$$

Using the last expression we can write

$$L'_{11} = E_{1k} L_{ki} E'_{i1} = (E'_{i1} E_{1k}) L_{ki}$$

and

$$L'_{22} = E_{2k} L_{ki} E'_{i2} = (E'_{i2} E_{2k}) L_{ki}\,.$$

Summing up, we obtain

$$L'_{11} + L'_{22} = \left(E'_{i1} E_{1k} + E'_{i2} E_{2k}\right) L_{ki}\,.$$

The sum in the parentheses can be made more compact using Einstein's summation rule:

$$E'_{i1} E_{1k} + E'_{i2} E_{2k} = E'_{ij} E_{jk}\,.$$

We showed that

$$E'_{ij} E_{jk} = \delta_{ik}\,,$$

therefore

$$L'_{11} + L'_{22} = \delta_{ik} L_{ki} = L_{ii} = L_{11} + L_{22}\,.$$

Exercise 5.1

The operator $\angle$ is not linear in either of its arguments. Indeed, scaling the first argument by an arbitrary factor α does not affect the measured angle:

$$\angle\,(\alpha\vec{a})\,\vec{b} = \angle\,\vec{a}\,\vec{b} \neq \alpha(\angle\,\vec{a}\,\vec{b})\,.$$

Same applies to the second argument.

Exercise 5.2

Components of any linear operator are defined as the coefficients in the expansion

$$\widehat{L}\,\vec{e}_i = L_{ij}\vec{e}_j\,.$$

For the potential operator $\widehat{\beta}$ this means

$$\widehat{\beta}\,\vec{e}_i = (a_i b_j)\vec{e}_j = a_i(b_j\vec{e}_j) = a_i\vec{b}\,.$$

Thus, the operator $\widehat{\beta}$ maps all basis vectors into vectors parallel to $\vec{b} = b_j\vec{e}_j$. Consequently, the operator $\widehat{\beta}$ maps *all* vectors into the same direction parallel to the vector $\vec{b}$. It is an example of a *degenerate linear operator.* See also Exercise 4.3.

Exercise 5.3

By definition

$$\overleftarrow{a} = \widehat{\sigma}\,\vec{a} = \widehat{\sigma}\,(a_1\vec{e}_1 + a_2\vec{e}_2)\,.$$

Using the linearity of $\widehat{\sigma}$, we first write

$$\widehat{\sigma}\,(a_1\vec{e}_1 + a_2\vec{e}_2) = [\widehat{\sigma}\,(a_1\vec{e}_1)] + [\widehat{\sigma}\,(a_2\vec{e}_2)]\,.$$

Using the linearity again, we get

$$\widehat{\sigma}\,(a_1\vec{e}_1) = a_1(\widehat{\sigma}\,\vec{e}_1) = a_1\overleftarrow{e}_1$$

and

$$\widehat{\sigma}\,(a_2\vec{e}_2) = a_2(\widehat{\sigma}\,\vec{e}_2) = a_2\overleftarrow{e}_2\,,$$

which lead to

$$\overleftarrow{a} = \widehat{\sigma}\,(a_1\vec{e}_1 + a_2\vec{e}_2) = a_i\overleftarrow{e}_i\,.$$

We showed that vector conjugate to $\vec{a}$ can be expanded in terms of basis conjugate to $\vec{e}_i$.

Exercise 5.4

The determinant of a linear operator $\widehat{L}$ can be calculated from its components according to:

$$det\,\widehat{L} = L_{11}L_{22} - L_{12}L_{21}\,.$$

For a projector $\widehat{\underline{\underline{A}}}$ we have

$$\underline{\underline{A}}_{11}\underline{\underline{A}}_{22} = \frac{a_1 a_1 a_2 a_2}{a^4} = \frac{a_1^2 a_2^2}{a^4}$$

and

$$\underline{\underline{A}}_{12}\underline{\underline{A}}_{21} = \frac{a_1 a_2 a_2 a_1}{a^4} = \frac{a_1^2 a_2^2}{a^4}\,.$$

It immediately follows that $det\,\widehat{L} = 0$.

A helpful related exercise is Exercise 5.2.

Exercise 5.5

(a) First, we can write symbolically:

$$\widehat{\underline{\underline{A}}} \circ \widehat{\underline{\underline{A}}} = \frac{(\overrightarrow{a}\overleftarrow{a})}{a^2}\,\frac{(\overrightarrow{a}\overleftarrow{a})}{a^2} = \frac{\overrightarrow{a}(\overleftarrow{a}\overrightarrow{a})\overleftarrow{a}}{a^4}\,,$$

which, using the fact $\overleftarrow{a}\overrightarrow{a} = a^2$ is reduced to

$$\widehat{\underline{\underline{A}}} \circ \widehat{\underline{\underline{A}}} = \frac{\overrightarrow{a}\overleftarrow{a}}{a^2} = \widehat{\underline{\underline{A}}}\,.$$

Second, using components, we write the product of two operators as follows:

$$(\underline{\underline{A}}_{ij})(\underline{\underline{A}}_{jk}) = \frac{a_i a_j a_j a_k}{a^4}\,.$$

Recalling that $a_j a_j = a^2$, we find

$$(\underline{\underline{A}}_{ij})(\underline{\underline{A}}_{jk}) = \frac{a_i a_k}{a^2} = \underline{\underline{A}}_{ik}\,.$$

Every projector of the type $L = \overrightarrow{d}\overleftarrow{d}/d^2$ has this property.

(b) For a composition of two projectors

$$\widehat{L} = \widehat{\underline{\underline{B}}} \circ \widehat{\underline{\underline{A}}}$$

the components are given by

$$L_{ik} = \lambda\, a_i b_k, \qquad \lambda = \frac{\vec{a}\cdot\vec{b}}{a^2 b^2}\,.$$

Composition of two such operators can be evaluated using their components:

$$L_{ik}L_{kj} = (\lambda\, a_i b_k)(\lambda\, a_k b_j) = \lambda^2 a_i (a_k b_k) b_j = \lambda^2(\vec{a}\cdot\vec{b}) a_i b_j\,.$$

We thus showed that

$$\widehat{L} \circ \widehat{L} = \lambda(\vec{a} \cdot \vec{b})\widehat{L}\,.$$

Now

$$\lambda(\vec{a} \cdot \vec{b}) = \frac{(\vec{a} \cdot \vec{b})^2}{a^2 b^2} = \cos\theta\,,$$

where θ is the angle between the vectors $\vec{a}$ and $\vec{b}$. Therefore,

$$\widehat{L} \circ \widehat{L} = \cos\theta \widehat{L} \neq \widehat{L}\,.$$

Only for $\vec{a} = \vec{b}$ we have the property $\widehat{L} \circ \widehat{L} = \widehat{L}$.

Exercise 5.6

The components of the composition

$$\widehat{L} = \underline{\underline{\widehat{B}}} \circ \underline{\underline{\widehat{A}}}\,.$$

are

$$L_{ik} = \lambda\, a_i b_k, \quad \lambda = \frac{\vec{a} \cdot \vec{b}}{a^2 b^2}\,.$$

Reversing the order of arguments of the composition results in

$$\widehat{M} = \underline{\underline{\widehat{A}}} \circ \underline{\underline{\widehat{B}}}\,,$$

with the components

$$M_{ik} = \lambda\, b_i a_k, \quad \lambda = \frac{\vec{b} \cdot \vec{a}}{a^2 b^2}\,.$$

In general, $M_{12} \neq L_{12}$ because $a_1 b_2 \neq b_1 a_2$.

Exercise 5.7

To arrive at the transformation rules for the components of different types of tensors, we will use a simple fact: A general tensor with components t^{ij} will behave just like the tensor product of two contra-variant vectors $a^i b^j$. Thus, we will study four types of tensor products:

$$a^i b^j\,, \quad a_i b_j\,, \quad a^i b_j\,, \quad a_i b^j\,.$$

It will be helpful to recall the transformation rules of contravariant and covariant vectors. An arrow-like contravariant vector can be expanded in a basis:

$$\vec{a} = a^i \vec{e}_i \,.$$

Every vectors from a "new" basis can be similarly expanded:

$$\vec{e}\,'_j = E_{j\bullet}^{\bullet k} \vec{e}_k \,.$$

Here we deliberately included the "dummy" symbol " $\bullet$ " to visually align the indices according to their order. Expanding the same vector $\vec{a}$ relative to the "new" basis has the form:

$$\vec{a} = a'^i \vec{e}\,'_i \,,$$

and similarly

$$\vec{a} = a^i \vec{e}_i \,.$$

$$\vec{e}_k = (E')_{k\bullet}^{\bullet j} \vec{e}_j \,.$$

We also obtained the relation between the components of the same vector $\vec{a}$ in different bases:

$$a'^i = (E')_{k\bullet}^{\bullet i} a^k \,.$$

Another contravariant vector will have similar relations:

$$b'^j = (E')_{l\bullet}^{\bullet j} b^l \,,$$

and their tensor product $t^{ij} = a^i b^j$ will be transformed according to

$$(t')^{ij} = (E')_{k\bullet}^{\bullet i} (E')_{l\bullet}^{\bullet j} t^{kl} \,.$$

Using the transformation rule of covariant components:

$$a'_i = E_{i\bullet}^{\bullet k} a_k \quad \text{and} \quad b'_j = E_{j\bullet}^{\bullet l} a_l \,,$$

we immediately arrive at transformation rule of the components of a doubly-covariant tensor:

$$t'_{ij} = E_{i\bullet}^{\bullet k} E_{j\bullet}^{\bullet l} t_{kl} \,.$$

In a similar fashion, by combining the transformation rules of contravariant and covariant vectors, we can obtain the transformation of contra-covariant tensor:

$$(t')^{i\bullet}_{\bullet j} = (E')_{k\bullet}^{\bullet i} E_{j\bullet}^{\bullet l} t^{k\bullet}_{\bullet l} \,,$$

and covariant-contravariant tensor:

$$(t')_{i\bullet}^{\bullet j} = E_{i\bullet}^{\bullet k} (E')_{l\bullet}^{\bullet j} t_{k\bullet}^{\bullet l} \ .$$

Exercise 6.1

(a) The action of the metric tensor on a tensor product $\vec{a} \otimes \vec{b}$ can be understood once we write it out using components. First, let's write the expression for the distance squared:

$$d^2 = \eta_{ij} d^i d^j \ .$$

In a similar way, we can write

$$\widehat{\eta}(\vec{a} \otimes \vec{b}) = \eta_{ij} a^i b^j \ .$$

Using the fact

$$\eta_{ij} = \widehat{\sigma}\, \vec{e}_i\, \vec{e}_j$$

and bilinear nature of the dol-operator $\widehat{\sigma}$, we deduce

$$\eta_{ij} a^i b^j = a^i b^j (\widehat{\sigma}\, \vec{e}_i\, \vec{e}_j) = \widehat{\sigma}\,(a^i \vec{e}_i)\,(b^j \vec{e}_j) = \widehat{\sigma} \vec{a}\, \vec{b} \ .$$

Thus, we showed that

$$\widehat{\eta}(\vec{a} \otimes \vec{b}) = \vec{a} \cdot \vec{b} \ .$$

The connection between the metric tensor and scalar products is not surprising. Indeed, if we recall that if vector $\vec{d}$ connects two points in a plane and is given by

$$\vec{d} = \vec{b} - \vec{a} \ , \qquad d^i = b^i - a^i \ ,$$

then

$$\eta_{ij} d^i d^j = \eta_{ij} (a^i - b^i)(a^j - b^j) = \eta_{ij} a^i a^j + \eta_{ij} b^i b^j + 2\eta_{ij} a^i b^j \ .$$

We arrived at the familiar theorem of planar geometry – theorem of cosine:

$$d^2 = a^2 + b^2 - 2ab\cos\theta \ .$$

(b) Vectors of orthonormal basis all have unit lengths (*normalized* vectors):

$$e_i^2 = \widehat{\sigma}\,\vec{e}_i\,\vec{e}_i = \eta_{ii} = 1\,.$$

Different vectors of orthonormal basis are perpendicular to each other (*orthogonal* vectors):

$$\vec{e}_i \cdot \vec{e}_j = \widehat{\sigma}\,\vec{e}_i\,\vec{e}_j = \eta_{ij} = 0\,.$$

We showed that

$$\eta_{ij} = \delta_{ij}$$

when the basis is orthonormal.

8. Appendix

Exercise 1.1

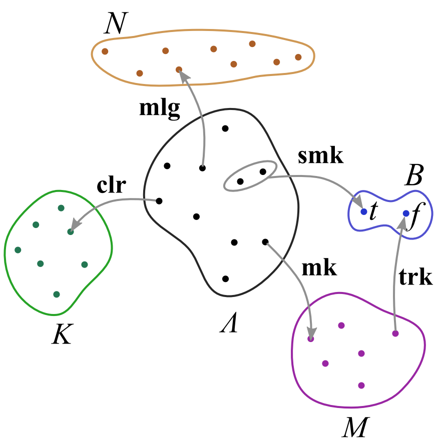


Fig. 8.1: The set M contains all possible makes of cars: Ford, Toyota, etc.

The diagram in the Figure 9.1 shows the set M – the set of all possible makes of cars. A mapping **trk** returns $true$ if a given car maker produces trucks.

Exercise 2.1

Any binary function can be viewed as a unary function if two inputs are replaced by a single input of a *pair of numbers*. Similarly for a function with two outputs. This idea is illustrated in the Figure 9.2(a): The function **swp** is viewed as a unary function which swaps the numbers in an *ordered pair*:

$$\mathbf{swp}\ (n, m) = (m, n)\,.$$

Given the set $\mathbb{Z}$ of whole numbers, we can create the set of all possible *ordered pairs* (n, m). This set can be denoted as follows:

$$(\mathbb{Z}, \mathbb{Z}) \text{ or } \mathbb{Z} \times \mathbb{Z}\,.$$

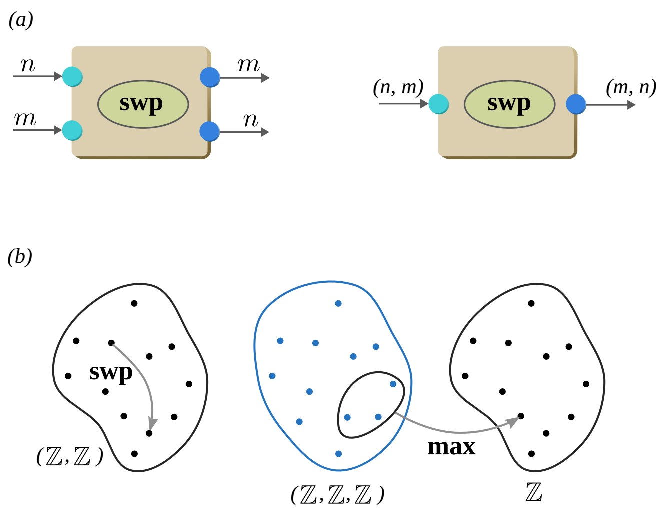


Fig. 8.2: (a) Two inputs (outputs) of a function can be replaced with a single input of a ***pair*** of numbers, turning a binary function into a unary one. (b) That.

The latter notation is standard in mathematics, but the former way of writing is also acceptable. We can similarly denote the set of all *ordered triples*:

$$(\mathbb{Z}, \mathbb{Z}, \mathbb{Z}) \text{ or } \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z} \, .$$

With the notation introduced above, the action of functions with multiple inputs or outputs can be depicted on the level of sets. The Figure 9.2(b) shows how this works for the functions **swp** and **max** .

Exercise 2.2

Consider a binary function that accepts a pair of natural numbers and returns the third natural number in the following way:

$$\textbf{rep}\ 3\,2 = 33 \qquad \textbf{rep}\ 1\,4 = 1111\, .$$

Thus, the output is a natural number with identical digits given by the first number, repeated a number of times specified by the second number.

Infix variant of this operation can be written, rather arbitrarily, like this:

$$\textbf{rep}\ n\, m = n \succ m = nnn \ldots n\, .$$

Exercise ??

A linear function f must satisfy the linearity condition

$$f\,(a * n) = a * (f\,n)\, .$$

For $a = 0$ we must have

$$f\,(0 * n) = 0 * (f\,n)\, ,$$

or, equivalently

$$f\,0 = 0\,.$$

Also, for $a = m$ and $n = 1$ we must have

$$f\,(m * 1) = m * (f\,1)\,,$$

from which follows

$$f\,m = m(f\,1)\,.$$

Exercise ??

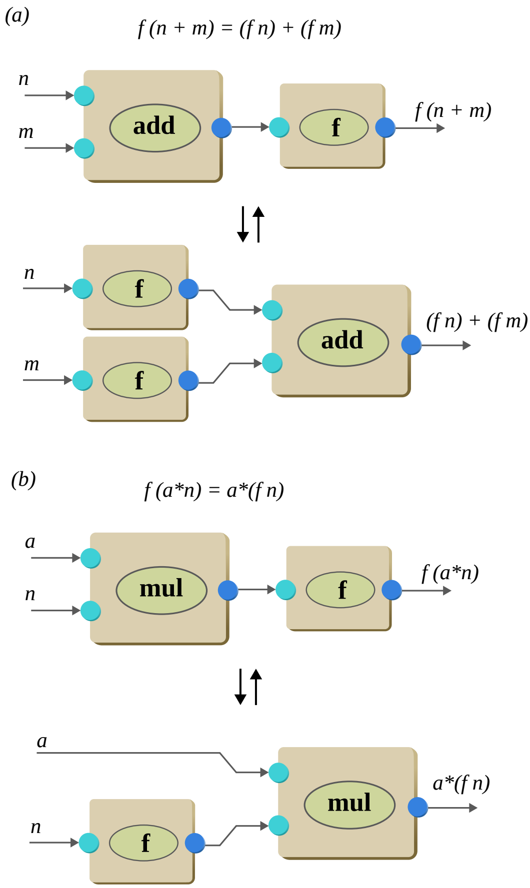


Fig. 8.3: The linearity conditions can be represented schematically with different relative configurations (order) of the "boxes".

The schematics in the Figure 8.3(a) and (b) demonstrate the two linearity requirements.

Exercise ??
Using Einstein's summation rule, a polynomial of degree n can be written as follows:

$$P_n\, x = a_i x^i\,, \quad i = 0, 1, 2 \ldots, n\,.$$

Exercise ??
The expression

$$b_i y_i\,, \quad i = 1, 2, 3, 4$$

represents the sum of four terms:

$$S = b_1 y_1 + b_2 y_2 + b_3 y_3 + b_4 y_4\,.$$

Similarly,

$$b_j y_j\,, \quad j = 1, 2, 3, 4$$

stands for the same sum S, just as the expression

$$b_k y_k = b_1 y_1 + b_2 y_2 + b_3 y_3 + b_4 y_4\,.$$

Exercise ??
In the expression

$$(a_i x_i)(a_j x_j)$$

both parentheses contain the identical sum:

$$a_i x_i = a_j x_j = a_1 x_1 + a_2 x_2\,.$$

Opening the parentheses, we obtain

$$(a_i x_i)(a_j x_j) = a_1^2 x_1^2 + a_2^2 x_2^2 + 2 a_1 a_2 x_1 x_2\,.$$

In contrast, the expression $a_i^2 x_i^2$ stands for

$$a_i^2 x_i^2 = a_1^2 x_1^2 + a_2^2 x_2^2\,.$$

Clearly,

$$a_i^2 x_i^2 \neq (a_i x_i)^2\,.$$

Exercise ??
The left-hand side of the expression

$$(a_i x_i)^2 = \frac{b_j y_j}{c_k c_k}$$

can be written as

$$(a_i x_i)^2 = \left(\sum_{i=1}^{i=N} a_i x_i \right)^2$$

The right-hand side takes the form:

$$\frac{b_j y_j}{c_k c_k} = \frac{\sum\limits_{j=1}^{j=N} b_j y_j}{\sum\limits_{k=1}^{k=N} c_k c_k}\,.$$

Therefore, the original equality can be re-written using the traditional summation sign:

$$\left(\sum_{i=1}^{i=N} a_i x_i\right)^2 = \frac{\sum_{j=1}^{j=N} b_j y_j}{\sum_{k=1}^{k=N} c_k c_k} \,.$$

Already one can see the advantage of Einstein's summation rule.

Exercise ??

The expression

$$\delta_{1i} a_i \,, \quad i = 1, 2, 3, 4$$

represents the sum

$$\delta_{11} a_1 + \delta_{12} a_2 + \delta_{13} a_3 + \delta_{14} a_4 \,.$$

The only non-zero term corresponds to $\delta_{11} = 1$, therefore

$$\delta_{1i} a_i = a_1 \,.$$

Similarly, we have

$$\delta_{3k} a_k \,, \quad k = 1, 2, 3, 4$$

representing

$$\delta_{31} a_1 + \delta_{32} a_2 + \delta_{33} a_3 + \delta_{34} a_4 = \delta_{33} a_3 \,.$$

Consequently,

$$\delta_{3k} a_k = a_3 \,.$$

Next,

$$\epsilon_{1j} a_j = \epsilon_{11} a_1 + \epsilon_{12} a_2 + \epsilon_{13} a_3 + \epsilon_{14} a_4$$

can be simplified to

$$\epsilon_{1j} a_j = a_2 + a_3 + a_4$$

using the definition of ϵ_{ij}.

Finally, the expression

$$\epsilon_{3j} a_j = \epsilon_{31} a_1 + \epsilon_{32} a_2 + \epsilon_{33} a_3 + \epsilon_{34} a_4$$

is reduced to

$$\epsilon_{3j} a_j = a_4 - a_1 - a_2 \,.$$

Exercise ??

The sum

$$a_i + a_j$$

can be rewritten using the facts $a_j = \delta_{ji} a_i$ and $\delta_{ij} = \delta_{ji}$:

$$a_i + a_j = a_i + \delta_{ij} a_i = (1 + \delta_{ij}) a_i = (1 + \delta_{ji}) a_i \,.$$

Exercise ??

(a) The expression $\delta_{ij} a_i b_j$ can be simplified using the fact $\delta_{ij} b_j = b_i$:

$$\delta_{ij} a_i b_j = a_i b_i = a_1 b_1 + a_2 b_2 \,.$$

(b) Fully writing out $\epsilon_{ij} a_i b_j$ results in

$$\epsilon_{11} a_1 b_1 + \epsilon_{12} a_1 b_2 + \epsilon_{21} a_2 b_1 + \epsilon_{22} a_2 b_2 \,.$$

From the definition of ϵ_{ij} follows that only the terms with $i \neq j$ survive:

$$\epsilon_{ij} a_i b_j = a_1 b_2 - a_2 b_1 \,.$$

Exercise ??

(a) Firstly, we can recall that when δ_{ij} is summed with a vector a_j it simply "renames" the index that is being used for summation:

$$\delta_{ij} a_j = a_i \,.$$

Using this property, we immediately get

$$\delta_{ij} \delta_{jk} = \delta_{ik} \,.$$

Another – and longer – way to get this result is to write out the summation fully:

$$\delta_{ij} \delta_{jk} = \delta_{i1} \delta_{1k} + \delta_{i2} \delta_{2k} + \ldots + \delta_{in} \delta_{nk} \,.$$

If $i \neq k$, all terms on the right are zero. Indeed, $\delta_{i1}\delta_{1k}$ is zero unless $i = 1$ and $k = 1$; similarly, $\delta_{i2}\delta_{2k}$ is zero unless $i = 2$ and $k = 2$ and so on. Therefore, the only non-zero value for $\delta_{ij}\delta_{jk}$ is when $i = k$. Let $i = k = m$, then in the sum

$$\delta_{m1} \delta_{1m} + \delta_{m2} \delta_{2m} + \ldots + \delta_{mm} \delta_{mm} + \ldots + \delta_{mn} \delta_{nm} \,.$$

there is only one non-zero term, namely

$$\delta_{mm} \delta_{mm} = 1 \cdot 1 = 1 \,.$$

Summarizing the above arguments, we conclude that

$$\delta_{ij} \delta_{jk} = 1 \text{ if } i = k \text{ and } 0 \text{ otherwise} \,.$$

This is equivalent to the expression

$$\delta_{ij} \delta_{jk} = \delta_{ik} \,.$$

(b) The expression $\epsilon_{ij}\epsilon_{jk}$, when fully expanded as a sum, takes the form

$$\epsilon_{ij} \epsilon_{jk} = \epsilon_{i1} \epsilon_{1k} + \epsilon_{i2} \epsilon_{2k} \,.$$

If $i = k = 1$, the sum is reduced to

$$\epsilon_{1j} \epsilon_{j1} = \epsilon_{11} \epsilon_{11} + \epsilon_{12} \epsilon_{21} = 1 \cdot (-1) = -1 \,.$$

Similarly, for $i = k = 2$ we get

$$\epsilon_{2j} \epsilon_{j2} = \epsilon_{21} \epsilon_{12} + \epsilon_{22} \epsilon_{22} = (-1) \cdot 1 = -1 \,.$$

On the other hand, if $i = 1$ and $k = 2$, we obtain

$$\epsilon_{1j} \epsilon_{j2} = \epsilon_{11} \epsilon_{12} + \epsilon_{12} \epsilon_{22} = 0 \,.$$

Same for $i = 2$ and $k = 1$:

$$\epsilon_{2j}\epsilon_{j1} = \epsilon_{21}\epsilon_{11} + \epsilon_{22}\epsilon_{21} = 0 \,.$$

We thus manually checked all cases and showed that

$$\epsilon_{ij}\epsilon_{jk} = -\delta_{ik} \quad i, j, k = 1, 2 \,.$$

Exercise 2.3

Let us denote:

$$x = \epsilon_{ij} a_i a_j \,.$$

Since we can rename the summation indices, we can write

$$\epsilon_{ij} a_i a_j = \epsilon_{ik} a_i a_k = \epsilon_{jk} a_j a_k = \epsilon_{ji} a_j a_i \,.$$

Now we have $\epsilon_{ji} = -\epsilon_{ij}$ and this leads to

$$\epsilon_{ji} a_j a_i = -\epsilon_{ij} a_i a_j \,.$$

We thus showed that $x = -x$ and therefore $x = 0$.

Exercise ??

An expansion of an arbitrary vector $\vec{a}$ in terms of the basis vectors is given by

$$\vec{a} = a_1 \vec{e}_1 + a_2 \vec{e}_2 + a_3 \vec{e}_3 + \ldots a_n \vec{e}_n \,.$$

This can be compactly written using Einstein's summation rule:

$$\vec{a} = a_i \vec{e}_i \quad i = 1, 2, \ldots, n \,.$$

If the number of basis vectors is known and fixed, as is usually the case, we can omit the range of the summation index and simply write

$$\vec{a} = a_i \vec{e}_i \,.$$

Exercise ??

The expression

$$\vec{e}\,'_1 = E_{11} \vec{e}_1 + E_{12} \vec{e}_2 \,,$$

can be written using Einstein's summation rule as follows:

$$\vec{e}\,'_1 = E_{1j} \vec{e}_j \,.$$

Similarly for the second basis vector:

$$\vec{e}\,'_2 = E_{2j} \vec{e}_j \,.$$

Combining both results, we obtain

$$\vec{e}\,'_i = E_{ij} \vec{e}_j \,.$$

Exercise ??

(a) Writing the expansion of the "new" basis as follows:

$$\vec{e}\,'_1 = \mu\vec{e}_1 + 0\vec{e}_2\,,$$

$$\vec{e}\,'_2 = 0\vec{e}_1 + \nu\vec{e}_2\,,$$

we can immediately find the components E_{ij}:

$$E_{11} = \mu,\ E_{12} = 0,\ E_{21} = 0,\ E_{22} = \nu\,.$$

We note that the "new" basis vectors are simply scaled version of the "old" ones: $\vec{e}\,'_i$ is parallel to $\vec{e}_i$ but may have different length (if $\mu, \nu \neq 1$). (b) The simple relations between the "new" and "old" basis vectors allow us to find

$$\vec{e}_1 = \vec{e}\,'_1/\mu$$

and

$$\vec{e}_2 = \vec{e}\,'_2/\nu\,.$$

If the vector $\vec{a}$ is expanded using the "old" basis:

$$\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2\,,$$

then we can write

$$\vec{a} = (a'_1/\mu)\vec{e}\,'_1 + (a'_2/\nu)\vec{e}\,'_2\,,$$

and immediately find

$$a'_1 = a_1/\mu\,, \quad a'_2 = a_2/\nu\,.$$

Therefore, when the "new" basis vectors are scaled by factors μ and ν, the corresponding "new" components of the vectors are scaled by $1/\mu$ and $1/\nu$ – *in the opposite direction*, to counter the effect of basis variation. The arrow-like vectors are thus called *contravariant vectors.*

Exercise ??

The compact expression

$$E'_{ij}E_{jk}$$

for $i = 1$ and $k = 2$ can be expanded into a sum:

$$E'_{1j}E_{j2} = E'_{11}E_{12} + E'_{12}E_{22}\,.$$

Exercise ??

The system of four equations

$$\begin{aligned} aw + cx &= 1\,, && (8.1)\\ bw + dx &= 0\,, && (8.2)\\ ay + cz &= 0\,, && (8.3)\\ by + dz &= 1 && (8.4) \end{aligned}$$

can be solved by noticing that the first two equations do not involve the unknowns from the second pair of equations, and vice versa.

From the equation

$$bw + dx = 0$$

we first find $w = -dx/b$ and substitute in into the first equation:

$$-adx/b + cx = 1\,,$$

from which we easily find

$$x = \frac{b}{cb - ad} = -\frac{b}{\Delta}\,,$$

where we introduced the notation $\Delta = ad - bc$. Then

$$w = -\frac{dx}{b} = \frac{d}{\Delta}\,.$$

The second pair of equations can be solved similarly. First, we get

$$z = -\frac{ay}{c}\,,$$

and substitute it into the last of four equations:

$$by - \frac{ady}{c} = 1\,.$$

From the last expression follows

$$y = -\frac{c}{\Delta}\,.$$

Consequently,

$$z = \frac{a}{\Delta}\,.$$

Exercise ??

Firstly, we start with the compact expression

$$E_{ij}E'_{jk} = \delta_{ik}$$

and write it out fully for all four combinations of the indices i and k:

$$\begin{aligned}
E_{11}E'_{11} + E_{12}E'_{21} &= 1\,,\\
E_{11}E'_{12} + E_{12}E'_{22} &= 0\,,\\
E_{21}E'_{11} + E_{22}E'_{21} &= 0\,,\\
E_{21}E'_{12} + E_{22}E'_{22} &= 1\,.
\end{aligned}$$

Secondly, using the notation

$$E_{11} = a\,,\quad E_{12} = b\,,\quad E_{21} = c\,,\quad E_{22} = d\,,$$

and

$$E'_{11} = w\,,\quad E'_{12} = x\,,\quad E'_{21} = w\,,\quad E'_{22} = z\,,$$

we arrive at the four equations which we can group into two pairs of equations, each pair involves only two unknowns. The first pair is

$$\begin{aligned} aw + by &= 1\,, \\ cw + dy &= 0\,; \end{aligned}$$

the second pair:

$$\begin{aligned} cx + dz &= 1\,, \\ ax + bz &= 0\,. \end{aligned}$$

The first pair is easily solved when we find

$$w = -\frac{dy}{c}\,,$$

and substitute it into the first equation of the first pair:

$$-\frac{ady}{c} + by = 1\,,$$

from which follows:

$$y = -\frac{c}{\Delta} \qquad \Delta = ad - bc\,.$$

Immediately we get

$$w = \frac{d}{\Delta}\,.$$

Similarly, we first find

$$z = -\frac{ax}{b}\,,$$

and substitute into the first equation of the second pair:

$$cx - \frac{adx}{b} = 1\,.$$

Solving for x, we get

$$x = \frac{b}{\Delta}\,,$$

and therefore

$$z = \frac{a}{\Delta}\,.$$

We conclude that although two conditions $E'_{ij}E_{jk} = \delta_{ik}$ and $E_{ij}E'_{jk} = \delta_{ik}$ result in slightly different equations, they put the same constraints on the relations between the coefficients E_{ij} (a, b, c, d) and E'_{nm} (w, x, y, z).

Exercise 4.1

The equation of a circle with the radius R can be written using Cartesian coordinates:

$$x^2 + y^2 = R^2\,.$$

The transformation

$$b_1 = a_1 + a_2\,, \quad b_2 = a_1 * a_2$$

moves every point (x, y) into a new point (x', y') related by the same equations:

$$x' = x + y\,, \quad y' = xy\,.$$

Squaring x', we get

$$(x')^2 = x^2 + y^2 + 2xy = R^2 + 2y'\,.$$

Therefore, the components of the transformed vector are related as follows:

$$y' = (x')^2/2 - R^2/2 \quad \Leftrightarrow \quad b_2 = b_1^2/2 - R^2/2\,.$$

Exercise 4.2

The operator of normalization $\widehat{N}$ fails to satisfy the first linearity condition because

$$\widehat{N}\,(\alpha\vec{a}) \neq \alpha(\widehat{N}\,\vec{a})\,.$$

Indeed, the left-hand side must be a unit vector in the direction of $\alpha\vec{a}$, which is the same as the direction of $\vec{a}$:

$$\widehat{N}\,(\alpha\vec{a}) = \vec{u}_a = \widehat{N}\,\vec{a}\,.$$

In addition, the operator $\widehat{N}$ does not satisfy the second linearity condition:

$$\widehat{N}\,(\vec{a} + \vec{b}) = (\widehat{N}\,\vec{a}) + (\widehat{N}\,\vec{b})\,.$$

Take, for instance, $\vec{a} = \vec{e}_1$ and $\vec{b} = 1000\vec{e}_2$. The sum-vector $\vec{a} + \vec{b}$ will be pointing almost along the second basis vector $\vec{e}_2$, therefore

$$\widehat{N}\,(\vec{e}_1 + 1000\vec{e}_2)$$

will be a unit vector *almost parallel* to $\vec{e}_2$. However, the vector

$$(\widehat{N}\,\vec{e}_1) + (\widehat{N}\,[1000\vec{e}_2]) = \vec{e}_1 + \vec{e}_2$$

will go diagonally between $\vec{e}_1$ and $\vec{e}_2$.

Exercise 4.3

The condition for degeneracy of a linear operator can be written as follows:

$$\widehat{L}\,\vec{e}_1 = \lambda\left(\widehat{L}\,\vec{e}_2\right)\,.$$

For simplicity, let us denote $\vec{v} = \widehat{L}\,\vec{e}_2$. Then we have

$$\widehat{L}\,\vec{e}_1 = \lambda\vec{v}\,.$$

For an arbitrary vector $\vec{a}$ we can find the action of the degenerate linear operator $\widehat{L}$

$$\widehat{L}\,\vec{a} = \widehat{L}\,(a_1\vec{e}_1 + a_2\vec{e}_2) = a_1(\widehat{L}\,\vec{e}_1) + a_2(\widehat{L}\,\vec{e}_2)\,.$$

Now we can use the degeneracy condition to find

$$\widehat{L}\,\vec{a} = a_1\lambda\vec{v} + a_2\vec{v} = (a_1\lambda + a_2)\vec{v} = \alpha\vec{v}\,.$$

Thus, any vector is mapped into a vector parallel to $\vec{v}$. In other words, degenerate linear operator "collapses" all vectors onto a single line.

Exercise 4.4

The relation

$$\widehat{L}\,\vec{a} = \vec{b}$$

can be written using components relative to the "new" basis $\{\vec{e}\,'_i\}$. Expanding the vector $\vec{a}$, we get:

$$\widehat{L}\,(a'_i\vec{e}\,'_i) = a'_i(\widehat{L}\,\vec{e}\,'_i)\,.$$

Now the components of the operator $\widehat{L}$ relative to the "new" basis are defined similarly to the components relative to the "old" basis:

$$\widehat{L}\,\vec{e}\,'_i = L'_{ij}\,\vec{e}\,'_j\,.$$

Combining the last two expressions, we obtain

$$\widehat{L}\,\vec{a} = (a'_i L'_{ij})\vec{e}\,'_j\,,$$

where we indicated that the summation with the components of the vector $\vec{a}$ happens first. Comparing this result with the expansion of the vector $\vec{b}$ relative to the "new" basis

$$\vec{b} = b'_j\,\vec{e}\,'_j\,,$$

we can see that the following relation holds

$$a'_i L'_{ij} = b'_j\,.$$

Exercise 4.5

To evaluate the right-hand side of the expression

$$L_{11} + L_{22} = L'_{11} + L'_{22}\,,$$

we need to recall to rule of transformation of components of the linear operator:

$$L'_{mj} = E_{mk} L_{ki} E'_{ij}\,.$$

Using the last expression we can write

$$L'_{11} = E_{1k} L_{ki} E'_{i1} = (E'_{i1} E_{1k}) L_{ki}$$

and

$$L'_{22} = E_{2k} L_{ki} E'_{i2} = (E'_{i2} E_{2k}) L_{ki}\,.$$

Summing up, we obtain

$$L'_{11} + L'_{22} = \left(E'_{i1} E_{1k} + E'_{i2} E_{2k}\right) L_{ki}\,.$$

The sum in the parentheses can be made more compact using Einstein's summation rule:

$$E'_{i1} E_{1k} + E'_{i2} E_{2k} = E'_{ij} E_{jk}\,.$$

We showed that

$$E'_{ij} E_{jk} = \delta_{ik} \,,$$

therefore

$$L'_{11} + L'_{22} = \delta_{ik} L_{ki} = L_{ii} = L_{11} + L_{22} \,.$$

Exercise 5.1

The operator $\angle$ is not linear in either of its arguments. Indeed, scaling the first argument by an arbitrary factor α does not affect the measured angle:

$$\angle\,(\alpha\vec{a})\,\vec{b} = \angle\,\vec{a}\,\vec{b} \neq \alpha(\angle\,\vec{a}\,\vec{b}) \,.$$

Same applies to the second argument.

Exercise 5.2

Components of any linear operator are defined as the coefficients in the expansion

$$\widehat{L}\,\vec{e}_i = L_{ij}\,\vec{e}_j \,.$$

For the potential operator $\widehat{\beta}$ this means

$$\widehat{\beta}\,\vec{e}_i = (a_i b_j)\vec{e}_j = a_i(b_j\,\vec{e}_j) = a_i\,\vec{b} \,.$$

Thus, the operator $\widehat{\beta}$ maps all basis vectors into vectors parallel to $\vec{b} = b_j\,\vec{e}_j$. Consequently, the operator $\widehat{\beta}$ maps *all* vectors into the same direction parallel to the vector $\vec{b}$. It is an example of a *degenerate linear operator*. See also Exercise 4.3.

Exercise 5.3

By definition

$$\overleftarrow{a} = \widehat{\sigma}\,\vec{a} = \widehat{\sigma}\,(a_1\,\vec{e}_1 + a_2\,\vec{e}_2) \,.$$

Using the linearity of $\widehat{\sigma}$, we first write

$$\widehat{\sigma}\,(a_1\,\vec{e}_1 + a_2\,\vec{e}_2) = [\widehat{\sigma}\,(a_1\,\vec{e}_1)] + [\widehat{\sigma}\,(a_2\,\vec{e}_2)] \,.$$

Using the linearity again, we get

$$\widehat{\sigma}\,(a_1\,\vec{e}_1) = a_1(\widehat{\sigma}\,\vec{e}_1) = a_1\,\overleftarrow{e}_1$$

and

$$\widehat{\sigma}\,(a_2\,\vec{e}_2) = a_2(\widehat{\sigma}\,\vec{e}_2) = a_2\,\overleftarrow{e}_2 \,,$$

which lead to

$$\overleftarrow{a} = \widehat{\sigma}\,(a_1\,\vec{e}_1 + a_2\,\vec{e}_2) = a_i\,\overleftarrow{e}_i \,.$$

We showed that vector conjugate to $\vec{a}$ can be expanded in terms of basis conjugate to $\vec{e}_i$.

Exercise 5.4

The determinant of a linear operator $\widehat{L}$ can be calculated from its components according to:

$$det\,\widehat{L} = L_{11}L_{22} - L_{12}L_{21} \,.$$

For a projector $\widehat{\underline{\underline{A}}}$ we have

$$\underline{\underline{A}}_{11}\underline{\underline{A}}_{22} = \frac{a_1 a_1 a_2 a_2}{a^4} = \frac{a_1^2 a_2^2}{a^4}$$

and

$$\underline{\underline{A}}_{12}\underline{\underline{A}}_{21} = \frac{a_1 a_2 a_2 a_1}{a^4} = \frac{a_1^2 a_2^2}{a^4}\,.$$

It immediately follows that $det\,\widehat{L} = 0$.

A helpful related exercise is Exercise 5.2.

Exercise 5.5

(a) First, we can write symbolically:

$$\widehat{\underline{\underline{A}}} \circ \widehat{\underline{\underline{A}}} = \frac{(\vec{a}\overleftarrow{a})}{a^2}\frac{(\vec{a}\overleftarrow{a})}{a^2} = \frac{\vec{a}(\overleftarrow{a}\vec{a})\overleftarrow{a}}{a^4}\,,$$

which, using the fact $\overleftarrow{a}\vec{a} = a^2$ is reduced to

$$\widehat{\underline{\underline{A}}} \circ \widehat{\underline{\underline{A}}} = \frac{\vec{a}\overleftarrow{a}}{a^2} = \widehat{\underline{\underline{A}}}\,.$$

Second, using components, we write the product of two operators as follows:

$$(\underline{\underline{A}}_{ij})(\underline{\underline{A}}_{jk}) = \frac{a_i a_j a_j a_k}{a^4}\,.$$

Recalling that $a_j a_j = a^2$, we find

$$(\underline{\underline{A}}_{ij})(\underline{\underline{A}}_{jk}) = \frac{a_i a_k}{a^2} = \underline{\underline{A}}_{ik}\,.$$

Every projector of the type $L = \vec{d}\overleftarrow{d}/d^2$ has this property.

(b) For a composition of two projectors

$$\widehat{L} = \widehat{\underline{\underline{B}}} \circ \widehat{\underline{\underline{A}}}$$

the components are given by

$$L_{ik} = \lambda\, a_i b_k\,, \qquad \lambda = \frac{\vec{a}\cdot\vec{b}}{a^2 b^2}\,.$$

Composition of two such operators can be evaluated using their components:

$$L_{ik}L_{kj} = (\lambda\, a_i b_k)(\lambda\, a_k b_j) = \lambda^2 a_i (a_k b_k) b_j = \lambda^2 (\vec{a}\cdot\vec{b}) a_i b_j\,.$$

We thus showed that

$$\widehat{L}\circ\widehat{L} = \lambda(\vec{a}\cdot\vec{b})\widehat{L}\,.$$

Now

$$\lambda(\vec{a}\cdot\vec{b}) = \frac{(\vec{a}\cdot\vec{b})^2}{a^2 b^2} = \cos\theta\,,$$

where θ is the angle between the vectors $\vec{a}$ and $\vec{b}$. Therefore,

$$\widehat{L} \circ \widehat{L} = \cos\theta \widehat{L} \neq \widehat{L} \,.$$

Only for $\vec{a} = \vec{b}$ we have the property $\widehat{L} \circ \widehat{L} = \widehat{L}$.

Exercise 5.6

The components of the composition

$$\widehat{L} = \underline{\underline{\widehat{B}}} \circ \underline{\underline{\widehat{A}}} \,.$$

are

$$L_{ik} = \lambda\, a_i b_k \,, \qquad \lambda = \frac{\vec{a} \cdot \vec{b}}{a^2 b^2} \,.$$

Reversing the order of arguments of the composition results in

$$\widehat{M} = \underline{\underline{\widehat{A}}} \circ \underline{\underline{\widehat{B}}} \,,$$

with the components

$$M_{ik} = \lambda\, b_i a_k \,, \qquad \lambda = \frac{\vec{b} \cdot \vec{a}}{a^2 b^2} \,.$$

In general, $M_{12} \neq L_{12}$ because $a_1 b_2 \neq b_1 a_2$.

Exercise 5.7

To arrive at the transformation rules for the components of different types of tensors, we will use a simple fact: A general tensor with components t^{ij} will behave just like the tensor product of two contra-variant vectors $a^i b^j$. Thus, we will study four types of tensor products:

$$a^i b^j \,, \quad a_i b_j \,, \quad a^i b_j \,, \quad a_i b^j \,.$$

It will be helpful to recall the transformation rules of contravariant and covariant vectors. An arrow-like contravariant vector can be expanded in a basis:

$$\vec{a} = a^i \vec{e}_i \,.$$

Every vectors from a "new" basis can be similarly expanded:

$$\vec{e}\,'_j = E_{j \bullet}^{\bullet k} \vec{e}_k \,.$$

Here we deliberately included the "dummy" symbol "$\bullet$" to visually align the indices according to their order. Expanding the same vector $\vec{a}$ relative to the "new" basis has the form:

$$\vec{a} = a'^i \vec{e}\,'_i \,,$$

and similarly

$$\vec{a} = a^i \vec{e}_i \,.$$

$$\vec{e}_k = (E')_{k \bullet}^{\bullet j} \vec{e}_j \,.$$

We also obtained the relation between the components of the same vector $\vec{a}$ in different bases:

$$a'^i = (E')_{k \bullet}^{\bullet i} a^k \,.$$

Another contravariant vector will have similar relations:

$$b'^j = (E')_{l\bullet}^{\bullet j} b^l \,,$$

and their tensor product $t^{ij} = a^i b^j$ will be transformed according to

$$(t')^{ij} = (E')_{k\bullet}^{\bullet i} (E')_{l\bullet}^{\bullet j} t^{kl} \,.$$

Using the transformation rule of covariant components:

$$a'_i = E_{i\bullet}^{\bullet k} a_k \quad \text{and} \quad b'_j = E_{j\bullet}^{\bullet l} a_l \,,$$

we immediately arrive at transformation rule of the components of a doubly-covariant tensor:

$$t'_{ij} = E_{i\bullet}^{\bullet k} E_{j\bullet}^{\bullet l} t_{kl} \,.$$

In a similar fashion, by combining the transformation rules of contravariant and covariant vectors, we can obtain the transformation of contra-covariant tensor:

$$(t')^{i\bullet}_{\bullet j} = (E')_{k\bullet}^{\bullet i} E_{j\bullet}^{\bullet l} t^{k\bullet}_{\bullet l} \,,$$

and covariant-contravariant tensor:

$$(t')_{i\bullet}^{\bullet j} = E_{i\bullet}^{\bullet k} (E')_{l\bullet}^{\bullet j} t_{k\bullet}^{\bullet l} \,.$$

Exercise 6.1

(a) The action of the metric tensor on a tensor product $\vec{a} \otimes \vec{b}$ can be understood once we write it out using components. First, let's write the expression for the distance squared:

$$d^2 = \eta_{ij} d^i d^j \,.$$

In a similar way, we can write

$$\widehat{\eta}(\vec{a} \otimes \vec{b}) = \eta_{ij} a^i b^j \,.$$

Using the fact

$$\eta_{ij} = \widehat{\sigma}\, \vec{e}_i\, \vec{e}_j$$

and bilinear nature of the dol-operator $\widehat{\sigma}$, we deduce

$$\eta_{ij} a^i b^j = a^i b^j (\widehat{\sigma}\, \vec{e}_i\, \vec{e}_j) = \widehat{\sigma}\,(a^i \vec{e}_i)\,(b^j \vec{e}_j) = \widehat{\sigma} \vec{a}\, \vec{b} \,.$$

Thus, we showed that

$$\widehat{\eta}(\vec{a} \otimes \vec{b}) = \vec{a} \cdot \vec{b} \,.$$

The connection between the metric tensor and scalar products is not surprising. Indeed, if we recall that if vector $\vec{d}$ connects two points in a plane and is given by

$$\vec{d} = \vec{b} - \vec{a} \,, \quad d^i = b^i - a^i \,,$$

then

$$\eta_{ij} d^i d^j = \eta_{ij} (a^i - b^i)(a^j - b^j) = \eta_{ij} a^i a^j + \eta_{ij} b^i b^j + 2\eta_{ij} a^i b^j \,.$$

We arrived at the familiar theorem of planar geometry − theorem of cosine:

$$d^2 = a^2 + b^2 - 2ab\cos\theta\,.$$

(b) Vectors of orthonormal basis all have unit lengths (*normalized* vectors):

$$e_i^2 = \widehat{\sigma}\,\vec{e}_i\,\vec{e}_i = \eta_{ii} = 1\,.$$

Different vectors of orthonormal basis are perpendicular to each other (*orthogonal* vectors):

$$\vec{e}_i \cdot \vec{e}_j = \widehat{\sigma}\,\vec{e}_i\,\vec{e}_j = \eta_{ij} = 0\,.$$

We showed that

$$\eta_{ij} = \delta_{ij}$$

when the basis is orthonormal.

9. Solutions

Exercise 1.1

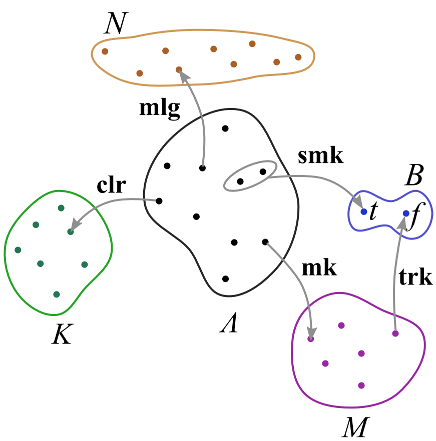


Fig. 9.1: The set M contains all possible makes of cars: Ford, Toyota, etc.

The diagram in the Figure 9.1 shows the set M – the set of all possible makes of cars. A mapping **trk** returns $true$ if a given car maker produces trucks.

Exercise 2.1

Any binary function can be viewed as a unary function if two inputs are replaced by a single input of a *pair of numbers*. Similarly for a function with two outputs. This idea is illustrated in the Figure 9.2(a): The function **swp** is viewed as a unary function which swaps the numbers in an *ordered pair*:

$$\mathbf{swp}\ (n, m) = (m, n)\,.$$

Given the set $\mathbb{Z}$ of whole numbers, we can create the set of all possible *ordered pairs* (n, m). This set can be denoted as follows:

$$(\mathbb{Z}, \mathbb{Z}) \text{ or } \mathbb{Z} \times \mathbb{Z}\,.$$

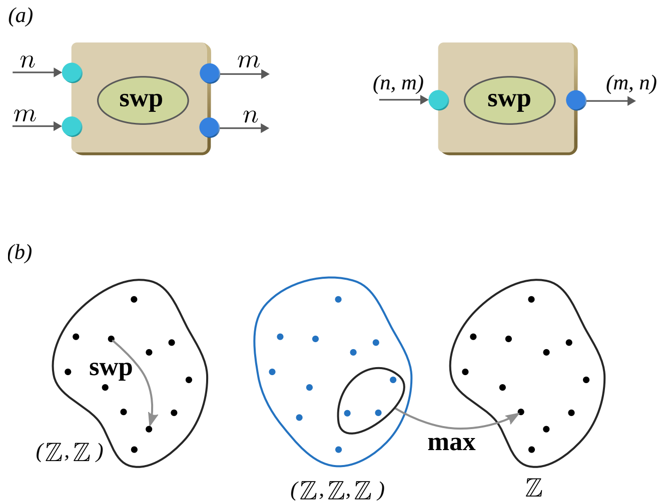


Fig. 9.2: (a) Two inputs (outputs) of a function can be replaced with a single input of a ***pair*** of numbers, turning a binary function into a unary one. (b) That.

The latter notation is standard in mathematics, but the former way of writing is also acceptable. We can similarly denote the set of all *ordered triples*:

$$(\mathbb{Z}, \mathbb{Z}, \mathbb{Z}) \text{ or } \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}\,.$$

With the notation introduced above, the action of functions with multiple inputs or outputs can be depicted on the level of sets. The Figure 9.2(b) shows how this works for the functions **swp** and **max** .

Index