Continuum Mechanics

by

Donald E. Carlson

1991

These notes are intended to be a textbook for an introductory but careful study of modern continuum mechanics at the advanced undergraduate or beginning graduate level. They are directed at American engineering students who have been exposed to the mathematics courses (calculus of one and several real variables, matrix algebra, ordinary differential equations) and the engineering mechanics courses (statics and dynamics of particles and rigid bodies, elementary mechanics of fluids and deformable solids) of the traditional curriculum. However, with respect to mechanics, the notes are completely self-contained; and, consequently, they should be accessible to students of mathematics and the sciences.

Contents

V

Chapter A8. Isotropic Tensors and Isotropic Tensor Functions

Part B. General Continuum Mechanics

Chapter B1. Kinematics

Part C. Examples of Simple Materials

Chapter C1. Some Fluids

Chapter C2. Elasticity

Special Symbols and Notations	Page
$a \in A$	A1.1.1
$\mathbb{R}$	"
$A = B$	A1.1.2
$A \subset B$, $B \supset A$	"
$\Longrightarrow$	"
$/$	"
iff	A1.1.3
$\Longleftrightarrow$	A1.1.4
$\square$	A1.1.5
$\{a, b, c\}$, $\{a\}$, $\{x : P(x)\}$	"
$\{x \in A : P(x)\}$	A1.1.6
$A \cup B$, $A \cap B$	"
ϕ	A1.1.7
w.r.t.	A1.1.8
$A - B$	A1.1.11
$(a_1, a_2, \cdots, a_n)$, $(a_1, a_2, \cdots, a_n) = (b_1, b_2, \cdots, b_n)$	A1.1.12
$\forall$	"
$A \times B$	A1.1.14
$\exists$, $\ni$	A1.2.2
$(x, y) \in f$, $y = f(x)$, $f : X \to Y$	A1.2.3
$x \mapsto f(x) = y$	A1.2.4
$f(S)$	"
f^{-1}	A1.2.7
$:=$, $=:$	"
$\therefore$	A1.2.8
$g \circ f$	A1.2.12

b

$f\vert_A$	A1.2.14
$a * b$, n , $\overset{r}{\hat{a}}$	A1.3.1
$\mathbb{R}^{\times}$	A1.3.3
$\underset{\sim}{u}+\underset{\sim}{v}$, $\underset{\sim}{0}$, $-\underset{\sim}{u}$	A2.1.5
$\alpha\underset{\sim}{u}$	A2.1.6
$\mathbb{F}$, $\mathbb{C}$	"
$\mathbb{R}^n$	A2.1.7
$[a,b]$	A2.1.11
$C^0([a,b])$	"
$\underset{\sim}{u}-\underset{\sim}{v}$	A2.1.12
$Lsp\,\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$	A2.1.22
$\underset{\sim}{u}\cdot\underset{\sim}{v}$, $\langle \underset{\sim}{u}, \underset{\sim}{v}\rangle$	A2.3.3
$\vert\underset{\sim}{u}\vert$	A2.3.10
$\sqrt{x}$	"
$\Vert\underset{\sim}{u}\Vert$	A2.3.14
$\mathbb{R}^+$	"
δ_{ij}	A2.3.22
$\{\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle n\rangle}\}$, $u^{\langle i\rangle}$	A2.3.25
$Lin(\mathcal{U}, \mathcal{V})$	A2.4.3
$\mathcal{V}^*$	A2.5.1
$\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$	A2.5.5
u_i	A2.5.10
$\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$	A2.2.5
$\mathcal{V}_n$	A2.2.8
u^i	A2.2.12
g_{ij} , g^{ij} , g^i_j	A2.5.13
α^i_j , β^i_j , $\bar{\underset{\sim}{e}}_i$, $\bar{\underset{\sim}{e}}^i$	A2.6.1
$C^1([a,b])$	A2.4.2a

C

$\mathbb{I}_n$	A2.7.4
$\varepsilon_{\sigma_1\sigma_2\cdots\sigma_n}$	A2.7.5
Δ, $\Delta(\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_n)$	"
$\mathbf{u}\times\mathbf{v}$	A2.8.2
$\mathbf{T}$	A3.1.1
$\mathbf{S}=\mathbf{T}$, $\mathbf{S}+\mathbf{T}$, $\alpha\mathbf{T}$, $\mathbf{0}$	A3.1.4
$\mathrm{Lin}\,\mathcal{V}$	A3.1.6
$\mathbf{I}$	"
$\mathbf{u}\otimes\mathbf{v}$	A3.1.7
T^{ij}, T_{ij}, $T^{i}_{\cdot j}$, $T_{i}^{\cdot j}$	A3.2.4
$\{\mathbf{e}_i\otimes\mathbf{e}_j\}$, $\{\mathbf{e}^i\otimes\mathbf{e}^j\}$, $\{\mathbf{e}_i\otimes\mathbf{e}^j\}$, $\{\mathbf{e}^i\otimes\mathbf{e}_j\}$	"
$\mathrm{Bd}(\mathcal{U},\mathcal{V})$	A3.4.5
glb A	A3.4.6
lub B	A3.4.7
$\Vert\mathbf{L}\Vert$	"
$\mathbf{T}\mathbf{S}$	A3.5.1
$\mathbf{u}\odot\mathbf{v}$, $\mathbf{i}$	A3.5.4
$\mathbf{T}^T$	A3.6.1
sym $\mathbf{T}$, skw $\mathbf{T}$	A3.6.7a
Sym $\mathcal{V}$, Skw $\mathcal{V}$	A3.6.9
ax $\mathbf{W}$, ax $\mathbf{w}$	A3.6.11
$\mathbf{u}\wedge\mathbf{v}$	A3.6.13
tr $\mathbf{T}$	A3.7.1
$\mathbf{S}\cdot\mathbf{T}$	A3.7.4
$\lvert\mathbf{T}\rvert$	A3.7.5
det $\mathbf{T}$	A3.8.1
$\mathbf{T}^{-1}$	A3.9.1
$\mathcal{N}(\mathbf{T})$	A3.9.3

$\mathrm{Inv}\, \mathcal{V}$	A3, 9.10
$\mathrm{Orth}\, \mathcal{V}_n$, $\mathrm{Orth}^+ \mathcal{V}_n$	A3.10.7
$(\lambda, \underset{\sim}{u})$	A3.11.1
$\iota_i(\underset{\sim}{T})$	A3.11.6
$\mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)$	A3.11.7
$\mathcal{U} + \mathcal{V}$	A3.12.1
$\mathcal{U} \oplus \mathcal{V}$	A3.12.2
$\mathcal{S}^{\perp}$	A3.12.6
$\mathrm{Psym}\, \mathcal{V}$	A3.15.2
$\sqrt{\underset{\sim}{S}}$	A3.15.3
$d(x, y)$	A5.1, 4
$B_r(x_0)$	A5.1, 6
$\mathcal{S}$	A5.1.10
$o(\underset{\sim}{u})$	A5, 2.3
$D\underset{\sim}{f}(\underset{\sim}{x})$, $D\underset{\sim}{f}(\underset{\sim}{x})[\underset{\sim}{u}]$	A5, 2.6
$\delta \underset{\sim}{f}(\underset{\sim}{x}; \underset{\sim}{u})$	A5, 2.8
$\underset{\sim}{f}'(x)$	A5, 3.1
$\nabla f(\underset{\sim}{x})$	A5, 4.1
$D^2 \underset{\sim}{f}(\underset{\sim}{x})$	A5, 10.1
$\underset{\sim}{u} \otimes \underset{\sim}{v}$ for $\underset{\sim}{u} \in \mathcal{U}$, $\underset{\sim}{v} \in \mathcal{V}$	A5, 10, 2
$D^2 \underset{\sim}{f}(\underset{\sim}{x})[\underset{\sim}{u}, \underset{\sim}{v}]$	A5.10, 3
$CP(\theta, \mathcal{V})$	A5.10.5
$D_i \underset{\sim}{f}(\underset{\sim}{x})$	A5.11.5
$\underset{\sim}{f}_{,i}(\underset{\sim}{x})$	A5.11.12
$\underset{\sim}{f}(\underset{\sim}{x}) = \hat{\underset{\sim}{f}}(\hat{\underset{\sim}{x}})$	A5, 13.2
$\hat{\underset{\sim}{f}}(\underset{\sim}{x}) = \sum_{j=1}^{n} \hat{f}^j(\hat{\underset{\sim}{x}})\, \underset{\sim}{e}_j$	A5, 13.7

e

$\mathcal{A} = \{\mathbb{p}, \mathbb{q}, \cdots\}$ A6.1.1

$\overrightarrow{\mathbb{p}\mathbb{q}} = \mathbb{q} - \mathbb{p}$ A6.1.2

$\mathbb{p} + \underset{\sim}{v}$ A6.1.3

$\{\mathbb{o}; \underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ A6.2.2

$d(\mathbb{p}, \mathbb{q})$ A6.3.1

$\mathcal{E}$ A6.3.2

$B_r(\mathbb{q})$ A6.3.3

$\bar{A}$ A6.3.5

∂A A6.3.5

$Df(x)$, $D\underset{\sim}{f}(x)[\underset{\sim}{u}]$ A6.4.2

$\underset{\sim}{f}(x) = \sum_{i=1}^{m} \hat{f}^{\langle i \rangle}(x^1, x^2, \cdots, x^n)\, \underset{\sim}{e}_{\langle i \rangle}$ A6.4.4

$\operatorname{div} \underset{\sim}{v}(x)$, $\operatorname{curl} \underset{\sim}{v}(x)$ A6.6.3

$\operatorname{div} \underset{\sim}{T}(x)$ A6.7.3

$D\mathbb{p}(x)$, $D\mathbb{p}(x)[\underset{\sim}{u}]$ A6.8.1

Part A

Mathematical Preliminaries

This long first part provides the additional mathematical background needed for our study. Some of the points of view adopted here are more common in mathematics than in engineering, but they are particularly useful in continuum mechanics. Keeping the mathematical preliminaries separate from the development of the mechanics has the advantage of permitting a cleaner treatment, and it also serves to emphasize that these mathematical notions are very general and are applicable to areas other than continuum mechanics. However, so as not to unduly delay reaching continuum mechanics proper, the study of some of the sections could be postponed until just before they are needed. E.g., Curvilinear coordinate systems are not employed until the development of particular solutions in Part C. Such matters will be pointed out at the appropriate places in the text.

Chapter A1

Some Basic Concepts

A1.1. Sets and Lists

A *set* is a collection of objects viewed as a single entity. A bushel of apples, a herd of cows, the positive integers, the real numbers, the exercises of Calculus III are all examples of sets. The objects in the collection are called the *elements* or *members* of the set, and the notation

$$a \in A$$

is used to indicate that a is an element of set A. E.g., $3/2$ is an element of the set of real numbers $\mathbb{R}$, and we write

$$3/2 \in \mathbb{R}.$$

It is important to realize that the above remarks do not constitute an adequate definition of set unless we know the meaning of collection and object. Here, we follow the tradition of taking the notion of set to be a *primitive* concept; i.e., an undefined but intuitively natural concept. Collection and object are just everyday nontechnical words that help us elucidate the primitive concept of set. Once primitive notions have been laid down, then mathematics customarily proceeds more formally through precise definitions and theorems in terms of the primitive concepts. This will be our approach throughout these notes.

Definition A1.1.1. Two sets A and B are said to be *equal*, and we write $A = B$, if they contain the same elements.

Here, and subsequently, definitions are understood as "if and only if" statements. Thus, the above definition says that if the sets A and B contain the same elements, then $A = B$, and conversely if $A = B$, then A and B contain the same elements.

Definition A1.1.2. A set A is said to be a *subset* of a set B, and we write $A \subset B$ or $B \supset A$, if every element of A is also an element of B; i.e., if

$$a \in A \Rightarrow a \in B.$$

If $A \subset B$ but $A \neq B$, then A is a *proper subset* of B.

As usual, the symbol $\Rightarrow$ stands for "implies", and the slash / is used for negation; thus, $A \neq B$ means that the sets A and B are not equal.

The following theorem provides the standard tool for proving that two sets are equal.

Theorem A1.1.1. Let A and B be sets. Then $A = B$ iff both $A \subset B$ and $B \subset A$.

Here, iff is an abbreviation for "if and only if". Thus, the theorem really consists in the two statements:

(i) $A = B$ if both $A \subset B$ and $B \subset A$;

(ii) $A = B$ only if both $A \subset B$ and $B \subset A$.

The first statement is straightforward, although, at least for me, it reads more smoothly in the form

(i)' If both $A \subset B$ and $B \subset A$, then $A = B$.

The second statement is more subtle. It means that the truth of $A = B$ requires that both $A \subset B$ and $B \subset A$. Thus, in phrasing parallel to that used in (i)', statement (ii) can be cast as

(ii)' If $A = B$, then both $A \subset B$ and $B \subset A$.

In the terminology of elementary logic, statement (ii)' is the converse of statement (i)', and vice versa.

Common alternative ways of stating Theorem A1.1.1 would be

$A = B$ is equivalent to both $A \subset B$ and $B \subset A$;

Necessary and sufficient for $A = B$ is that $A \subset B$ and $B \subset A$;

$A = B \Leftrightarrow$ both $A \subset B$ and $B \subset A$.

Of course, the symbol $\Leftrightarrow$ is a generalization of $\Rightarrow$; it is read as "implies and is implied by".

Since Theorem A1.1.1 provides a condition which is equivalent to $A = B$, it is often taken as the definition of $A = B$. In such an approach, our Definition A1.1.1 would then be a theorem. We shall always attempt to adopt the definitions that are the most natural rather than those which lead to the most efficient development. Of course, considerable subjectivity is involved in such decisions.

Before going on, we must not forget to turn to the

Proof of Theorem A1.1.1. Suppose that both $A \subset B$ and $B \subset A$. Since $A \subset B$, it follows from Definition A1.1.2 that every element of A is also in B, but since $B \subset A$, every element of B is also in A. Thus, there are no elements of B that are not also in A. In other words, A and B contain exactly the same elements. Therefore, $A = B$ by Definition A1.1.1.

Conversely, suppose that $A = B$. Consider a typical element

$a \in A$. Since $A = B$, A and B contain the same elements. Therefore, we also have $a \in B$. Thus, $a \in A \Rightarrow a \in B$; i.e., $A \subset B$ by Definition A1.1.2. In the same manner, we see that $B \subset A$. Hence, $A = B \Rightarrow$ both $A \subset B$ and $B \subset A$. □

The symbol □ is used here to indicate that the end of a proof has been reached.

Sets which contain just a few elements are usually indicated by listing all of their elements between between braces; e.g.,

$$\{a, b, c\} \,.$$

No significance is attached to the order in which the elements are listed. A set which contains a single element, say $\{a\}$, is called a *singleton*.

In more complicated cases, particular sets are specified by their elements possessing some defining property. The notation

$$\{x : P(x)\} \qquad)^1$$

stands for the set of all elements x which have the property

---

[1] Often, $\{x \mid P(x)\}$.

$P(x)$. E.g.,

$$\{x : x \in \mathbb{R},\ |x| < 1\}$$

is the set of all real numbers with absolute value less than 1. Sometimes it is advantageous to delimit the possibilities for candidacy in a set at the outset. The notation

$$\{x \in A : P(x)\}$$

stands for the set which consists in all of those elements x in A which possess the property $P(x)$. Thus, our previous example could have been expressed as

$$\{x \in \mathbb{R} : |x| < 1\}.$$

Next we turn to the common operations with sets.

Definition A1.1.3. The union of two sets A and B is the set

$$A \cup B = \{x : x \in A \text{ or } x \in B\}.$$

We will always use the conjunction or in its "weak" sense. Thus, in the above definition, we could have both $x \in A$ and $x \in B$.

Definition A1.1.4. The intersection of two sets A and B is the set

$$A \cap B = \{x : x \in A \text{ and } x \in B\}.$$

Of course, it could happen that the sets A and B would have no elements in common; then their intersection would be a set with no elements. This possibility leads us to the next two definitions.

Definition A1.1.5. <u>The set with no elements is called the</u> *empty set*[1]. <u>It is denoted by</u> ϕ.

Definition A1.1.6. <u>Let</u> A <u>and</u> B <u>be sets. If</u>

$$A \cap B = \phi,$$

<u>then</u> A <u>and</u> B <u>are said to be</u> *disjoint* <u>or</u> *nonintersecting*.

The next theorem gives the most important properties of the set operations of union and intersection.

Theorem A1.1.2. <u>Let</u> A, B, <u>and</u> C <u>be sets. Then we have the</u> *idempotent* <u>properties</u>

$$\text{(i)} \quad A \cup A = A, \quad A \cap A = A;$$

<u>the</u> *commutative* <u>properties</u>

---

[1] Often, the *void set* or the *vacuous set*.

(ii) $A \cup B = B \cup A$, $A \cap B = B \cap A$;

the associative properties

(iii) $A \cup (B \cup C) = (A \cup B) \cup C$, $(A \cap B) \cap C = A \cap (B \cap C)$;

and the distributive properties[1]

(iv) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Proof. Properties (i) and (ii) are so obvious that we do not supply formal proofs of them here.

To establish $(iii)_1$, we first suppose that $x \in (A \cup B) \cup C$. Then by repeated use Definition A1.1.3, we have $x \in (A \cup B)$ or $x \in C$, so that $x \in A$ or $x \in B$ or $x \in C$; consequently, $x \in A$ or $x \in (B \cup C)$, which means that $x \in A \cup (B \cup C)$. Therefore, by Definition A1.1.2, $(A \cup B) \cup C \subset A \cup (B \cup C)$. In the same way, we see that $A \cup (B \cup C) \subset (A \cup B) \cup C$. Hence, by Theorem A1.1.1, $A \cup (B \cup C) = (A \cup B) \cup C$.

The proof of $(iii)_2$ is left as Exercise A1.1.1.

The proofs of the distributive properties are a little more

---

[1] More precisely, $(iv)_1$ says that the operation of union is distributive w.r.t. the operation of intersection. Similarly, for $(iv)_2$. Here, w.r.t. stands for "with respect to".

difficult. Let us consider (iv)$_1$. Suppose that $x \in A \cup (B \cap C)$. Then by Definition A1.1.3, $x \in A$ or $x \in B \cap C$, so that by Definition A1.1.4, $x \in A$ or ($x \in B$ and $x \in C$). Note that the use of the parenthesis or some such device is essential here if we are to say what we mean. Next we examine the two possibilities at which we have arrived.

If $x \in A$, then it is correct to say that $x \in A \cup B$. This is not a natural step, because there is a loss of information; nonetheless, it is correct. Similarly, $x \in A \Rightarrow x \in A \cup C$. Putting these implications together, we have that $x \in A \Rightarrow x \in (A \cup B) \cap (A \cup C)$.

If $x \in B$ and $x \in C$, then $x \in A \cup B$ and $x \in A \cup C$; and these last two results $\Rightarrow x \in (A \cup B) \cap (A \cup C)$.

Thus, both of the possibilites under consideration lead to the same conclusion, and we have that

$$x \in A \cup (B \cap C) \Rightarrow x \in (A \cup B) \cap (A \cup C),$$

or

$$A \cup (B \cap C) \subset (A \cup B) \cap (A \cup C)$$

by Definition A1.1.2.

Next suppose that $x \in (A \cup B) \cap (A \cup C)$. In accordance with Definitions A1.1.3 and 4, this means that $x \in (A \cup B)$ and $x \in (A \cup C)$, which is to say that ($x \in A$ or $x \in B$) and

$(x \in A$ or $x \in C)$. Let us write out all of the possibilities:

$$x \in A \text{ and } x \in A \Rightarrow x \in A \cup (B \cap C);$$

$$x \in A \text{ and } x \in C \Rightarrow x \in A \cup (B \cap C);$$

$$x \in B \text{ and } x \in A \Rightarrow x \in A \cup (B \cap C);$$

$$x \in B \text{ and } x \in C \Rightarrow x \in B \cap C \Rightarrow x \in A \cup (B \cap C).$$

Again all of the possibilities lead to the same conclusion, and we have that

$$x \in (A \cup B) \cap (A \cup C) \Rightarrow x \in A \cup (B \cap C),$$

or

$$(A \cup B) \cap (A \cup C) \subset A \cup (B \cap C)$$

by Definition A1.1.2. Therefore, by Theorem A1.1.1,

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

The proof of $(iv)_2$ is left as Exercise A1.1.2. □

In view of the associative properties of union and intersection, we can drop the parentheses and write

$$A \cup B \cup C \quad \text{and} \quad A \cap B \cap C$$

without serious ambiguity.

On occasion we shall need to remove the elements of one set from another set. This leads to

Definition A1.1.7. The *set-difference* of two sets A and B is the set

$$A - B = \{x \in A : x \notin B\}. \qquad)^{1}$$

Of course, $x \notin B$ means that x is not an element of B.

The operation of set-difference has many interesting properties in conjunction with the operations of union and intersection, but we shall need only the definition.

In general, a set which consists of the two elements a and b would be denoted by $\{a, b\}$; and since no meaning is assigned to the order of the listing, $\{b, a\}$ is exactly the same set. However, on occasion, it is useful to order the elements, and this leads us to

---

[1] Usually, the set-difference $A-B$ is called the *complement of B relative to A* and is denoted by $A \setminus B$. Less often, it is denoted by $A \sim B$. Sometimes the term complement is restricted to the case where $B \subset A$.

the primitive concept of a "list".

Let n be a strictly positive integer.[1] Then a list of length n)[2] is an object denoted by

$$(a_1, a_2, \cdots, a_n)$$

which consists of a first element a_1, a second element a_2, $\cdots$, and an nth element a_n.

Definition A1.1.8. Two lists of order n $(a_1, a_2, \cdots, a_n)$ and $(b_1, b_2, \cdots, b_n)$ are equal, and we write

$$(a_1, a_2, \cdots, a_n) = (b_1, b_2, \cdots, b_n),$$

if

$$a_i = b_i \;\; \forall \; i \in \{1, 2, \cdots, n\}.$$

The symbol $\forall$ is read as "for all" or "for every". Of course, $a_i = b_i$ means that a_i and b_i are the same object.

Precisely because of the ordering built into the notion of a list, a list is not a set. Comparison of of the definitions of equality makes this especially

---

[1] Strictly positive means that 0 is not included.

[2] Often, n-tuple.

clear; e.g.,

$$\{\pi, e\} = \{e, \pi\} \quad \text{but} \quad (\pi, e) \neq (e, \pi).$$

When the length of a list is small[1], special terminology is customary.

Definition A1.1.9. Lists of length 2, 3, and 4 are called *ordered pairs*, *ordered triples*, and *ordered quadruples*, respectively.

Exercise A1.1.3. Is a list of length 1 a singleton?

See the book by Halmos included in the Supplementary Reading at the end of this section for an approach to lists which is entirely in terms of sets.

The following definition will play a major role in our treatment of functions in the next section.

Definition A1.1.10. The *set-product*[2] of two sets A and B

---

[1] You should resist the temptation to to speak of short lists, unless you know of some short integers.

[2] Almost always, *Cartesian product*; sometimes, *direct product*. This is the first instance of our general practice (cont.)

is the set of ordered pairs

$$A\times B = \{(a,b) : a\in A,\ b\in B\}.$$

Note that if either A or B is empty, then we cannot form any ordered pairs (a,b). Thus,

$$\phi\times B = A\times\phi = \phi.$$

---

of giving descriptive rather than proper names to concepts and theorems even when the proper name is used almost universally. In this regard, as well as many others, our thinking has been heavily influenced by the work of WALTER NOLL.

Supplementary Reading

BARTLE, The Elements of Real Analysis

BISHOP and GOLDBERG, Tensor Analysis on Manifolds

BOWEN and WANG, Introduction to Vectors and Tensors, Vol. 1, Linear and Multilinear Algebra
* (mechanics people)

HALMOS, Naive Set Theory

KOLMOGOROV and FOMIN, Introductory Real Analysis

LOOMIS and STERNBERG, Advanced Calculus

MICHEL and HERGET, Mathematical Foundations in Engineering and Science

NAYLOR and SELL, Linear Operator Theory in Engineering and Science

NOLL, Finite-Dimensional Spaces

ODEN, Applied Functional Analysis

SIMMONS, Introduction to Topology and Modern Analysis

A1.2. Functions

Beginning students usually think of "functions" as definite formulas relating real numbers, such as

$$y = f(x) = x^2 - 1 \quad \text{for } x \in \mathbb{R}.$$

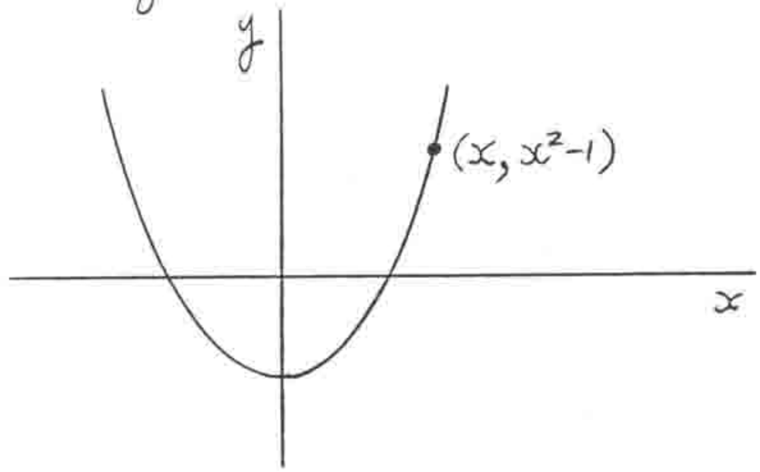
Of course, to limit oneself to real numbers and definite formulas is clearly too restrictive for the purposes of either mathematics or its applications.

Those who use mathematics typically generalize this along the following lines: "Given two sets A and B, a function f on A to B is a rule of correspondence that assigns to each $x \in A$ a unique element $f(x) \in B$." This definition works quite well, but it has the defect of being based on another primitive notion; namely, "rule of correspondence". Here, we have a dilemma. Obviously, we would be wise to minimize the number of primitive concepts introduced; indeed, it is the current fashion to found all of mathematics on sets. On the other hand, we would like to progress rapidly, perhaps even eventually getting to continuum mechanics, without being buried under a tangle of technical developments. Actually, we were in a similar position at the end of the previous section. There we chose to take the notion of list as a primitive concept, eventhough a set-theoretic definition was available. However, the concept of list is much more transparent than "rule of correspondence".

To gain some insight into a set-theoretic approach to "functions", we return to the pedestrian example mentioned at the beginning of the section:

$$y = f(x) = x^2 - 1 \quad , \ x \in \mathbb{R}.$$

The correspondence between x and y can be depicted by the "graph" of x^2-1.



But the graph is just a special set of ordered pairs, and thus we are lead to the following definition. In order to state it concisely, we first introduce the symbols $\exists$ and $\ni$, which are read as "there exists" and "such that", respectively.[1]

[1] Our use of $\ni$ for "such that" disallows the use of the symbol $\in$ in the backward fashion $A \ni a$, which is sometimes read as "the set A contains the element a".

Definition A1.2.1. Given two sets X and Y, a function[1] f on X to Y is a subset of $X \times Y$ with the property that

$$\text{for each } x \in X \; \exists \text{ a unique } y \in Y \ni (x,y) \in f.$$

The sets X and Y are called the domain and the codomain of the function f, respectively. For $x \in X$, the unique $y \in Y \ni (x,y) \in f$ is called the value of f at x; it is denoted by

$$y = f(x).$$

The uniqueness condition above is conveniently expressed as

$$(x,y) \in f \text{ and } (x,y') \in f \Rightarrow y' = y.$$

If this requirement is dropped, then we are left with the definition of a relation.

We shall often use the notation

$$f : X \to Y$$

in place of the statement "f is a function on X to Y".

---

[1] Often, map, mapping, transformation, operator.

Occasionally, the notation

$$x \mapsto f(x) = y$$

is a useful substitute for

$$(x,y) \in f \quad \text{or} \quad y = f(x).$$

Definition A1.2.2. Let $f: X \to Y$. If $S \subset X$, then the set

$$f(S) = \{f(x) : x \in S\}$$

is called the *direct image*[1] of S under f. The set $f(X)$ is called the *range* of f. If $f(X) = Y$, then f is said to be *onto*[2].

The context will usually prevent the confusion which is possible between function values and direct images.

Let us return to our earlier example of the function $f: R \to R$ defined by the rule

$$y = f(x) = x^2 - 1.$$

---

[1] Often, simply *image*.

[2] Often, *surjective*.

Is this really a function? Here, the domain and codomain each are $\mathbb{R}$ as claimed; indeed, if $x \in \mathbb{R}$, then $y = f(x) = x^2 - 1 \in \mathbb{R}$. To see that the given rule actually defines a function, we must establish the uniqueness condition that

$$(x,y) \in f \text{ and } (x,y') \in f \Rightarrow y' = y$$

or

$$y = f(x) \text{ and } y' = f(x) \Rightarrow y' = y$$

or

$$y = x^2 - 1 \text{ and } y' = x^2 - 1 \Rightarrow y' = y .$$

Since the meaning of $x^2 - 1$ is unambiguous, the last implication obviously is valid; and the rule $f(x) = x^2 - 1$ defines a function $f : \mathbb{R} \to \mathbb{R}$ in accordance with Definition A1.2.1. Of course, the actual function f is the set of ordered pairs $f = \{(x, x^2-1) : x \in \mathbb{R}\}$.

Since most functions in continuum mechanics are defined by such rules, we formalize the above result as

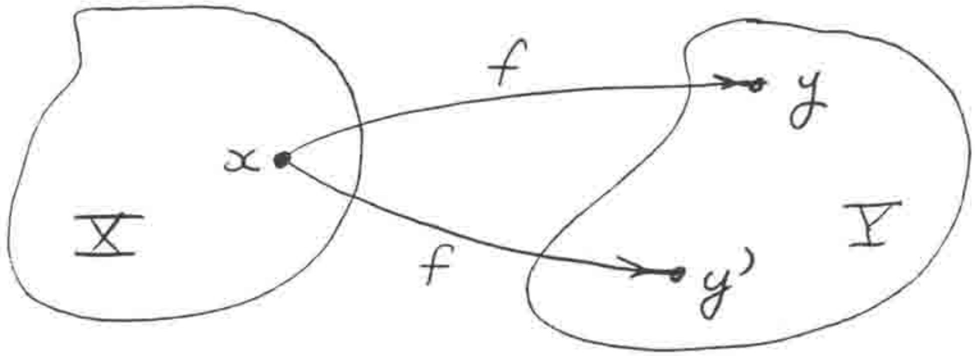
Theorem A1.2.1. <u>An unambiguous evaluation rule,</u>

$$\text{for } x \in X,\ y = f(x) \in Y,$$

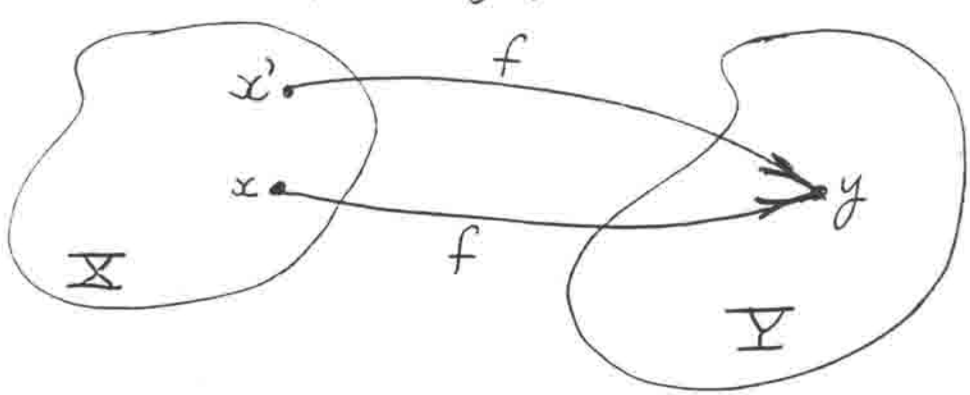
<u>defines a function</u> $f : X \to Y$ <u>through</u> $f = \{(x, f(x)) : x \in X\}$

Exercise A1.2.1. Is the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2 - 1$ <u>onto</u> $\mathbb{R}$?

By definition, as has been just emphasized, a function $f: X \to Y$ associates to a given element $x \in X$ a <u>unique</u> element $y = f(x) \in Y$; thus, the following picture <u>is</u> <u>not</u> possible.



But this does not rule out the possibility that y is also the value of f at some other element $x' \in X$; i.e., the following picture <u>is</u> possible.



For the important class of "one-to-one" functions, the situation pictured lastly above also is ruled out. Formally, we have

Definition A1.2.3. A function $f: \mathbb{X} \to \mathbb{Y}$ is said to be *one-to-one*[1] if $\forall\, x, x' \in \mathbb{X}$

$$f(x') = f(x) \Longrightarrow x' = x.$$

For one-to-one functions, it is possible to "run the functional correspondence backwards".

Theorem A1.2.2. If $f: \mathbb{X} \to \mathbb{Y}$ is one-to-one, then the set of ordered pairs

$$f^{-1} := \{ (y, x) \in f(\mathbb{X}) \times \mathbb{X} : (x, y) \in f \} \qquad)^2$$

is a function on $f(\mathbb{X})$ to $\mathbb{X}$. Moreover, f^{-1} is

---

[1] Often, *one-one*, 1-1, *injective*.

[2] The symbol $:=$ means that the left-hand side is defined by the right-hand side; similarly, for $=:$.

one-to-one and onto[1], and

$$f^{-1}(y) = x \iff f(x) = y .$$

The function f^{-1} is said to be the *inverse function* to the one-to-one function f.

Proof. Refer to Definition A1.2.1. First, we note that by its construction f^{-1} is a subset of $f(\mathfrak{X}) \times \mathfrak{X}$, and thus it is a candidate for a function on $f(\mathfrak{X})$ to $\mathfrak{X}$. We must show that

$$\text{for each } y \in f(\mathfrak{X}) \ \exists \text{ a unique } x \in \mathfrak{X} \ni (y, x) \in f^{-1}.$$

Accordingly, let $y \in f(\mathfrak{X})$. By Definition A1.2.2,

$$f(\mathfrak{X}) = \{ f(x) : x \in \mathfrak{X} \} ;$$

and $\therefore$)[2] $\exists$ an $x \in \mathfrak{X} \ni f(x) = y$. But this equation is just another way of stating that $(x, y) \in f$. Looking back to the definition of f^{-1}, we see that this proves that

$$\text{for each } y \in f(\mathfrak{X}) \ \exists \text{ an } x \in \mathfrak{X} \ni (y, x) \in f^{-1} .$$

---

[1] Functions that are one-to-one and onto, i.e., injective and surjective, are often said to be *bijective*.

[2] As always, the symbol $\therefore$ stands for "therefore".

The issue now is the uniqueness of x. Suppose we have both $(y,x) \in f^{-1}$ and $(y,x') \in f^{-1}$. By the construction of f^{-1}, this means that $(x,y) \in f$ and $(x',y) \in f$ or $y = f(x)$ and $y = f(x')$. Hence, $f(x') = f(x)$; but f is one-to-one, so by Definition A1.2.3, $x' = x$. Thus, x is unique, and f^{-1} is a function.

Next we turn to the proof of

$$f^{-1}(y) = x \iff f(x) = y .$$

Now that we know that f^{-1} is a function so that the notation $f^{-1}(y) = x$ makes sense, the double implication above is really tautologous with

$$f^{-1} = \{ (y,x) \in f(\mathfrak{X}) \times \mathfrak{X} : (x,y) \in f \} .$$

We have

$$f^{-1}(y) = x \Rightarrow (y,x) \in f^{-1} \Rightarrow (x,y) \in f \Rightarrow y = f(x)$$

and

$$f(x) = y \Rightarrow (x,y) \in f \Rightarrow (y,x) \in f^{-1} \Rightarrow x = f^{-1}(y) .$$

To prove that f^{-1} is one-to-one, we must show that $\forall\, y, y' \in f(\mathfrak{X})$

$$f^{-1}(y') = f^{-1}(y) \Rightarrow y' = y .$$

Accordingly, let y and y' be arbitrary elements of $f(\mathfrak{X})$, and set $f^{-1}(y) = x$ and $f^{-1}(y') = x'$. Then

$$f^{-1}(y') = f^{-1}(y) \Rightarrow x' = x \Rightarrow f(x') = f(x) \Rightarrow y' = y.$$

To prove that $f^{-1}: f(X) \to X$ is onto, we must show in accordance with Definition A1.2.2 that

$$f^{-1}(f(X)) = X.$$

By this same definition

$$f^{-1}(f(X)) = \{f^{-1}(y) : y \in f(X)\}.$$

The problem here is to show that two sets are equal; as usual, we use theorem A1.1.1. Suppose $x \in X$. Set $y = f(x)$. Then $y \in f(X)$ and $x = f^{-1}(y)$. Thus,

$$x \in \{f^{-1}(y) : y \in f(X)\} = f^{-1}(f(X));$$

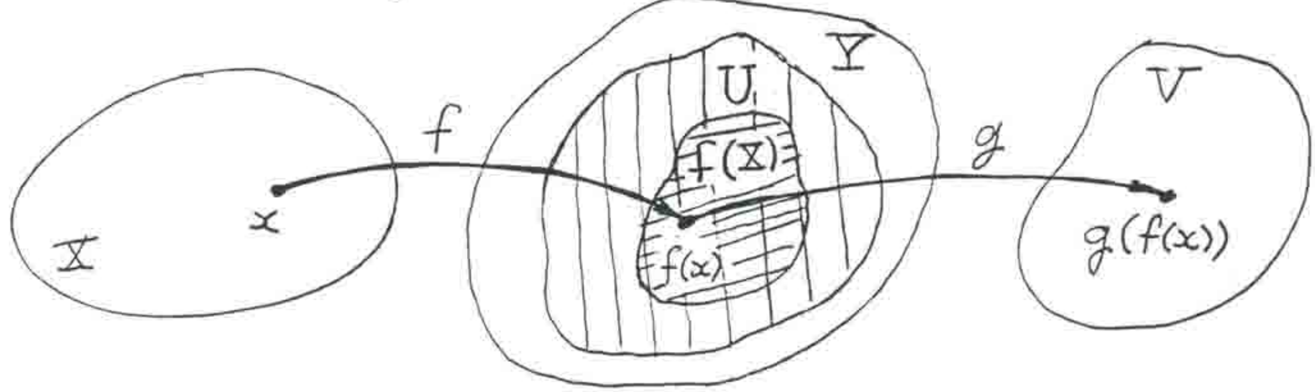
and since x is an arbitrary element of X, $X \subset f^{-1}(f(X))$. Next suppose $x \in f^{-1}(f(X))$. Since $f^{-1}(f(X)) = \{f^{-1}(y) : y \in f(X)\}$, $\exists$ some $y \in f(X) \ni$

$$x = f^{-1}(y) \Rightarrow (y, x) \in f^{-1} = \{(y, x) \in f(X) \times X : (x, y) \in f\} \Rightarrow x \in X,$$

so that $f^{-1}(f(X)) \subset X$. $\therefore f^{-1}(f(X)) = X$. $\square$

Exercise A1.2.2. Show that the function $f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2 - 1$ is not one-to-one.

Often in continuum mechanics it is necessary to "compose" functions in the following sense. Suppose we have two functions $f : \mathbb{X} \to \mathbb{Y}$ and $g : \mathbb{U} \to \mathbb{V}$. If $x \in \mathbb{X}$, then f maps x into some element $f(x)$ in the range of f. If $f(x)$ is in the domain $\mathbb{U}$ of g, then g maps $f(x)$ into some element of $\mathbb{V}$. Thus, we have a scheme for mapping from $\mathbb{X}$ to $\mathbb{V}$.



The following theorem makes this precise.

Theorem A1.2.3. Let $f : \mathbb{X} \to \mathbb{Y}$ and $g : \mathbb{U} \to \mathbb{V}$ with $f(\mathbb{X}) \subset \mathbb{U}$. Then

$$g \circ f := \{(x, v) : x \in \mathbb{X},\ v = g(f(x))\}$$

is a function on $\mathbb{X}$ to $\mathbb{V}$, and

$$(g \circ f)(x) = g(f(x)).$$

$g \circ f$ is called the *composition* of f and g.

Proof. *Exercise* A1.2.3. $\square$

We shall have need of

Theorem A1.2.4. The composition of functions is *associative*; i.e., let the domains, codomains, and ranges of the functions f, g, and h be such that the compositions $(h\circ g)\circ f$ and $h\circ(g\circ f)$ make sense, then

$$(h\circ g)\circ f = h\circ(g\circ f) \ .\ )^1$$

Proof. *Exercise* A1.2.4. □

Because of Theorem A1.2.4, we can write $h\circ g\circ f$ without serious ambiguity.

Exercise A1.2.6. Given a set X, the *identity map* on X is the function $i_X : X \to X$ defined by the evaluation rule

$$i_X(x) = x \ .\ )^2$$

For a one-to-one function $f: X \to Y$, show that

$$f^{-1}\circ f = i_X \quad \text{and} \quad f\circ f^{-1} = i_Y \ .$$

---

[1] Of course, the *equality* of two functions means that their respective domains and codomains are equal and that the defining subsets of ordered pairs (cf. Definition A1.2.1) are equal. It is left as *Exercise* A1.2.5 to show that (cont.)

Occasionally, we need to restrict our attention to a subset of the domain of a given function, and then the following definition will be useful.

Definition A1.2.4. Given a function $f: X \to Y$, let $A \subset X \subset B$. Then the restriction of f to A, denoted by $f|_A$, is the function $f|_A : A \to Y$ defined by the evaluation rule[1]

$$\text{for } x \in A, \ f|_A(x) = f(x);$$

and a function $g: B \to C$ is an extension of f to B if

$$g|_X = f.)^2$$

Supplementary Reading

Same as for § A1.1

---

[1] Cf. Theorem A1.2.1.

[2] This, of course, requires that $Y \subset C$.

(cont. from p. A1.2.13) two functions f and $\bar{f}$ on X to Y are equal iff $\bar{f}(x) = f(x) \ \forall x \in X$.

[2] Cf. Theorem A1.2.1.

A1.3. Groups

General theories are built up in mathematics by imposing additional structure on sets. The "group structure" is one which occurs repeatedly.

Definition A1.3.1. A group is a set, say $\mathcal{G}$, equipped with a function from $\mathcal{G}\times\mathcal{G}$ to $\mathcal{G}$, called combination and denoted by

$$(a,b) \longmapsto a*b,$$

which satisfies the following axioms:

(G1) Associativity. $\forall\, a,b,c \in \mathcal{G}$

$$(a*b)*c = a*(b*c)\ ;$$

(G2) Existence of a neutral element. $\exists\, n \in \mathcal{G} \ni$

$$n*a = a*n = a \quad \forall\, a \in \mathcal{G};$$

(G3) Existence of reverse elements. For each $a \in \mathcal{G}\ \exists$ $\overset{r}{a} \in \mathcal{G} \ni$

$$\overset{r}{a}*a = a*\overset{r}{a} = n\,.$$

If, in addition, the commutative property

$$a * b = b * a \quad \forall\, a, b \in \mathcal{G}$$

is satisfied, then $\mathcal{G}$ is said to be a commutative[1] group.

Since the combination function in the definition of a group may be thought of as taking two elements from $\mathcal{G}$ and producing an element also in $\mathcal{G}$, it is often referred to as a closed binary operation. The closure property, i.e., the fact that the function value $a*b$ is also in $\mathcal{G}$, is often called the fundamental group property.

There are many familiar examples of groups, but they are not always identified as such. The set of real numbers $\mathbb{R}$ is a group w.r.t. addition. Of course, in this concrete example, we write $a+b$ instead of $a*b$. I.e., the combination function under consideration is defined on $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$ by

$$(a, b) \longmapsto a + b.$$

Since this rule is unambiguous, the operation of addition in $\mathbb{R}$ really is a function by Theorem A1.2.1. What about the three axioms? It is a familiar property of the real numbers that

$$(a+b)+c = a+(b+c) \quad \forall\, a, b, c \in \mathbb{R},$$

---

[1] Sometimes, Abelian.

so associativity is satisfied. The number 0 is the neutral element; i.e.,

$$0+a = a+0 \quad \forall\, a \in \mathbb{R}.$$

The reverse elements are the negatives; i.e., given any $a \in \mathbb{R}$, the number $-a$ has the property

$$(-a) + a = a + (-a) = 0.$$

Thus, $\mathbb{R}$ is, indeed, a group w.r.t. addition; in fact, since

$$a+b = b+a \quad \forall\, a, b \in \mathbb{R},$$

it is a commutative group.

Exercise A1.3.1. Show that $\mathbb{R} - \{0\} =: \mathbb{R}^{\times}$ is a commutative group w.r.t. multiplication.

The following result is often useful in manipulations in a group. It is so transparent that usually it is not stated formally. It simply says that if equals are combined with equals the results are equal. We state it here in the language of Definition A1.3.1.

Theorem A1.3.1. Let a, b, c, and d be elements of a group. Then

$$a=b \text{ and } c=d \implies a*c = b*d.$$

Proof. We start with the identity

$$\begin{aligned} a * c &= a * c \\ &= b * c \quad (\text{substitution of } a = b) \\ &= b * d \quad (\text{substitution of } c = d). \quad \square \end{aligned}$$

Our next theorem asserts that the equation $a * x = b$ has the unique solution $x = \hat{a} * b$, with a similar result for $y * a = b$.

Theorem A1.3.2. Let $\mathcal{G}$ be a group. Then for any $a, b \in \mathcal{G}$, $\exists$ a unique $x \in \mathcal{G} \ni$

$$a * x = b.$$

In fact,

$$x = \hat{a} * b.$$

Similarly, $\exists$ a unique $y \in \mathcal{G} \ni$

$$y * a = b.$$

In fact,

$$y = b * \hat{a}.$$

Proof. Suppose, tentatively, that $\exists$ an $x \in \mathcal{G} \ni a * x = b$. Then combining on the left with $\hat{a}$, we have

$$\hat{a} * (a * x) = \hat{a} * b,$$

in accordance with Theorem A1.3.1. Now

$$\overset{r}{a} * (a * x) = (\overset{r}{a} * a) * x \quad \text{(associativity)}$$
$$= n * x \quad \text{(reverse)}$$
$$= x \quad \text{(neutral)},$$

and $\therefore$ $x = \overset{r}{a} * b$. Thus, if x exists, it must be precisely $\overset{r}{a} * b$. This establishes the uniqueness. To prove existence, we must show that this x has the desired property. First of all, we note that by closure $\overset{r}{a} * b \in \mathcal{G}$. Next, consider

$$a * (\overset{r}{a} * b) = (a * \overset{r}{a}) * b \quad \text{(associativity)}$$
$$= n * b \quad \text{(reverse)}$$
$$= b \quad \text{(neutral)}.$$

Hence, $x = \overset{r}{a} * b$, does the job.

The "y-problem" can be handled in the same way. Work this out as *Exercise* A1.3.2. $\square$

The next five results are all corollaries of Theorem A1.3.2. They also can be viewed as direct consequences of Theorem A1.3.1. We continue to use the language of Definition A1.3.1.

Theorem A1.3.3. (*Cancellation Property*) Let $\mathcal{G}$ be a group. Then for $a, b, c \in \mathcal{G}$

$$b * a = c * a \Rightarrow b = c$$

and

$$a * b = a * c \Rightarrow b = c.$$

Proof. Suppose $b * a = c * a$. We apply the second part of Theorem A1.3.2 with b playing the role of the unknown y and $c * a$ being the right-hand side b to get

$$
\begin{aligned}
b &= (c * a) * \overset{r}{a} \\
&= c * (a * \overset{r}{a}) && \text{(associativity)} \\
&= c * n && \text{(reverse)} \\
&= c && \text{(neutral).}
\end{aligned}
$$

Obviously, the second assertion can be handled in the same fashion. But, for diversity, we leave it for the student as Exercise A1.3.3 to establish it by combining $a * b = a * c$ on the left with $\overset{r}{a}$ in accordance with Theorem A1.3.1. $\square$

Theorem A1.3.4. <u>A given group has only one neutral element.</u>

Proof. Suppose n and n' are both neutral elements for a group $\mathcal{G}$. I.e., $n, n' \in \mathcal{G}$ and $\forall\, a \in \mathcal{G}$

$$n * a = a * n = a \quad \text{and} \quad n' * a = a * n' = a .$$

Consider $a * n' = a$. By Theorem A1.3.2, this $\Rightarrow$

$$\begin{aligned} n' &= \overset{r}{a} * a \\ &= n \quad (\text{reverse}). \ \square \end{aligned}$$

Theorem A1.3.5. <u>A given element of a group has only one reverse.</u>

Proof. Let $a \in \mathcal{G}$, where $\mathcal{G}$ is a group. Suppose $\overset{r}{a}$ and $\overset{*}{a}$ are both reverses of a. I.e., $\overset{r}{a}, \overset{*}{a} \in \mathcal{G}$ and

$$\overset{r}{a} * a = a * \overset{r}{a} = n \quad \text{and} \quad \overset{*}{a} * a = a * \overset{*}{a} = n .$$

Consider $a * \overset{*}{a} = n$. By Theorem A1.3.2, this $\Rightarrow$

$$\begin{aligned} \overset{*}{a} &= \overset{r}{a} * n \\ &= \overset{r}{a} \quad (\text{neutral}). \ \square \end{aligned}$$

Theorem A1.3.6. Let n be the identity element of a group. Then

$$\overset{r}{n} = n .$$

Proof. For any element a, the reverse axiom requires $a*\overset{r}{a} = n$. Taking $a = n$, we get $n*\overset{r}{n} = n$. By Theorem A1.3.2, this $\Rightarrow$

$$\begin{aligned} \overset{r}{n} &= \overset{r}{n}*n \\ &= n \quad (\text{reverse}). \quad \square \end{aligned}$$

Theorem A1.3.7. Let a be any element of a group. Then

$$\overset{r}{\overline{\overset{r}{a}}} = a .$$

Proof. The reverse axiom requires $\overset{r}{a}*a = n$. By Theorem A1.3.2, this $\Rightarrow$

$$\begin{aligned} a &= \overset{r}{\overline{\overset{r}{a}}}*n \\ &= \overset{r}{\overline{\overset{r}{a}}} \quad (\text{neutral}). \quad \square \end{aligned}$$

Another useful result is that the reverse of a combination is the combination of the reverses in the reverse order. Formally, we have

Theorem A1.3.8. Let a and b be any two elements of a group. Then

$$\overset{r}{\overline{a*b}} = \overset{r}{b}*\overset{r}{a} .$$

Proof. Consider

$$\begin{aligned}(a*b)*(\overset{r}{b}*\overset{r}{a}) &= a*(b*(\overset{r}{b}*\overset{r}{a})) && \text{(associativity)}\\ &= a*((b*\overset{r}{b})*\overset{r}{a}) && \text{(associativity)}\\ &= a*(n*\overset{r}{a}) && \text{(reverse)}\\ &= a*\overset{r}{a} && \text{(neutral)}\\ &= n && \text{(reverse)}.\end{aligned}$$

In the same way,

$$(\overset{r}{b}*\overset{r}{a})*(a*b) = n.$$

But these are the two properties that the reverse axiom requires. ∴ $\overset{r}{b}*\overset{r}{a}$ is <u>an</u> inverse element to $a*b$. In view of Theorem A1.3.5, it is <u>the</u> reverse. □

Since a group $\mathcal{G}$ is first of all a set, it is easy to consider subsets of $\mathcal{G}$. An important question is whether or not the subsets are themselves groups. In general, they are not. E.g., consider the subset $\mathcal{G} - \{n\}$. The following terminology will be useful in this regard.

Definition A1.3.2. Let $\mathcal{H}$ be a nonempty [proper] subset of a group $\mathcal{G}$. Then if $\mathcal{H}$ is a group w.r.t. the combination function of $\mathcal{G}$, it is called a [proper] subgroup of $\mathcal{G}$.

Several comments are in order here. The bracket device is a standard way to make two similar statements at once. To get the first statement, leave out the the bracketed material. To get the second statement, include the bracketed material. Since $\mathcal{H} \subset \mathcal{G}$, $\mathcal{H} \times \mathcal{H} \subset \mathcal{G} \times \mathcal{G}$, and there is no problem in evaluating the combination function of $\mathcal{G}$ at elements of $\mathcal{H}$. However, the function values, which necessarily belong to $\mathcal{G}$, are not automatically in $\mathcal{H}$; i.e., for an arbitrary subset, closure could fail.

Theorem A1.3.9. Let $\mathcal{G}$ be a group with neutral element n. Then the singleton $\{n\}$ is a subgroups of $\mathcal{G}$.

Proof. Exercise A1.3.4. $\square$

Exercise A1.3.5. Refer to Theorem A1.3.9. Is $\{n\}$ a <u>proper</u> subgroup of $\mathcal{G}$?

The following theorem provides a useful criterion sufficient for a subset to be a subgroup.

Theorem A1.3.10. <u>Let</u> $\mathcal{H}$ <u>be a nonempty subset of a group</u> $\mathcal{G}$. <u>If</u>

$$\forall a, b \in \mathcal{H}, \quad a * \overset{r}{b} \in \mathcal{H},$$

<u>then</u> $\mathcal{H}$ <u>is a subgroup of</u> $\mathcal{G}$.

Proof. As noted in the discussion following Definition A1.3.2, the combination function for $\mathcal{G}$ makes sense for $\mathcal{H}$. Moreover, the associativity requirement is $\therefore$ automatically met. Assume that

$$(*) \quad a, b \in \mathcal{H} \Rightarrow a * \overset{r}{b} \in \mathcal{H}.$$

Let $a \in \mathcal{H}$ and choose $b = a$. Then $(*) \Rightarrow a * \overset{r}{a} \in \mathcal{H}$. But $a * \overset{r}{a} = n$, the neutral element of $\mathcal{G}$. Thus, $n \in \mathcal{H}$.

Next, let $b \in \mathcal{H}$ and choose $a = n$. Then $(*) \Rightarrow n * \overset{r}{b} \in \mathcal{H}$. But $n * \overset{r}{b} = \overset{r}{b}$. Then $b \in \mathcal{H} \Rightarrow \overset{r}{b} \in \mathcal{H}$, so $\mathcal{H}$ contains the reverses of all of its elements. The reverses necessarily exist because of the properties of the "parent" group $\mathcal{G}$.

Finally, let $a, b \in \mathcal{H}$. Then by the preceding step, $\overset{r}{b} \in \mathcal{H}$. Letting $\overset{r}{b}$ play the role of b in $(*)$, we get

$$a * \overset{r}{\overset{r}{b}} \in \mathcal{H}.$$

But by Theorem A3.1.7, $\overset{r}{\overset{r}{b}} = b$. Thus, $a * b \in \mathcal{H}$, and we have that $\mathcal{H}$ is closed under the combination operation. $\square$

Supplementary Reading

BIRKHOFF and MAC LANE, A Survey of Modern Algebra

BOWEN and WANG, Introduction to Vectors and Tensors, Vol. 1, Linear and Multilinear Algebra

MICHEL and HERGET, Mathematical Foundations in Engineering and Science

MOSTOW, SAMPSON, and MEYER, Fundamental Structures of Algebra

NOLL, Finite-Dimensional Spaces

ODEN, Applied Functional Analysis

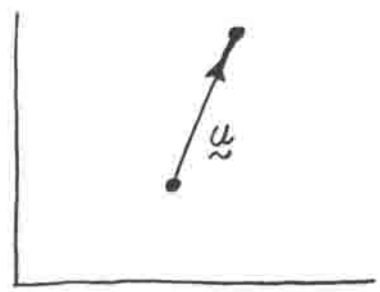
SIMMONS, Introduction to Topology and Modern Analysis

Chapter A2

Linear Spaces

A2.1. Definition of a Linear Space

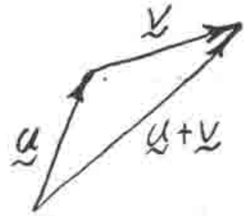
For the purpose of motivation, recall your experience with vectors in the Euclidean plane. Here a vector is just the directed line segment or arrow that goes from one point in the plane to another point in the plane.



We will follow the common practice in mechanics of denoting vectors by bold face lower case letters.

Two vectors in the plane are regarded as equal if they are parallel, point in the same direction, and have the same length. The actual location of a vector in the plane is not important, at least for most purposes.

The addition of vectors is defined by the parallelogram rule or the "tip to tail" method depicted below. Note that the sum is itself a vector; i.e., vector addition is closed.



Vector Addition

The addition of vectors is "immediately seen" to be <u>commutative</u> ($\underset{\sim}{u}+\underset{\sim}{v}=\underset{\sim}{v}+\underset{\sim}{u}$) and <u>associative</u> ($(\underset{\sim}{u}+\underset{\sim}{v})+\underset{\sim}{w}=\underset{\sim}{u}+(\underset{\sim}{v}+\underset{\sim}{w})$). The "proofs" of these properties are pictured next.

Commutative Property

Associative Property

In this scheme, any two distinct points in the plane define a vector. What if the two points are the same? This means the arrow collapses to a single point. We naturally think of this peculiar vector as having zero length, but we cannot assign it a direction. It is called the <u>zero vector</u> and is denoted by $\underset{\sim}{0}$. Because of its zero length, it clearly has the property that for any vector $\underset{\sim}{u}$, $\underset{\sim}{0}+\underset{\sim}{u}=\underset{\sim}{u}+\underset{\sim}{0}=\underset{\sim}{u}$. Thus, for vector addition, the zero vector plays the role of a neutral element.

Given any nonzero vector $\underset{\sim}{u}$, we can generate a new vector by simply reversing the direction of $\underset{\sim}{u}$. This new vector is called the <u>negative</u> of $\underset{\sim}{u}$, and is denoted by $-\underset{\sim}{u}$.

Clearly, the negative has the property that $(-\underset{\sim}{u})+\underset{\sim}{u}=\underset{\sim}{u}+(-\underset{\sim}{u})=\underset{\sim}{0}$. If we take the negative of $\underset{\sim}{0}$ to be $\underset{\sim}{0}$ itself, then we can say that all vectors have this property. Thus, for vector addition, the negative vectors play the roles of reverse elements.

On comparison with Definition A1.3.1, we can say that the set of plane vectors is a group w.r.t. vector addition.

You also learned another operation with vectors. Given any nonzero vector $\underset{\sim}{u}$ and any nonzero real number α, then the <u>scalar multiple</u> of $\underset{\sim}{u}$ by α, denoted by $\alpha\underset{\sim}{u}$, is defined to be the vector which is parallel to $\underset{\sim}{u}$, points in the same or opposite direction as $\underset{\sim}{u}$ according as $\alpha > 0$ or $\alpha < 0$, and has length $|\alpha||\underset{\sim}{u}|$, where $|\underset{\sim}{u}|$ is the length of $\underset{\sim}{u}$. If either $\alpha = 0$ or $\underset{\sim}{u} = \underset{\sim}{0}$, then $\alpha\underset{\sim}{u} := \underset{\sim}{0}$.

This operation also has important properties. First, we note the obvious notational delights that for any vector $\underset{\sim}{u}$, $1\underset{\sim}{u} = \underset{\sim}{u}$ and $(-1)\underset{\sim}{u} = -\underset{\sim}{u}$. It is important to realize that these are theorems — not notational agreements.

With a little more effort, we can see, as surely you have at some point in your studies, that for any $\alpha, \beta \in \mathbb{R}$ and for any vectors $\underset{\sim}{u}$ and $\underset{\sim}{v}$, scalar multiplication is

<u>associative</u>, $(\alpha\beta)\underset{\sim}{u} = \alpha(\beta\underset{\sim}{u})$,

<u>distributive w.r.t. scalar addition</u>, $(\alpha+\beta)\underset{\sim}{u} = \alpha\underset{\sim}{u} + \beta\underset{\sim}{u}$,

and <u>distributive w.r.t. vector addition</u>, $\alpha(\underset{\sim}{u}+\underset{\sim}{v}) = \alpha\underset{\sim}{u} + \alpha\underset{\sim}{v}$.

In our study of continuum mechanics, we shall encounter many different sets and operations that have essentially the same structure that we have observed above for vectors in the Euclidean plane. Thus, it will be efficient to make the following generalization.

Definition A2.1.1. A *real linear space*[1] is a set, say $\mathcal{V}$, equipped with two special operations. The first operation, called *addition*, is a function on $\mathcal{V}\times\mathcal{V}$ to $\mathcal{V}$, denoted here by

$$(\underset{\sim}{u}, \underset{\sim}{v}) \longmapsto \underset{\sim}{u} + \underset{\sim}{v}$$

with the function value $\underset{\sim}{u}+\underset{\sim}{v}$ called the *sum* of $\underset{\sim}{u}$ and $\underset{\sim}{v}$, which satisfies the following axioms:

(A1) *Associativity*. $\forall\, \underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{w} \in \mathcal{V}$

$$(\underset{\sim}{u}+\underset{\sim}{v}) + \underset{\sim}{w} = \underset{\sim}{u} + (\underset{\sim}{v}+\underset{\sim}{w}) \;;$$

(A2) Existence of a *zero element*. $\exists\ \underset{\sim}{0} \in \mathcal{V} \ni$

$$\underset{\sim}{0} + \underset{\sim}{u} = \underset{\sim}{u} + \underset{\sim}{0} = \underset{\sim}{u} \quad \forall\ \underset{\sim}{u} \in \mathcal{V};$$

(A3) Existence of *negative elements*. For each $\underset{\sim}{u} \in \mathcal{V}$ $\exists$ $-\underset{\sim}{u} \in \mathcal{V} \ni$

$$-\underset{\sim}{u} + \underset{\sim}{u} = \underset{\sim}{u} + (-\underset{\sim}{u}) = \underset{\sim}{0} \;;$$

(A4) *Commutativity*. $\forall\ \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{V}$

$$\underset{\sim}{u} + \underset{\sim}{v} = \underset{\sim}{v} + \underset{\sim}{u}.$$

The second operation, called *scalar multiplication*, is a function

---

[1] Often, *real vector space*, and then the elements of $\mathcal{V}$ are called vectors.

on $\mathbb{R} \times \mathcal{V}$ to $\mathcal{V}$, denoted here by

$$(\alpha, \underset{\sim}{u}) \longmapsto \alpha \underset{\sim}{u}$$

with the function value $\alpha \underset{\sim}{u}$ called the scalar multiple of $\underset{\sim}{u}$ by α, which satisfies the following axioms:

(SM 1) Associativity. $\forall \alpha, \beta \in \mathbb{R}$ and $\forall \underset{\sim}{u} \in \mathcal{V}$

$$(\alpha\beta)\underset{\sim}{u} = \alpha(\beta\underset{\sim}{u})\ ;$$

(SM 2) Distributivity w.r.t. scalar addition. $\forall \alpha, \beta \in \mathbb{R}$ and $\forall \underset{\sim}{u} \in \mathcal{V}$

$$(\alpha+\beta)\underset{\sim}{u} = \alpha\underset{\sim}{u} + \beta\underset{\sim}{u}\ ;$$

(SM 3) Distributivity w.r.t. addition in $\mathcal{V}$. $\forall \alpha \in \mathbb{R}$ and $\forall \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{V}$

$$\alpha(\underset{\sim}{u} + \underset{\sim}{v}) = \alpha\underset{\sim}{u} + \alpha\underset{\sim}{v}\ ;$$

(SM 4) $\forall \underset{\sim}{u} \in \mathcal{V}$

$$1\underset{\sim}{u} = \underset{\sim}{u}.$$

More generally, $\mathbb{R}$ in the above definition could be any "scalar field" $\mathbb{F}$. We would then speak of a linear space over the field $\mathbb{F}$. When $\mathbb{F}$ is the complex number field $\mathbb{C}$, then we have a complex linear space. We will use only real linear spaces, and the adjective real will usually be omitted. Generally, the elements of linear spaces will be denoted by lower case bold face letters, and light face letters will be used for elements of $\mathbb{R}$; sometimes, entrenched custom or the likelihood of confusion will dictate otherwise.

Of course, it is no accident that a linear space is defined in such a way that it is a commutative group w.r.t. addition. Speaking of addition, note that the + on the l.h.s.[1] of (SM 2) refers to ordinary real number addition, while on the r.h.s. + refers to the addition operation on the linear space. If this causes confusion, return to GO and ask for a \$200 tuition refund.

We have already seen that the set of vectors in the Euclidean plane is a linear space. Another important and familiar example is given in the following theorem.

Theorem A2.1.1. *Let n be a strictly positive integer. Consider the set of all lists of length n whose elements are real numbers; i.e.,*

$$\mathbb{R}^n := \{(u_1, u_2, \cdots, u_n) : u_i \in \mathbb{R}\ ,\ i = 1, 2, \cdots, n\}$$

Write $\mathbf{u} = (u_1, u_2, \cdots, u_n)$, $\mathbf{v} = (v_1, v_2, \cdots, v_n)$, *etc., and*

---

[1] l.h.s. stands for "left-hand side"; similarly r.h.s. means "right-hand side".

define addition, scalar multiplication, and equality by

$$\underset{\sim}{u} + \underset{\sim}{v} = (u_1 + v_1, u_2 + v_2, \cdots, u_n + v_n),$$

$$\alpha \underset{\sim}{u} = (\alpha u_1, \alpha u_2, \cdots, \alpha u_n),$$

$$\underset{\sim}{u} = \underset{\sim}{v} \Leftrightarrow u_i = v_i, \quad i = 1, 2, \cdots, n.)$$ [1]

This system is a real linear space. It is called *n-dimensional number space*[2].

Proof. First, we examine the addition operation. Obviously, the domain is $\mathbb{R}^n \times \mathbb{R}^n$, the codomain is $\mathbb{R}^n$, and the evaluation rule is unambiguous. Thus, by Theorem A1.2.1, the operation of addition is, indeed, a function on $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^n$.

---

[1] Note that this is consistent with Definition A1.1.8.

[2] Often, *Cartesian n-space*, *n-dimensional coordinate space*, *n-dimensional arithmetic space*.

Next we check the group axioms for addition.

(A1):

$$(\underset{\sim}{u}+\underset{\sim}{v})+\underset{\sim}{w} = ((u_1+v_1)+w_1, \cdots, (u_n+v_n)+w_n) \quad (\text{definition of } + \text{ in } \mathbb{R}^n)$$

$$= (u_1+(v_1+w_1), \cdots, u_n+(v_n+w_n)) \quad (\text{associativity of } + \text{ in } \mathbb{R} \text{ and definition of } = \text{ in } \mathbb{R}^n)$$

$$= \underset{\sim}{u}+(\underset{\sim}{v}+\underset{\sim}{w}) \quad (\text{definition of } + \text{ in } \mathbb{R}^n).$$

(A2): Define $\underset{\sim}{0} \in \mathbb{R}^n$ to be $\underset{\sim}{0} = (0, 0, \cdots, 0)$. Then $\forall\, \underset{\sim}{u} \in \mathbb{R}^n$

$$\underset{\sim}{0}+\underset{\sim}{u} = (0+u_1, \cdots, 0+u_n) \quad (\text{definition of } + \text{ in } \mathbb{R}^n)$$

$$= (u_1, \cdots, u_n) \quad (\text{property of } 0 \in \mathbb{R} \text{ and definition of } = \text{ in } \mathbb{R}^n)$$

$$= \underset{\sim}{u} \quad (\text{notation}).$$

Similarly, $\underset{\sim}{u}+\underset{\sim}{0} = \underset{\sim}{u}$.

(A3): Given $\underset{\sim}{u} = (u_1, u_2, \cdots, u_n) \in \mathbb{R}^n$, define $-\underset{\sim}{u} \in \mathbb{R}^n$ to be $-\underset{\sim}{u} = (-u_1, -u_2, \cdots, -u_n)$. Then

$$-\underset{\sim}{u} + \underset{\sim}{u} = (-u_1 + u_1, \cdots, -u_n + u_n) \quad (\text{definition of } + \text{ in } \mathbb{R}^n)$$

$$= (0, \cdots, 0) \quad (\text{property of negatives in } \mathbb{R} \text{ and definition of } = \text{ in } \mathbb{R}^n)$$

$$= \underset{\sim}{0} \quad (\text{definition of } \underset{\sim}{0} \in \mathbb{R}^n).$$

Similarly, $\underset{\sim}{u} + (-\underset{\sim}{u}) = \underset{\sim}{0}$.

(A4):

$$\underset{\sim}{u} + \underset{\sim}{v} = (u_1 + v_1, \cdots, u_n + v_n) \quad (\text{definition of } + \text{ in } \mathbb{R}^n)$$

$$= (v_1 + u_1, \cdots, v_n + u_n) \quad (\text{commutivity of } + \text{ in } \mathbb{R} \text{ and definition of } = \text{ in } \mathbb{R}^n)$$

$$= \underset{\sim}{v} + \underset{\sim}{u} \quad (\text{definition of } + \text{ in } \mathbb{R}^n)$$

Scalar multiplication can be checked out in a similar fashion, and this constitutes *Exercise* A2.1.1. □

Exercise A2.1.2. Consider the set of all real numbers $\mathbb{R}$. Denote them by x, y, etc. Define addition and scalar multiplication by ordinary real number addition and multiplication. Of course, the scalars are also elements of $\mathbb{R}$; to minimize confusion, denote them by α, β, etc. Show that this system is a linear space. Is it different from $\mathbb{R}^1$ of Theorem A2.1.1?

Exercise A2.1.3. Consider the set of all continuous, real-valued functions on the closed interval $[a,b] := \{x \in \mathbb{R} : a \le x \le b\}$; i.e.,

$$C^0([a,b]) := \{f : f \text{ is a continuous function on } [a,b] \text{ to } \mathbb{R}\}$$

Define addition, scalar multiplication, and equality by

$$(f+g)(x) = f(x) + g(x) \quad , \quad x \in [a,b] \quad ,$$

$$(\alpha f)(x) = \alpha[f(x)] \quad , \quad x \in [a,b] \quad ,$$

and

$$f = g \iff f(x) = g(x) \;\; \forall x \in [a,b] \, . \quad)^1$$

Show that $C^0([a,b])$ is a linear space. Feel free to draw on theorems from calculus.

Since a linear space is a group w.r.t. addition (with $\underset{\sim}{0}$ and $-\underset{\sim}{u}$ playing the roles of the neutral and reverse elements, respectively), all of the results of §A1.3 apply. From Theorems A1.3.4 and 5, we have

Theorem A2.1.2. <u>A given linear space has only one zero element. A given element of a linear space has only one negative.</u>

In the context of a linear space, the cancellation property (Theorem A1.3.3) becomes

---

[1] Cf. Footnote 1 on p. A1.2.13.

Theorem A2.1.3. For $\underset{\sim}{u}$, $\underset{\sim}{v}$, and $\underset{\sim}{w}$ elements of a linear space

$$\underset{\sim}{u} + \underset{\sim}{v} = \underset{\sim}{u} + \underset{\sim}{w} \Rightarrow \underset{\sim}{v} = \underset{\sim}{w}.$$

Since addition in a linear space is commutative, we do not bother to formally state the companion result that

$$\underset{\sim}{v} + \underset{\sim}{u} = \underset{\sim}{w} + \underset{\sim}{u} \Rightarrow \underset{\sim}{v} = \underset{\sim}{w} ;$$

and, of course, this can be written in still other variants.

Theorems A1.3.6 and 7 give us the following interesting property of negatives

Theorem A2.1.4. Let $\underset{\sim}{u}$ be any element of a linear space, and let $\underset{\sim}{0}$ be the zero element. Then

$$-\underset{\sim}{0} = \underset{\sim}{0} \quad \text{and} \quad -(-\underset{\sim}{u}) = \underset{\sim}{u}.$$

For stating the remaining consequences of the group structure, it will be convenient to introduce

Definition A2.1.2. The *difference* $\underset{\sim}{u} - \underset{\sim}{v}$ of two elements $\underset{\sim}{u}$ and $\underset{\sim}{v}$ of a linear space is the sum of $\underset{\sim}{u}$ and $-\underset{\sim}{v}$; i.e.,

$$\underset{\sim}{u} - \underset{\sim}{v} = \underset{\sim}{u} + (-\underset{\sim}{v}).$$

The operation $(\underset{\sim}{u}, \underset{\sim}{v}) \mapsto \underset{\sim}{u} - \underset{\sim}{v}$ is called *subtraction*.

Of course, by closure the difference of two elements belongs to the underlying linear space. By Axiom (A3), the difference has the following property.

Theorem A2.1.5. Let $\underset{\sim}{u}$ be any element of a linear space. Then

$$\underset{\sim}{u} - \underset{\sim}{u} = \underset{\sim}{0}.$$

Note that the above result is not true just because it looks right in terms of our experience with minus signs.

Theorem A1.3.8 has the following interpretation for linear spaces.

Theorem A2.1.6. Let $\underset{\sim}{u}$ and $\underset{\sim}{v}$ be any two elements of a linear space. Then

$$-(\underset{\sim}{u} + \underset{\sim}{v}) = -\underset{\sim}{u} + (-\underset{\sim}{v}) = -\underset{\sim}{u} - \underset{\sim}{v}.$$

Theorem A1.3.2 gives us

Theorem A2.1.7. Let $\mathcal{V}$ be a linear space. Then for any $\underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{V}$, $\exists$ a unique $\underset{\sim}{x} \in \mathcal{V} \ni$

$$\underset{\sim}{u} + \underset{\sim}{x} = \underset{\sim}{v}.$$

In fact,

$$\underset{\sim}{x} = \underset{\sim}{v} + (-\underset{\sim}{u}) = \underset{\sim}{v} - \underset{\sim}{u}.$$

Finally, we note that in a linear space when equals are added to or subtracted from equals, the results are equal.

Theorem A2.1.8. Let $\underset{\sim}{s}$, $\underset{\sim}{t}$, $\underset{\sim}{u}$, and $\underset{\sim}{v}$ be elements of a linear space. Then

$$\underset{\sim}{s} = \underset{\sim}{t} \text{ and } \underset{\sim}{u} = \underset{\sim}{v} \Rightarrow \underset{\sim}{s} + \underset{\sim}{u} = \underset{\sim}{t} + \underset{\sim}{v} \text{ and } \underset{\sim}{s} - \underset{\sim}{u} = \underset{\sim}{t} - \underset{\sim}{v}$$

Proof. The result for sums is simply the linear space counterpart of Theorem A1.3.1.

To get the subtraction version, we note that $\underset{\sim}{u} = \underset{\sim}{v}$ means that $\underset{\sim}{u}$ and $\underset{\sim}{v}$ are really the same element of the linear space; $\therefore$ $\underset{\sim}{u} = \underset{\sim}{v} \Rightarrow -\underset{\sim}{u} = -\underset{\sim}{v}$. Then by the result for sums

$$\underset{\sim}{s} = \underset{\sim}{t} \text{ and } \underset{\sim}{u} = \underset{\sim}{v} \Rightarrow \underset{\sim}{s} + (-\underset{\sim}{u}) = \underset{\sim}{t} + (-\underset{\sim}{v})$$

$$\Rightarrow \underset{\sim}{s} - \underset{\sim}{u} = \underset{\sim}{t} - \underset{\sim}{v} \quad \text{(Definition A2.1.2)}. \square$$

The scalar multiplication analogue of the above result is

Theorem A2.1.9. Let $\alpha, \beta \in \mathbb{R}$, and let $\underset{\sim}{u}$ and $\underset{\sim}{v}$ be elements of a linear space. Then

$$\alpha = \beta \text{ and } \underset{\sim}{u} = \underset{\sim}{v} \Rightarrow \alpha \underset{\sim}{u} = \beta \underset{\sim}{v}.$$

Proof. Exercise A2.1.4. $\square$

For the (remaining) properties of scalar multiplication, we have nothing to fall back on except for our experience with particular examples such as vectors in the Euclidean plane. The more common properties are gathered together in

Theorem A2.1.10. Let $\underset{\sim}{u}$ and $\underset{\sim}{v}$ be any two elements of a real linear space, and let α and β be any two real numbers. Then

(i) $0\underset{\sim}{u} = \underset{\sim}{0}$;

(ii) $\alpha \underset{\sim}{0} = \underset{\sim}{0}$;

(iii) $\alpha \underset{\sim}{u} = \underset{\sim}{0} \Rightarrow$ either $\alpha = 0$ or $\underset{\sim}{u} = \underset{\sim}{0}$;

(iv) $-(\alpha \underset{\sim}{u}) = \alpha(-\underset{\sim}{u})$;

(v) $-(\alpha \underset{\sim}{u}) = (-\alpha)\underset{\sim}{u}$;

(vi) $-\underset{\sim}{u} = (-1)\underset{\sim}{u}$;

(vii) $(\alpha - \beta)\underset{\sim}{u} = \alpha \underset{\sim}{u} - \beta \underset{\sim}{u}$;

(viii) $\alpha(\underset{\sim}{u} - \underset{\sim}{v}) = \alpha \underset{\sim}{u} - \alpha \underset{\sim}{v}$.

Proof. (i): $\underset{\sim}{u} = 1\underset{\sim}{u}$ (Axiom (SM4))

$= (1+0)\underset{\sim}{u}$ (zero in $\mathbb{R}$)

$$= 1\underset{\sim}{u} + 0\underset{\sim}{u} \qquad (SM2)$$

$$= \underset{\sim}{u} + 0\underset{\sim}{u} \qquad (SM4).$$

It is tempting to conclude at this point on the basis of (A2) and the uniqueness of $\underset{\sim}{0}$ that $0\underset{\sim}{u} = \underset{\sim}{0}$. However, the $\underset{\sim}{0}$ in (A2) must work $\forall$ elements; i.e.,

(Do not have $\underset{\sim}{v} = \underset{\sim}{v} + 0\underset{\sim}{u}$ from (A2) to conclude that $0\underset{\sim}{u} = \underset{\sim}{0}$)

$$\underset{\sim}{u} = \underset{\sim}{u} + \underset{\sim}{0}, \quad \underset{\sim}{v} = \underset{\sim}{v} + \underset{\sim}{0}, \quad \underset{\sim}{w} = \underset{\sim}{w} + \underset{\sim}{0}, \text{ etc.}$$

But as far as we know now, $0\underset{\sim}{u}$ is only "a zero for $\underset{\sim}{u}$". To continue, we subtract $\underset{\sim}{u}$ from both sides of the above result to get

$$\underset{\sim}{u} - \underset{\sim}{u} = (\underset{\sim}{u} + 0\underset{\sim}{u}) - \underset{\sim}{u} \qquad (\text{Theorem A2.1.8})$$

$$\Rightarrow$$

If $\underset{\sim}{v} = \underset{\sim}{x} + \underset{\sim}{v}$ for some $\underset{\sim}{v}$ then $\underset{\sim}{x} = \underset{\sim}{0}$. Proof

$$\underset{\sim}{0} = (\underset{\sim}{u} + 0\underset{\sim}{u}) - \underset{\sim}{u} \qquad (\text{Theorem A2.1.5})$$

$$= -\underset{\sim}{u} + (\underset{\sim}{u} + 0\underset{\sim}{u}) \qquad (\text{Definition A2.1.2 and commutivity})$$

$$= (-\underset{\sim}{u} + \underset{\sim}{u}) + 0\underset{\sim}{u} \qquad (\text{associativity})$$

$$= \underset{\sim}{0} + 0\underset{\sim}{u} \qquad (\text{negative})$$

$$= 0\underset{\sim}{u} \qquad (\text{zero element}).$$

For reasons which will become apparent, it is convenient to prove the remaining assertions in an order which differs from that in which they were stated.

(v): $\underset{\sim}{0} = \alpha\underset{\sim}{u} + [-(\alpha\underset{\sim}{u})]$ (negative)

$\Rightarrow (-\alpha)\underset{\sim}{u} + \underset{\sim}{0} = (-\alpha)\underset{\sim}{u} + \{\alpha\underset{\sim}{u} + [-(\alpha\underset{\sim}{u})]\}$ (Theorem A2.1.8)

$\Rightarrow (-\alpha)\underset{\sim}{u} = [(-\alpha)\underset{\sim}{u} + \alpha\underset{\sim}{u}] + [-(\alpha\underset{\sim}{u})]$ (zero and associativity)

$= [(-\alpha) + \alpha]\underset{\sim}{u} + [-(\alpha\underset{\sim}{u})]$ (distributivity)

$= 0\underset{\sim}{u} + [-(\alpha\underset{\sim}{u})]$ (negative in $\mathbb{R}$)

$= \underset{\sim}{0} + [-(\alpha\underset{\sim}{u})]$ (i)

$= -(\alpha\underset{\sim}{u})$ (zero).

(vi): Take $\alpha = 1$ in (v) to get

$(-1)\underset{\sim}{u} = -(1\underset{\sim}{u})$

$= -\underset{\sim}{u}$ (SM4).

(vii): $(\alpha - \beta)\underset{\sim}{u} = [\alpha + (-\beta)]\underset{\sim}{u}$ (subtraction in $\mathbb{R}$)

$= \alpha\underset{\sim}{u} + (-\beta)\underset{\sim}{u}$ (distributivity)

$= \alpha\underset{\sim}{u} + [-(\beta\underset{\sim}{u})]$ (v)

$$= \alpha \underset{\sim}{u} - \beta \underset{\sim}{u} \quad \text{(Definition A2.1.2)}.$$

(iv): Take $\alpha = 0$ in (vii) to get

$$(0-\beta)\underset{\sim}{u} = 0\underset{\sim}{u} - \beta\underset{\sim}{u}$$

$$\Rightarrow \quad (-\beta)\underset{\sim}{u} = 0\underset{\sim}{u} - \beta\underset{\sim}{u} \quad \text{(zero in } \mathbb{R}\text{)}$$

$$= 0\underset{\sim}{u} + [-(\beta\underset{\sim}{u})] \quad \text{(Definition A2.1.2)}$$

$$= \underset{\sim}{0} + [-(\beta\underset{\sim}{u})] \quad \text{(i)}$$

$$= -(\beta\underset{\sim}{u}) \quad \text{(zero element)}.$$

(viii): $\alpha(\underset{\sim}{u} - \underset{\sim}{v}) = \alpha[\underset{\sim}{u} + (-\underset{\sim}{v})]$ (Definition A2.1.2)

$$= \alpha\underset{\sim}{u} + \alpha(-\underset{\sim}{v}) \quad \text{(distributivity)}$$

$$= \alpha\underset{\sim}{u} + \alpha[(-1)\underset{\sim}{v}] \quad \text{(vi)}$$

$$= \alpha\underset{\sim}{u} + [(-1)\alpha]\underset{\sim}{v} \quad \text{(associativity)}$$

$$= \alpha\underset{\sim}{u} + (-\alpha)\underset{\sim}{v} \quad \text{(property of } \mathbb{R}\text{)}$$

$$= \alpha\underset{\sim}{u} + [-(\alpha\underset{\sim}{v})] \quad \text{(iv)}$$

$$= \alpha\underset{\sim}{u} - \alpha\underset{\sim}{v} \quad \text{(Definition A2.1.2)}.$$

(ii): Take $\underset{\sim}{u} = \underset{\sim}{v}$ in (viii) to get

$$\alpha(\underset{\sim}{u} - \underset{\sim}{u}) = \alpha\underset{\sim}{u} - \alpha\underset{\sim}{u}$$

$$\Rightarrow \quad \alpha(\underset{\sim}{0}) = \underset{\sim}{0} \qquad \text{(Theorem A2.1.5).}$$

(iii): Suppose that $\alpha\underset{\sim}{u} = \underset{\sim}{0}$. If $\alpha \neq 0$, then we can use Theorem A2.1.9 to multiply both sides by $1/\alpha$ to get

$$1/\alpha\, \underset{\sim}{0} = 1/\alpha\,(\alpha\underset{\sim}{u})$$

$$\begin{aligned}
\Rightarrow \quad \underset{\sim}{0} &= 1/\alpha\,(\alpha\underset{\sim}{u}) && \text{(ii)} \\
&= (1/\alpha\,\alpha)\underset{\sim}{u} && \text{(associativity)} \\
&= (1)\underset{\sim}{u} && \text{(reciprocal in } \mathbb{R}) \\
&= \underset{\sim}{u} && \text{(SM4).}
\end{aligned}$$

Thus, if $\alpha \neq 0$, then $\alpha\underset{\sim}{u} = \underset{\sim}{0} \Rightarrow \underset{\sim}{u} = \underset{\sim}{0}$.

Now assume that $\alpha\underset{\sim}{u} = \underset{\sim}{0}$ and $\underset{\sim}{u} \neq \underset{\sim}{0}$. We show that in this case $\alpha = 0$ by <u>reductio</u> <u>ad</u> <u>absurdum</u>; i.e., by argument by contradiction. Suppose $\alpha \neq 0$. Then we can use the argument above to conclude that $\underset{\sim}{u} = \underset{\sim}{0}$ — contradicting the hypothesis that $\underset{\sim}{u} \neq \underset{\sim}{0}$. $\therefore$ $\alpha \neq 0$ is false, which means that $\alpha = 0$.

Of course, if either $\alpha = 0$ or $\underset{\sim}{u} = \underset{\sim}{0}$ (or both) to begin with (and, in view of (i) and (ii), these are consistent with $\alpha\underset{\sim}{u} = \underset{\sim}{0}$), then there is nothing to prove. □

It is useful to realize that if we have a candidate for a linear space in the sense of having the elements and the operations of addition and scalar multiplication, then (i) and (vi) of Theorem A2.1.10 tell us exactly what the zero and the negatives must be; i.e., for any $\underset{\sim}{u}$

$$\boxed{\underset{\sim}{0} = 0\underset{\sim}{u}} \quad \text{and} \quad -\underset{\sim}{u} = (-1)\underset{\sim}{u}\,.$$

In view of the uniqueness theorem, Theorem A2.1.2, there are no other possibilities. Often the candidate fails because the zero element or the negatives are not contained in the given set.

We generally will use the linear space axioms and Theorems A2.1.3 – 10 without further comment. I.e., we will not meticulously write out every step such as we did in the proof of Theorem A2.1.10. However, one must, of course, always be able to do this.

Definition A2.1.3. A nonempty subset $\mathcal{S}$ of a linear space $\mathcal{V}$ is said to be a *linear subspace*[1] of $\mathcal{V}$ if it is closed w.r.t. addition and scalar multiplication; i.e.,

[1] Often, simply *subspace* or *linear manifold*.

$$(S1) \quad \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{S} \Rightarrow (\underset{\sim}{u} + \underset{\sim}{v}) \in \mathcal{S}$$

and

$$(S2) \quad \underset{\sim}{u} \in \mathcal{S},\ \alpha \in \mathbb{R} \Rightarrow \alpha \underset{\sim}{u} \in \mathcal{S}.$$

Exercise A2.1.5. Refer to Definition A2.1.3. Prove that (S1) and (S2) together are equivalent to

$$\underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{S} \text{ and } \alpha, \beta \in \mathbb{R} \Rightarrow (\alpha \underset{\sim}{u} + \beta \underset{\sim}{v}) \in \mathcal{S}$$

Theorem A2.1.11. A linear subspace of a linear space is itself a linear space.[1]

Proof. *Exercise* A2.1.6. □

Definition A2.1.4. Let $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\} \subset \mathcal{V}$, where $\mathcal{V}$ is a linear space. Then an element of $\mathcal{V}$ of the form

$$\alpha_1 \underset{\sim}{u}_1 + \alpha_2 \underset{\sim}{u}_2 + \cdots + \alpha_k \underset{\sim}{u}_k ,$$

where $\alpha_1, \alpha_2, \cdots, \alpha_k \in \mathbb{R}$, is said to be a *linear combination* of the subset $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$.

The set of all possible linear combinations of a given subset is of special interest.

---

[1] Of course, the operations of addition and scalar multiplication for the subspace are borrowed from the "parent" space.

Definition A2.1.5. Let $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_k\} \subset V$, where V is a linear space. Then the set

$$Lsp\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_k\} = \{\underset{\sim}{u} : \underset{\sim}{u} = \alpha_1 \underset{\sim}{u}_1 + \alpha_2 \underset{\sim}{u}_2 + \cdots + \alpha_k \underset{\sim}{u}_k,\ \alpha_i \in \mathbb{R}\}$$

is called the *linear span*[1] of $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_k\}$.

Theorem A2.1.12. $Lsp\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_k\}$ as defined above is a linear subspace of V.

Proof. *Exercise* A2.1.7. □

---

[1] More often, simply *span*, with the notation $Sp\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_k\}$.

Supplementary Reading

BISHOP and GOLDBERG, Tensor Analysis on Manifolds

BOWEN and WANG, Introduction to Vectors and Tensors, Vol. 1, Linear and Multilinear Algebra

GEL'FAND, Lectures on Linear Algebra

GREUB, Linear Algebra

HALMOS, Finite-Dimensional Vector Spaces

LOOMIS and STERNBERG, Advanced Calculus

MARTIN and MIZEL, Introduction to Linear Algebra

MICHEL and HERGET, Mathematical Foundations in Engineering and Science

MOSTOW, SAMPSON, and MEYER, Fudamental Structures of Algebra

NAYLOR and SELL, Linear Operator Theory in Engineering and Science
NICKERSON, SPENCER, and STEENROD, Advanced Calculus
NOLL, Finite-Dimensional Spaces

NOMIZU, Fundamentals of Linear Algebra
ODEN, Applied Functional Analysis
[illegible]

A2.2. Finite-Dimensional Linear Spaces

There are many different ways to organize the material presented in this section. Consequently, some of our theorems will be definitions in alternative treatments and vice versa. We shall adopt what seems to us to be the most natural approach; it is not necessarily the most efficient.

Definition A2.2.1. A subset $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$ of a linear space is said to be *linearly dependent* if at least one of the elements of the subset is a linear combination[1] of the remaining elements of the subset. Otherwise, the subset is *linearly independent*.

The next theorem provides a standard tool for proving that a subset is linearly dependent.

Theorem A2.2.1. A subset $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$ of a linear space is linearly dependent iff $\exists\ \alpha_1, \alpha_2, \cdots, \alpha_k \in \mathbb{R}$ not all zero $\ni$

$$\alpha_1 \underset{\sim}{u}_1 + \alpha_2 \underset{\sim}{u}_2 + \cdots + \alpha_k \underset{\sim}{u}_k = \underset{\sim}{0}. \quad)^2$$

Proof. Suppose that $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$ is linearly dependent.

---

[1] Of course, the operations of addition and scalar multiplication from the underlying linear space are used to form the linear combination; cf. Definition A2.1.4.

[2] Of course, $\underset{\sim}{0}$ is the zero element of the underlying linear space.

Then one of the $\underset{\sim}{u}$'s, say $\underset{\sim}{u}_1$, can be written as a linear combination of the others; i.e., $\exists\ \beta_2, \beta_3, \cdots, \beta_k \in \mathbb{R} \ni$

$$\underset{\sim}{u}_1 = \beta_2 \underset{\sim}{u}_2 + \beta_3 \underset{\sim}{u}_3 + \cdots + \beta_k \underset{\sim}{u}_k .$$

$$\Rightarrow \quad \underset{\sim}{u}_1 - \beta_2 \underset{\sim}{u}_2 - \beta_3 \underset{\sim}{u}_3 - \cdots - \beta_k \underset{\sim}{u}_k = \underset{\sim}{0} .$$

Thus, $\exists\ \alpha$'s not all zero $(\alpha_1 = 1, \alpha_2 = -\beta_2, \alpha_3 = -\beta_3, \cdots, \alpha_k = -\beta_k)$ $\ni$

$$\alpha_1 \underset{\sim}{u}_1 + \alpha_2 \underset{\sim}{u}_2 + \cdots + \alpha_k \underset{\sim}{u}_k = \underset{\sim}{0} .$$

Conversely, suppose that $\exists\ \alpha$'s not all zero $\ni$ this last equation holds. For definiteness, take $\alpha_1 \neq 0$. Then the above equation $\Rightarrow$

$$\underset{\sim}{u}_1 = -\frac{\alpha_2}{\alpha_1} \underset{\sim}{u}_2 - \frac{\alpha_3}{\alpha_1} \underset{\sim}{u}_3 - \cdots - \frac{\alpha_k}{\alpha_1} \underset{\sim}{u}_k .$$

Here, one of the $\underset{\sim}{u}$'s can be expressed as a linear combination of the others; and $\therefore$ the set $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$ is linearly dependent. $\square$

The contrapositive of Theorem A2.2.1 provides a necessary and sufficient condition for linear independence. Even though it is less transparent than the prior content of this section, it is so useful that it is often taken as the starting point for this material.

<u>Theorem A2.2.2.</u> <u>A subset $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$ of a linear space is is linearly independent iff the equation</u>

$$\alpha_1 \underset{\sim}{u}_1 + \alpha_2 \underset{\sim}{u}_2 + \cdots + \alpha_k \underset{\sim}{u}_k = \underset{\sim}{0} \qquad (\alpha_i \in \mathbb{R})$$

$$\Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0.$$

Proof. Suppose that

$$\alpha_1 \underset{\sim}{u}_1 + \alpha_2 \underset{\sim}{u}_2 + \cdots + \alpha_k \underset{\sim}{u}_k = \underset{\sim}{0} \Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0.$$

Then it is easy to see that the subset $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$ is linearly independent. For suppose not; then the subset is linearly dependent, and by Theorem A2.2.1 $\exists$ β's not all zero $\ni$ $\beta_1 \underset{\sim}{u}_1 + \beta_2 \underset{\sim}{u}_2 + \cdots + \beta_k \underset{\sim}{u}_k = \underset{\sim}{0}$ — a contradiction of the starting hypothesis. $\therefore$ $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$ is linearly independent.

Conversely, suppose that the subset $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$ is linearly independent. Then

$$\alpha_1 \underset{\sim}{u}_1 + \alpha_2 \underset{\sim}{u}_2 + \cdots + \alpha_k \underset{\sim}{u}_k = \underset{\sim}{0} \Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0.$$

For if not, $\exists$ α's not all zero $\ni$ $\alpha_1 \underset{\sim}{u}_1 + \alpha_2 \underset{\sim}{u}_2 + \cdots + \alpha_k \underset{\sim}{u}_k = \underset{\sim}{0}$; and by Theorem A2.2.1, $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$ is linearly dependent — a contradiction. $\square$

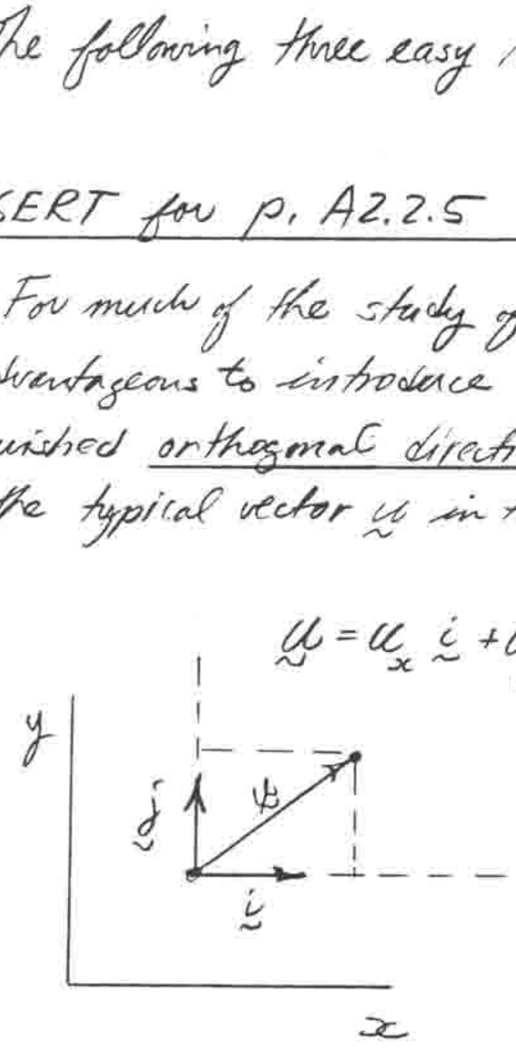
Exercise A2.2.1. Let $\underset{\sim}{u} \in V$, where V is a linear space. Show that if $\underset{\sim}{u} \neq \underset{\sim}{0}$, then the singleton $\{\underset{\sim}{u}\}$ is linearly independent.

The following three easy results are often useful.

INSERT for p. A2.2.5

For much of the study of vectors in the Euclidean plane, it is advantageous to introduce unit vectors, say $\underset{\sim}{i}$ and $\underset{\sim}{j}$, along distinguished orthogonal directions, say x and y, so as to express the typical vector $\underset{\sim}{u}$ in the form

$$\underset{\sim}{u} = u_x \underset{\sim}{i} + u_y \underset{\sim}{j}.$$



This, of course, is the first step in the development of the subject of analytic geometry. Our first step in the generalization of this line of thought to abstract linear spaces is

Theorem A2.2.3. Any finite[1] subset (of a linear space) which contains the zero element is linearly dependent.

Proof. Exercise A2.2.2 □

Theorem A2.2.4. Any subset of a linearly independent subset is linearly independent.

Proof. Exercise A2.2.3. □

Theorem A2.2.5. Any finite subset (of a linear space) which contains a linearly dependent subset is linearly dependent.

Proof. Exercise A2.2.4. □

INSERT material on p. A2.2.4

Definition A2.2.2. Let $\mathcal{V}$ be a linear space. A subset $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\} \subset \mathcal{V}$ is said to be a basis for $\mathcal{V}$ if

(B1) $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ is linearly independent

and

(B2) $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ spans[2] $\mathcal{V}$ in the sense that

$$\mathcal{V} \subset \mathcal{L}sp\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}.$$

---

[1] Our definition of linear dependence is restricted to finite sets. See the book by NAYLOR and SELL listed in the Supplementary reading for an extension of the notion to infinite sets.

[2] Cf. Definition A2.1.5.

Necessarily, $Lsp\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\} \subset \mathcal{V}$, so we could have stated (B2) as $\mathcal{V} = Lsp\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$.

Theorem A2.2.6. <u>All bases for a given linear space[1] contain the same number of elements.</u>

Proof. Suppose that both $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ and $\{\underset{\sim}{f}_1, \underset{\sim}{f}_2, \cdots, \underset{\sim}{f}_m\}$ are bases for the linear space $\mathcal{V}$. Since $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ spans $\mathcal{V}$ and since $\{\underset{\sim}{f}_1, \underset{\sim}{f}_2, \cdots, \underset{\sim}{f}_m\} \subset \mathcal{V}$, $\exists$ $\beta_{ij} \in \mathbb{R} \ni$

$$\underset{\sim}{f}_i = \beta_{1i}\underset{\sim}{e}_1 + \beta_{2i}\underset{\sim}{e}_2 + \cdots + \beta_{ni}\underset{\sim}{e}_n, \quad i = 1, 2, \cdots, m.$$

Then for any choice of $\alpha_1, \alpha_2, \cdots, \alpha_m \in \mathbb{R}$, we have

$$\alpha_1\underset{\sim}{f}_1 + \alpha_2\underset{\sim}{f}_2 + \cdots + \alpha_m\underset{\sim}{f}_m = \left(\sum_{j=1}^{m}\beta_{1j}\alpha_j\right)\underset{\sim}{e}_1 + \left(\sum_{j=1}^{m}\beta_{2j}\alpha_j\right)\underset{\sim}{e}_2 + \cdots + \left(\sum_{j=1}^{m}\beta_{nj}\alpha_j\right)\underset{\sim}{e}_n$$

Consider the system of homogeneous linear algebraic equations

$$\sum_{j=1}^{m}\beta_{ij}\alpha_j = 0, \quad i = 1, 2, \cdots, n,$$

and think of the α's as unknowns. Suppose that $m > n$. Then there are more unknowns than equations, and it is always

---

[1] It is possible that a given linear space will have no basis; this, in fact, is the common circumstance.

possible to find a nontrivial solution.[1] More precisely, $\exists$ $\alpha_1, \alpha_2, \cdots, \alpha_m \in \mathbb{R}$ not all zero $\ni$

$$\sum_{j=1}^{m} \beta_{ij}\alpha_j = 0 \quad , \quad i = 1, 2, \cdots, n .$$

With such a choice of the α's

$$\alpha_1 \underset{\sim}{f}_1 + \alpha_2 \underset{\sim}{f}_2 + \cdots + \alpha_m \underset{\sim}{f}_m = \underset{\sim}{0} \; ,$$

which by Theorem A2.2.2 contradicts the linear independence of $\{\underset{\sim}{f}_1, \underset{\sim}{f}_2, \cdots, \underset{\sim}{f}_m\}$. $\therefore$ $m \le n$.

Next suppose that $m < n$ and reverse the roles of $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ and $\{\underset{\sim}{f}_1, \underset{\sim}{f}_2, \cdots, \underset{\sim}{f}_m\}$ in the above argument to conclude that $m \ge n$. The details are left as *Exercise* A2.2.5. Hence, $m = n$. $\square$

In view of Theorem A2.2.6, the following definition is meaningful.

Definition A2.2.3. A linear space V is *n-dimensional*[2] if it contains a basis with n elements; If no such integer n exists, then the linear space is *infinite-dimensional*. Of course, n is called the *dimension* of V, and we write $n =: \dim V$.

---

[1] See almost any text that treats linear algebraic equations. A particularly good one is HOHN's Elementary Matrix Algebra.

[2] Often, simply *finite-dimensional*.

Symbols such as V_n are often used to denote an n-dimensional space.

Exercise A2.2.6. Consider n-dimensional number space $\mathbb{R}^n$. Prove that the lists

$$\begin{aligned} \underset{\sim}{e}_1 &= (1, 0, 0, \cdots, 0) \\ \underset{\sim}{e}_2 &= (0, 1, 0, 0, \cdots, 0) \\ &\cdots \\ \underset{\sim}{e}_n &= (0, 0, \cdots, 0, 1) \end{aligned}$$

constitute a basis for $\mathbb{R}^n$. It then follows from Definition A2.2.3 that $\mathbb{R}^n$ is, indeed, n-dimensional as its name suggests. The above basis is called the standard basis for $\mathbb{R}^n$.

Exercise A2.2.7. In the spirit of Exercise A2.1.2, show that the singleton $\{1\}$ is a linearly independent subset which spans $\mathbb{R}$. Thus, $\mathbb{R}$, when viewed as a linear space, is 1-dimensional.

The following result will not be a surprise.

Theorem A2.2.7. Let V_n be an n-dimensional linear space. Then any subset of V_n which contains more than n elements is linearly dependent.

Proof. Consider $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_m\} \subset V_n$ with $m > n$. Since V_n is n-dimensional, it has a basis $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$. Since

$\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_m\} \subset V_n \subset Lsp\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$, $\exists\ \beta_{ij} \in \mathbb{R} \ni$

$$\underset{\sim}{u}_i = \beta_{1i}\underset{\sim}{e}_1 + \beta_{2i}\underset{\sim}{e}_2 + \cdots + \beta_{ni}\underset{\sim}{e}_n \ , \quad i = 1, 2, \cdots, m.$$

We leave it as Exercise A2.2.8 for the student to continue the argument as in the proof of Theorem A2.2.6 to the conclusion that $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_m\}$ is linearly dependent. □

The following theorem is helpful in finding a basis when the dimension of the space is already known.

Theorem A2.2.8. Let V_n be an n-dimensional linear space. Then a subset $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\} \subset V_n$ is a basis for V_n iff it is linearly independent.

Proof. Suppose that $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ is a basis for V_n. Then it is linearly independent by Definition A2.2.2.

Conversely, suppose that $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ is linearly independent. In order to prove that it is a basis for V_n, we must show that it spans V_n. Accordingly, let $\underset{\sim}{u}$ be an arbitrary element of V_n and consider the subset $\{\underset{\sim}{u}, \underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$. By Theorem A2.2.7, this subset of $n+1$ elements must be linearly dependent. $\therefore$ by Theorem A2.2.1, $\exists$ $\alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{R}$ not all zero $\ni$

$$\alpha_0\underset{\sim}{u} + \alpha_1\underset{\sim}{e}_1 + \alpha_2\underset{\sim}{e}_2 + \cdots + \alpha_n\underset{\sim}{e}_n = \underset{\sim}{0} \ .$$

Now $\alpha_0 \neq 0$. For if $\alpha_0 = 0$, then the above equation would become

$$\alpha_1 \underset{\sim}{e}_1 + \alpha_2 \underset{\sim}{e}_2 + \cdots + \alpha_n \underset{\sim}{e}_n = \underset{\sim}{0}$$

with not all of $\alpha_1, \alpha_2, \cdots, \alpha_n$ zero ; and by Theorem A2.2.2, this contradicts the linear independence of $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$. Since $\alpha_0 \neq 0$, we can manipulate

$$\alpha_0 \underset{\sim}{u} + \alpha_1 \underset{\sim}{e}_1 + \alpha_2 \underset{\sim}{e}_2 + \cdots + \alpha_n \underset{\sim}{e}_n = \underset{\sim}{0}$$

into

$$\underset{\sim}{u} = -\frac{\alpha_1}{\alpha_0} \underset{\sim}{e}_1 - \frac{\alpha_2}{\alpha_0} \underset{\sim}{e}_2 - \cdots - \frac{\alpha_n}{\alpha_0} \underset{\sim}{e}_n \in Lsp\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}.$$

$\therefore \; \mathcal{V}_n \subset Lsp\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$. $\square$

We shall have occasion to use the next two theorems.

Theorem A2.2.9. Let $\mathcal{V}_n$ be an n-dimensional linear space. Then any linearly independent subset $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_m\} \subset \mathcal{V}_n$ with $m < n$ can be extended to a basis for $\mathcal{V}_n$; i.e., $\exists$ $n-m$ elements $\underset{\sim}{e}_{m+1}, \underset{\sim}{e}_{m+2}, \cdots, \underset{\sim}{e}_n \in \mathcal{V}_n \ni$ $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_m, \underset{\sim}{e}_{m+1}, \underset{\sim}{e}_{m+2}, \cdots, \underset{\sim}{e}_n\}$ is a basis for $\mathcal{V}_n$.

Proof. Choose $\underset{\sim}{e}_{m+1} \in \mathcal{V}_n \ni \underset{\sim}{e}_{m+1} \notin Lsp\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_m\}$. This is always possible. For if not, then the linearly independent subset $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_m\}$ spans $\mathcal{V}_n$; consequently, it is a basis for $\mathcal{V}_n$, and $n = m$ — a contradiction. The subset $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_m, \underset{\sim}{e}_{m+1}\}$ is linearly independent. For if not, it is linearly dependent; then since $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_m\}$ is linearly independent, it is easy to show as in the proof of Theorem A2.2.8 (The details are left as Exercise A2.2.9.) that $\underset{\sim}{e}_{m+1} \in Lsp\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_m\}$ — a contradiction. If $m+1=n$, then by Theorem A2.2.8, we are done.

If $m+1<n$, then we repeat the procedure. Choose $\underset{\sim}{e}_{m+2} \in \mathcal{V}_n \ni \underset{\sim}{e}_{m+2} \notin Lsp\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_m, \underset{\sim}{e}_{m+1}\}$. Then $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_m, \underset{\sim}{e}_{m+1}, \underset{\sim}{e}_{m+2}\}$ is linearly independent. If $m+2=n$, we are done.

Clearly, after a total of precisely $n-m$ such steps, we have an extended basis. (Exercise A2.2.10. Devise a mathematical induction argument to show that the procedure never breaks down.) □

A useful complement to Theorem A2.2.9 is given by

Theorem A2.2.10. Let $\mathcal{U} \neq \{\underset{\sim}{0}\}$ be a linear subspace of an n-dimensional linear space $\mathcal{V}_n$. Then $\mathcal{U}$ is m-dimensional with $m \leq n$, and $\exists$ a basis $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_m, \underset{\sim}{e}_{m+1}, \underset{\sim}{e}_{m+2}, \cdots, \underset{\sim}{e}_n\}$ for $\mathcal{V}_n \ni \{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_m\}$ is a basis for $\mathcal{U}$.

Proof. Since $\mathcal{U}$ is a linear subspace, it is not empty; and since $\mathcal{U} \neq \{\underset{\sim}{0}\}$, $\mathcal{U}$ contains a nonzero (element), say $\underset{\sim}{e}_1$. By Exercise A2.2.1, $\{\underset{\sim}{e}_1\}$ is linearly independent. $\therefore$ if $\mathcal{U} \subset Lsp\{\underset{\sim}{e}_1\}$, $\{\underset{\sim}{e}_1\}$ is a basis for $\mathcal{U}$; and $\mathcal{U}$ is 1-dimensional. By Theorem A2.2.9, $\{\underset{\sim}{e}_1\}$ can be enlarged to a basis for $\mathcal{V}_n$; and the proof is complete.

If $\mathcal{U} \not\subset Lsp\{\underset{\sim}{e}_1\}$, $\exists\ \underset{\sim}{e}_2 \in \mathcal{U} \ni \underset{\sim}{e}_2 \notin Lsp\{\underset{\sim}{e}_1\}$. It is easy to show (The details are left as *Exercise* A2.2.11.) that $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2\}$ is linearly independent. If $\mathcal{U} \subset \mathrm{Lsp}\{\underset{\sim}{e}_1, \underset{\sim}{e}_2\}$, $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2\}$ is a basis for $\mathcal{U}$; and $\mathcal{U}$ is 2-dimensional. Again $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2\}$ can be extended to a basis for $\mathcal{V}_n$, and we are done.

If $\mathcal{U} \not\subset \mathrm{Lsp}\{\underset{\sim}{e}_1, \underset{\sim}{e}_2\}$, then we repeat the procedure. Since $\mathcal{U} \subset \mathcal{V}_n$, which is n-dimensional, there can be no more than n such steps. □

By Definition A2.2.2, each element of a finite-dimensional linear space can be expressed as a linear combination of basis elements. The following terminology will prove useful when we take advantage of this fact.

Definition A2.2.4. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ be a basis for an n-dimensional linear space $\mathcal{V}_n$. Then the scalars[1] $u^1, u^2, \cdots, u^n$ in the representation

$$\underset{\sim}{u} = u^1 \underset{\sim}{e}_1 + u^2 \underset{\sim}{e}_2 + \cdots + u^n \underset{\sim}{e}_n$$

of $\underset{\sim}{u} \in \mathcal{V}_n$ are the *components*[2] of $\underset{\sim}{u}$ w.r.t. the basis $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$.

---

[1] Here, *scalar* is just a synonym for real number. More generally, scalars are the elements of the scalar field which the linear space is over.

[2] Rarely, *scalar components*, and then the $u^i \underset{\sim}{e}_i$ are called *vector components*. Sometimes, *coordinates*;

We used superscripts rather than subscripts on the components for a reason which will become clear later. At this point, it seems like a silly practice; <u>adopt it anyway</u>.

The concept of components would be useless if it were not for

<u>Theorem A2.2.11.</u> <u>The components of a given element of a finite-dimensional linear space w.r.t. a given basis are unique.</u>

Proof. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ be a basis for an n-dimensional linear space $\mathcal{V}_n$. Let $\underset{\sim}{u} \in \mathcal{V}_n$, and suppose that we have both

$$\underset{\sim}{u} = u^1 \underset{\sim}{e}_1 + u^2 \underset{\sim}{e}_2 + \cdots + u^n \underset{\sim}{e}_n \quad \text{and} \quad \underset{\sim}{u} = \overset{*}{u}{}^1 \underset{\sim}{e}_1 + \overset{*}{u}{}^2 \underset{\sim}{e}_2 + \cdots + \overset{*}{u}{}^n \underset{\sim}{e}_n .$$

These two equations $\Rightarrow$

$$(\overset{*}{u}{}^1 - u^1)\underset{\sim}{e}_1 + (\overset{*}{u}{}^2 - u^2)\underset{\sim}{e}_2 + \cdots + (\overset{*}{u}{}^n - u^n)\underset{\sim}{e}_n = \underset{\sim}{0} ,$$

which in turn $\Rightarrow$

$$\overset{*}{u}{}^1 = u^1, \ \overset{*}{u}{}^2 = u^2, \ \cdots, \ \overset{*}{u}{}^n = u^n ,$$

by Theorem A2.2.2 since $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ is linearly independent. $\square$

The next result shows that addition, scalar multiplication, and equality of elements of a finite-dimensional linear space correspond to addition, scalar multiplication, and equality of their components w.r.t. a basis.

Theorem A2.2.12. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ be a basis for an n-dimensional linear space $\mathcal{V}_n$. Let $\underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{V}_n$ so that

$$\underset{\sim}{u} = u^1\underset{\sim}{e}_1 + u^2\underset{\sim}{e}_2 + \cdots + u^n\underset{\sim}{e}_n \, ; \quad \underset{\sim}{v} = v^1\underset{\sim}{e}_1 + v^2\underset{\sim}{e}_2 + \cdots + v^n\underset{\sim}{e}_n \, .$$

Then

(i) For $\alpha, \beta \in \mathbb{R}$,

$$\alpha\underset{\sim}{u} + \beta\underset{\sim}{v} = (\alpha u^1 + \beta v^1)\underset{\sim}{e}_1 + (\alpha u^2 + \beta v^2)\underset{\sim}{e}_2 + \cdots + (\alpha u^n + \beta v^n)\underset{\sim}{e}_n \, ;^{1}$$

(ii) $\underset{\sim}{0} = 0\underset{\sim}{e}_1 + 0\underset{\sim}{e}_2 + \cdots + 0\underset{\sim}{e}_n$;

(iii) $\underset{\sim}{u} = \underset{\sim}{v} \iff u^i = v^i, \; i = 1, 2, \cdots, n$.

Proof. *Exercises* A2.2.12 – 14. $\square$

This easy theorem is actually quite profound. In general, two "mathematical structures" are "essentially identical" if they can be put into one-to-one correspondence in a way that preserves their structure, i.e., they have the same mathematical structure even though the interpretations of their elements may differ. Technically, it is said that two such structures are "isomorphic", and the one-to-one correspondence which connects them is called an "isomorphism". In the case of a linear space, the (algebraic)

---

[1] I.e., in transparent notation, $(\alpha\underset{\sim}{u} + \beta\underset{\sim}{v})^i = \alpha u^i + \beta v^i$.

structure is provided by the operations of addition and scalar multiplication; and the above remarks lead us to

Definition A2.2.5. Two real linear spaces $\mathcal{U}$ and $\mathcal{V}$ are isomorphic if $\exists$ a one-to-one and onto function $f : \mathcal{U} \to \mathcal{V}$ which is additive,

$$\underset{\sim}{f}(\underset{\sim}{u}+\underset{\sim}{v}) = \underset{\sim}{f}(\underset{\sim}{u}) + \underset{\sim}{f}(\underset{\sim}{v}) \quad \forall\, \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{U},$$

and homogeneous,

$$\underset{\sim}{f}(\alpha \underset{\sim}{u}) = \alpha \underset{\sim}{f}(\underset{\sim}{u}) \quad \forall\, \underset{\sim}{u} \in \mathcal{U} \text{ and } \forall\, \alpha \in \mathbb{R}.$$

Any such function $\underset{\sim}{f}$ is an isomorphism of $\mathcal{U}$ onto $\mathcal{V}$.

Of course, the additivity is the preservation of the operation of addition, while the homogeneity is the preservation of scalar multiplication.

Theorem A2.2.13. Every n-dimensional linear space $\mathcal{V}_n$ is isomorphic to n-dimensional number space $\mathbb{R}^n$.

Proof. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ be a basis for $\mathcal{V}_n$. Then for $\underset{\sim}{u} \in \mathcal{V}_n$ $\exists$ $u^1, u^2, \cdots, u^n \in \mathbb{R} \ni$

$$\underset{\sim}{u} = u^1 \underset{\sim}{e}_1 + u^2 \underset{\sim}{e}_2 + \cdots + u^n \underset{\sim}{e}_n .$$

With this representation understood, define $\underset{\sim}{f} : V_n \to \mathbb{R}^n$ by

$$\underset{\sim}{f}(\underset{\sim}{u}) = (u^1, u^2, \cdots, u^n) .$$

Since for given $\underset{\sim}{u}$ the u^i are unique by Theorem A2.2.11, $\underset{\sim}{f}$ is, indeed, a function; cf. Theorem A1.2.1.

To see that $\underset{\sim}{f}$ is one-to-one, let $\overset{*}{\underset{\sim}{u}}, \underset{\sim}{u} \in V_n$ and suppose that $\underset{\sim}{f}(\overset{*}{\underset{\sim}{u}}) = \underset{\sim}{f}(\underset{\sim}{u})$. This $\Rightarrow$

$$(\overset{*}{u}{}^1, \overset{*}{u}{}^2, \cdots, \overset{*}{u}{}^n) = (u^1, u^2, \cdots, u^n)$$

$$\Rightarrow \overset{*}{u}{}^i = u^i, \quad i = 1, 2, \cdots, n \qquad (\text{equality in } \mathbb{R}^n)$$

$$\Rightarrow \overset{*}{\underset{\sim}{u}} = \underset{\sim}{u} \qquad (\text{Theorem A2.2.12 (iii)}).$$

$\therefore$ by Definition A1.2.3, $\underset{\sim}{f}$ is one-to-one.

To see that $\underset{\sim}{f}$ is onto, let $(u^1, u^2, \cdots, u^n) \in \mathbb{R}^n$. Then $\underset{\sim}{u} := u^1 \underset{\sim}{e}_1 + u^2 \underset{\sim}{e}_2 + \cdots + u^n \underset{\sim}{e}_n \in \mathrm{Lsp}\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\} = V_n$, and by the definition of $\underset{\sim}{f}$, $\underset{\sim}{f}(\underset{\sim}{u}) = (u^1, u^2, \cdots, u^n)$. $\therefore$ by Definition A1.2.2, $\mathbb{R}^n \subset \underset{\sim}{f}(V_n)$. But by the definition of $\underset{\sim}{f}$, $\underset{\sim}{f}(V_n) \subset \mathbb{R}^n$; so $\underset{\sim}{f}(V_n) = \mathbb{R}^n$, and $\underset{\sim}{f}$ is onto.

By Theorem A2.2.12 (i), for $\alpha, \beta \in \mathbb{R}$ and $\underset{\sim}{u}, \underset{\sim}{v} \in V_n$,

$$\alpha \underset{\sim}{u} + \beta \underset{\sim}{v} = (\alpha u^1 + \beta v^1) \underset{\sim}{e}_1 + (\alpha u^2 + \beta v^2) \underset{\sim}{e}_2 + \cdots + (\alpha u^n + \beta v^n) \underset{\sim}{e}_n .$$

Then by the definition of $\underset{\sim}{f}$,

$$\underset{\sim}{f}(\alpha\underset{\sim}{u}+\beta\underset{\sim}{v}) = (\alpha u^1+\beta v^1, \alpha u^2+\beta v^2, \cdots, \alpha u^n+\beta v^n)$$

$$= (\alpha u^1, \cdots, \alpha u^n) + (\beta v^1, \cdots, \beta v^n) \quad (\text{addition in } \mathbb{R}^n)$$

$$= \alpha(u^1, \cdots, u^n) + \beta(v^1, \cdots, v^n) \quad (\text{scalar multiplication in } \mathbb{R}^n)$$

$$= \alpha\underset{\sim}{f}(\underset{\sim}{u}) + \beta\underset{\sim}{f}(\underset{\sim}{v}) \quad (\text{definition of } \underset{\sim}{f}).$$

This gives both additivity ($\alpha=\beta=1$) and homogeneity ($\beta=0$).

Hence, $\underset{\sim}{f}$ is an isomorphism and $\mathcal{V}_n$ and $\mathbb{R}^n$ are isomorphic. □

Note that in the above proof each different choice of basis for $\mathcal{V}_n$ (and there are infinitely many) leads to a different isomorphism $\underset{\sim}{f}$ between $\mathcal{V}_n$ and $\mathbb{R}^n$. This messy feature is one of the reasons that we do not just restrict our study of n-dimensional linear spaces to $\mathbb{R}^n$ from this point forward.

Supplementary Reading

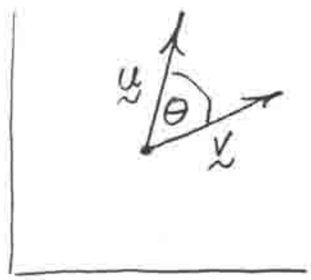
Same as for § A2.1

A2.3. Inner Product Spaces. Normed Linear Spaces

Recall again your experience with the Euclidean plane. A great help to the study of the geometry of the plane is the concept of the inner product of two vectors in the plane, which is defined as

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = |\underset{\sim}{u}||\underset{\sim}{v}| \cos\theta ,$$

where $|\underset{\sim}{u}|$ and $|\underset{\sim}{v}|$ are the lengths of the vectors, and θ is the (smaller) angle between them.



Here, we have an operation which takes two vectors and produces a scalar. What are its properties?

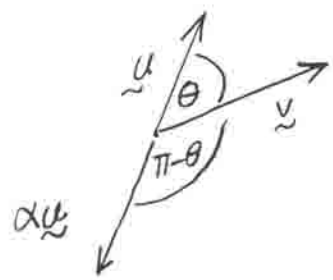
Obviously, it is commutative; i.e.,

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = \underset{\sim}{v} \cdot \underset{\sim}{u}$$

It is also homogeneous in the sense that for $\alpha \in \mathbb{R}$

$$(\alpha \underset{\sim}{u}) \cdot \underset{\sim}{v} = \alpha (\underset{\sim}{u} \cdot \underset{\sim}{v}) .$$

To see this, we (have) three cases to consider: $\alpha>0$; $\alpha<0$, $\alpha=0$. Let us look at the most difficult case, which is $\alpha<0$. In this case $\alpha\underset{\sim}{u}$ points in the direction opposite to $\underset{\sim}{u}$.



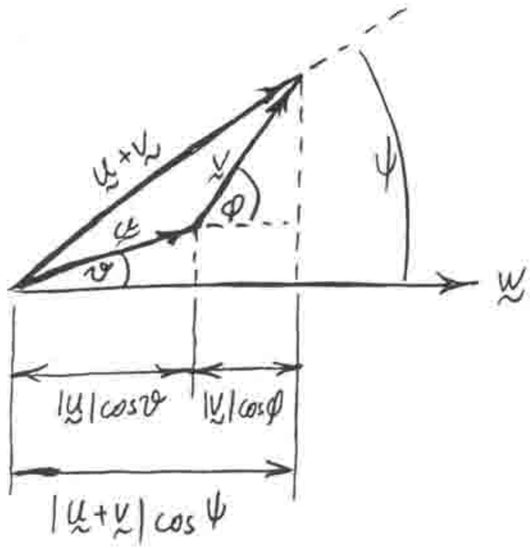
By definition,

$$
\begin{aligned}
(\alpha\underset{\sim}{u})\cdot\underset{\sim}{v} &= |\alpha\underset{\sim}{u}||\underset{\sim}{v}|\cos(\pi-\theta) \\
&= |\alpha||\underset{\sim}{u}||\underset{\sim}{v}|\cos(\pi-\theta) \quad (\text{definition of } \alpha\underset{\sim}{u}) \\
&= |\alpha||\underset{\sim}{u}||\underset{\sim}{v}|(-\cos\theta) \quad (\text{property of cosine}) \\
&= (-|\alpha|)|\underset{\sim}{u}||\underset{\sim}{v}|\cos\theta \quad (\text{associativity and commutativity in } \mathbb{R}) \\
&= \alpha|\underset{\sim}{u}||\underset{\sim}{v}|\cos\theta \quad (\text{definition of } |\alpha| \text{ for } \alpha<0) \\
&= \alpha(\underset{\sim}{u}\cdot\underset{\sim}{v}) \quad (\text{definition of } \underset{\sim}{u}\cdot\underset{\sim}{v}).
\end{aligned}
$$

The inner product of plane vectors is <u>distributive w.r.t. vector addition</u>:

$$(\underset{\sim}{u}+\underset{\sim}{v})\cdot\underset{\sim}{w} = \underset{\sim}{u}\cdot\underset{\sim}{w} + \underset{\sim}{v}\cdot\underset{\sim}{w},$$

A "proof" of this fact is depicted below.



Finally, since $\underset{\sim}{u}\cdot\underset{\sim}{u} = |\underset{\sim}{u}|^2$, we see that the inner product is positive definite in the sense that

$$\underset{\sim}{u}\cdot\underset{\sim}{u} \geq 0 \quad \text{and} \quad \underset{\sim}{u}\cdot\underset{\sim}{u} = 0 \text{ iff } \underset{\sim}{u} = \underset{\sim}{0}\ .$$

These are the key properties of the inner product of vectors in the Euclidean plane. For the general case, we have

Definition A2.3.1. A real inner product space[1] is a real linear space, say V, equipped with a function on $V \times V$ to $\mathbb{R}$, denoted by

$$(\underset{\sim}{u}, \underset{\sim}{v}) \longmapsto \underset{\sim}{u}\cdot\underset{\sim}{v} \qquad)^2$$

---

[1] Sometimes, Euclidean space.

[2] Mathematicians usually write $(\underset{\sim}{u}, \underset{\sim}{v})$ or $\langle \underset{\sim}{u}, \underset{\sim}{v} \rangle$ for $\underset{\sim}{u}\cdot\underset{\sim}{v}$.

with the function value $\underset{\sim}{u} \cdot \underset{\sim}{v}$ called the inner product[1] of $\underset{\sim}{u}$ and $\underset{\sim}{v}$, which satisfies the following axioms:

(I1) Commutativity. $\forall\, \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{V}$

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = \underset{\sim}{v} \cdot \underset{\sim}{u} \; ;$$

(I2) Homogeneity. $\forall\, \alpha \in \mathbb{R}$ and $\forall\, \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{V}$

$$(\alpha \underset{\sim}{u}) \cdot \underset{\sim}{v} = \alpha\, (\underset{\sim}{u} \cdot \underset{\sim}{v}) \; ;$$

(I3) Distributivity w.r.t. addition. $\forall\, \underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{w} \in \mathcal{V}$

$$(\underset{\sim}{u} + \underset{\sim}{v}) \cdot \underset{\sim}{w} = \underset{\sim}{u} \cdot \underset{\sim}{w} + \underset{\sim}{v} \cdot \underset{\sim}{w} \; ;$$

(I4) Positive Definiteness. $\forall\, \underset{\sim}{u} \in \mathcal{V}$

$$\underset{\sim}{u} \cdot \underset{\sim}{u} \geq 0 \text{ with } \underset{\sim}{u} \cdot \underset{\sim}{u} = 0 \text{ only if } \underset{\sim}{u} = \underset{\sim}{0} .$$

For a complex linear space, $\underset{\sim}{u} \cdot \underset{\sim}{v} \in \mathbb{C}$, and the commutativity axiom is modified to $\underset{\sim}{u} \cdot \underset{\sim}{v} = \overline{\underset{\sim}{v} \cdot \underset{\sim}{u}}$. This makes $\underset{\sim}{u} \cdot \underset{\sim}{u} \in \mathbb{R}$, so the inequality in the positive definiteness axiom makes sense. Complex inner product spaces are often called unitary linear spaces.

We have already seen that the set of vectors in the

---

[1] Often, scalar product or dot product.

Euclidean plane is an inner product space. Another important example is $\mathbb{R}^n$.

Theorem A2.3.1. The linear space

$$\mathbb{R}^n = \{(u_1, u_2, \cdots, u_n) : u_i \in \mathbb{R},\ i = 1, 2, \cdots, n\} \qquad)^1$$

with $\underset{\sim}{u} \cdot \underset{\sim}{v}$ defined by

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

is an inner product space. This inner product for $\mathbb{R}^n$ is called the standard inner product.

Proof. It is clear [2] that

$$(\underset{\sim}{u}, \underset{\sim}{v}) \longmapsto u_1 v_1 + u_2 v_2 + \cdots + u_n v_n =: \underset{\sim}{u} \cdot \underset{\sim}{v}$$

defines a function on $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$. It is left for the student as Exercise A2.3.1 to show that axioms (I1)–(I3) are satisfied.

To verify the positive definiteness, we note that

$$\underset{\sim}{u} \cdot \underset{\sim}{u} = u_1^2 + u_2^2 + \cdots + u_n^2$$

---

[1] Cf. Theorem A2.1.1.
[2] Cf. Theorem A1.2.1.

— a sum of squares. $\therefore$

$$\underset{\sim}{u} \cdot \underset{\sim}{u} \geq 0 \quad \forall \; \underset{\sim}{u} \in \mathbb{R}^n.$$

Since there is no possibility of cancellation,

$$\underset{\sim}{u} \cdot \underset{\sim}{u} = u_1^2 + u_2^2 + \cdots + u_n^2 = 0 \Rightarrow u_i = 0 \;, i = 1, 2, \cdots, n \Rightarrow \underset{\sim}{u} = \underset{\sim}{0},$$

because the zero element of $\mathbb{R}^n$ is $(0, 0, \cdots, 0)$. Thus,

$$\underset{\sim}{u} \cdot \underset{\sim}{u} = 0 \text{ only if } \underset{\sim}{u} = \underset{\sim}{0},$$

and (I4) is met. $\square$

The next exercise emphasizes that the standard inner product for $\mathbb{R}^n$ is not the only choice of inner product for $\mathbb{R}^n$.

Exercise A2.3.2. Let the $n \times n$ real matrix $[a_{ij}]$ be symmetric and positive definite. Look up the meanings of these terms in the matrix context if you do not know them. With reference to the notation of Theorem A2.3.1, show that

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} u_i v_j$$

defines an inner product for $\mathbb{R}^n$.

Exercise A2.3.3. In the spirit of Exercises A2.1.2 and A2.2.7, view $\mathbb{R}$ as a 1-dimensional linear space. Show that

$$\langle x, y \rangle = xy$$

defines an inner product on $\mathbb{R}$. Here, the notation $x \cdot y$ is avoided because of possible confusion with the product xy.

Since every n-dimensional linear space is isomorphic to $\mathbb{R}^n$ (cf. Theorem A2.2.13) and since $\mathbb{R}^n$ can be made an inner product space in many ways, the following result is not unexpected.

Theorem A2.3.2. Every n-dimensional linear space $\mathcal{V}_n$ can be assigned an inner product.

Proof. We make the same identification between $\mathcal{V}_n$ and $\mathbb{R}^n$ that was used to establish that they are isomorphic. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ be a basis for $\mathcal{V}_n$. Then for $\underset{\sim}{u} \in \mathcal{V}_n$ $\exists$ $u^1, u^2, \cdots, u^n \in \mathbb{R}$ $\ni$

$$\underset{\sim}{u} = u^1\underset{\sim}{e}_1 + u^2\underset{\sim}{e}_2 + \cdots + u^n\underset{\sim}{e}_n.$$

With this representation understood, define

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = u^1v^1 + u^2v^2 + \cdots + u^nv^n.$$

The remaining steps are identical to those used to prove

Theorem A2.3.1. Here, the meanings of $\underset{\sim}{u}+\underset{\sim}{v}$ and $\alpha\underset{\sim}{u}$ in terms of components follow from Theorem A2.2.12 as does the equivalence $\underset{\sim}{u}=\underset{\sim}{0} \Leftrightarrow u^i=0$, $i=1,2,\cdots,n$. □

Of course, in the above proof, each different choice of basis can be expected to lead to a different inner product. Surprisingly, we shall see in § A2.5 that for a fairly large class of bases these inner products will be identical.

So far, all of our examples of inner product spaces have been finite-dimensional. The next exercise shows that infinite-dimensional spaces also can have an inner product.

Exercise A2.3.4. Consider the infinite-dimensional linear space $C^0([a,b])$ of Exercise A2.1.3. Show that

$$<f,g> = \int_a^b f(x)\,g(x)\,dx$$

defines an inner product for $C^0([a,b])$. (Strictly speaking, we haven't shown that this space is infinite-dimensional, but don't spend too much time looking for a basis.)

The next three easy theorems provide useful results that apply to all inner product spaces.

Theorem A2.3.3. Let V be an inner product space. Then

$$\underset{\sim}{0}\cdot\underset{\sim}{u}=0 \quad \forall\, \underset{\sim}{u}\in V.$$

In particular,

$$\underset{\sim}{0} \cdot \underset{\sim}{0} = 0.$$

Proof. Exercise A2.3.5. (Hint: Make intelligent choices of the variables in either the homogeneity property or the distributivity property.) □

Since we always have $\underset{\sim}{0} \cdot \underset{\sim}{0} = 0$ in an inner product space, the positive definiteness axiom is sometimes stated in the redundant fashion:

$$\forall\, \underset{\sim}{u} \in \mathcal{V},\ \underset{\sim}{u} \cdot \underset{\sim}{u} \geq 0 \quad \text{with } \underset{\sim}{u} \cdot \underset{\sim}{u} = 0 \text{ iff } \underset{\sim}{u} = \underset{\sim}{0}.$$

Theorem A2.3.4. Let $\underset{\sim}{u} \in \mathcal{V}$, where $\mathcal{V}$ is an inner product space. Then

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = 0 \quad \forall\, \underset{\sim}{v} \in \mathcal{V} \Rightarrow \underset{\sim}{u} = \underset{\sim}{0}.$$

Proof. Since $\underset{\sim}{u} \cdot \underset{\sim}{v} = 0$ must hold $\forall\, \underset{\sim}{v} \in \mathcal{V}$, it must hold in particular for $\underset{\sim}{v} = \underset{\sim}{u}$. Hence, $\underset{\sim}{u} \cdot \underset{\sim}{u} = 0 \Rightarrow \underset{\sim}{u} = \underset{\sim}{0}$ by the positive definiteness of the inner product. □

INSERT material on p. A2.3.9a

In our motivational example of vectors in the Euclidean plane, $\underset{\sim}{u} \cdot \underset{\sim}{u} = |\underset{\sim}{u}|^2$, where $|\underset{\sim}{u}|$ is the length of the vector $\underset{\sim}{u}$. This leads us to

Definition A2.3.2. The magnitude[1] of an element $\underset{\sim}{u}$ of

---

[1] Often, length, modulus, norm.

INSERT for p. A2.3.9

As expected, cf. Theorems A1.3.1, A2.1.8, and A2.1.9, when equal elements of an inner product space are "dotted" with equals, the resulting scalars are equal.

Theorem A2.3.5. Let $\underset{\sim}{u}$, $\underset{\sim}{v}$, $\underset{\sim}{w}$, and $\underset{\sim}{x}$ be elements of an inner product space. Then

$$\underset{\sim}{u} = \underset{\sim}{v} \text{ and } \underset{\sim}{w} = \underset{\sim}{x} \implies \underset{\sim}{u} \cdot \underset{\sim}{w} = \underset{\sim}{v} \cdot \underset{\sim}{x}.$$

Proof. *Exercise* A2.3.6. □

Return to p. A2.3.9

of an inner product space is the is the nonnegative number

$$|\underset{\sim}{u}| = \sqrt{\underset{\sim}{u}\cdot\underset{\sim}{u}}\ . \quad)^{1,2}$$

Of course, this definition would not make sense if it were not for the fact that $\underset{\sim}{u}\cdot\underset{\sim}{u} \geq 0$ by the positive definiteness of the inner product.

Exercise A2.3.7. In the spirit of Exercises A2.1.2, A2.2.7, and A2.3.3, view $\mathbb{R}$ as the 1-dimensional inner product space with the inner product $\langle x,y\rangle = xy$. Show that the magnitude of x is the absolute value of x. I.e, $|x| = |x|$, where on the l.h.s. $|\cdot|$ means magnitude, while on the r.h.s. $|\cdot|$ means absolute value. What good fortune!

It is important to realize that in the definition of magnitude and in most of what follows the underlying inner product space need not be finite-dimensional.

Of course, the operation of calculating the magnitude of an element of an inner product space can be thought of as a function on the space to the nonnegative reals. This is our viewpoint in stating the following important properties.

---

[1] We will always use $\sqrt{\cdot}$ to denote the positive square root.

[2] Mathematicians usually write $\|\underset{\sim}{u}\|$ for $|\underset{\sim}{u}|$ (cf. Definition A2.3.3 and Theorem A2.3.9); engineers often use $|\underset{\sim}{u}| = u$.

Theorem A2.3.6. On any inner product space $\mathcal{V}$, magnitude is positive homogeneous,

(i) $|\alpha \underset{\sim}{u}| = |\alpha||\underset{\sim}{u}| \quad \forall \alpha \in \mathbb{R}$ and $\forall \underset{\sim}{u} \in \mathcal{V}$,

and positive definite,

(ii) $|\underset{\sim}{u}| \geq 0 \quad \forall \underset{\sim}{u} \in \mathcal{V}$ with $|\underset{\sim}{u}| = 0$ iff $\underset{\sim}{u} = \underset{\sim}{0}$.

Proof. Exercises A2.3.8 and 9. $\square$

The following inequality has major applications in widely different contexts.

Theorem A2.3.7. (Inner Product-Magnitude Inequality)[1] Let $\mathcal{V}$ be an inner product space. Then

$$|\underset{\sim}{u} \cdot \underset{\sim}{v}| \leq |\underset{\sim}{u}||\underset{\sim}{v}| \quad \forall \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{V}.$$

Proof. If either $\underset{\sim}{u} = \underset{\sim}{0}$ or $\underset{\sim}{v} = \underset{\sim}{0}$, it follows from our previous results that the inequality above is satisfied as an equality. $\therefore$ we assume that $\underset{\sim}{u} \neq \underset{\sim}{0}$ and $\underset{\sim}{v} \neq \underset{\sim}{0}$ in the remainder of the proof.

By the positive definiteness of the inner product,

Almost always labled with some combination of the names CAUCHY, BUNYAKOVSKII, and SCHWARZ.

$$(\alpha \underset{\sim}{u} + \beta \underset{\sim}{v}) \cdot (\alpha \underset{\sim}{u} + \beta \underset{\sim}{v}) \geq 0 \quad \forall \alpha, \beta \in \mathbb{R}$$

Expanding this (using distributivity, etc.), we find that

$$\alpha^2 |\underset{\sim}{u}|^2 + 2\alpha\beta\, \underset{\sim}{u} \cdot \underset{\sim}{v} + \beta^2 |\underset{\sim}{v}|^2 \geq 0.$$

Choosing $\alpha = |\underset{\sim}{v}|^2$ and $\beta = -\underset{\sim}{u} \cdot \underset{\sim}{v}$ yields

$$\left[|\underset{\sim}{u}|^2 |\underset{\sim}{v}|^2 - (\underset{\sim}{u} \cdot \underset{\sim}{v})^2 \right] |\underset{\sim}{v}|^2 \geq 0.$$

Since $|\underset{\sim}{v}|^2 > 0$, this $\Rightarrow$

$$|\underset{\sim}{u}|^2 |\underset{\sim}{v}|^2 \geq (\underset{\sim}{u} \cdot \underset{\sim}{v})^2 \;;$$

and since the square root function is monotonic, this $\Rightarrow$

$$|\underset{\sim}{u}||\underset{\sim}{v}| \geq \sqrt{(\underset{\sim}{u} \cdot \underset{\sim}{v})^2} = |\underset{\sim}{u} \cdot \underset{\sim}{v}| \;. \;\square$$

It is interesting to note that when the inner product-magnitude inequality is applied to $\mathbb{R}^n$ with the standard inner product (cf. Theorem A2.3.1) and to $C^0[(a,b)]$ with the inner product introduced in Exercise A2.3.4, the following seemingly unrelated results are obtained:

$$\left| \sum_{i=1}^{n} a_i b_i \right| \leq \left(\sum_{i=1}^{n} a_i^2 \right)^{1/2} \left(\sum_{i=1}^{n} b_i^2 \right)^{1/2} \qquad)^1$$

---

[1] Here, of course, $(\cdot)^{1/2} := \sqrt{\cdot}$.

for any lists $(a_1, a_2, \cdots, a_n)$ and $(b_1, b_2, \cdots, b_n)$ of real numbers; and

$$\left| \int_a^b f(x)g(x)\,dx \right| \le \left(\int_a^b [f(x)]^2 dx \right)^{1/2} \left(\int_a^b [g(x)]^2 dx \right)^{1/2}$$

for any real-valued functions f and g continuous on $[a,b]$. This is indicative of the power of abstract methods.

In using the inner product - magnitude inequality, it is often helpful to recognize that

$$\underset{\sim}{u} \cdot \underset{\sim}{v} \le |\underset{\sim}{u} \cdot \underset{\sim}{v}| \le |\underset{\sim}{u}||\underset{\sim}{v}| ;$$

in fact, we shall do this in our proof of

Theorem A2.3.8. (**Triangle Inequality**) Let $\mathcal{V}$ be an inner product space. Then

$$|\underset{\sim}{u} + \underset{\sim}{v}| \le |\underset{\sim}{u}| + |\underset{\sim}{v}| \quad \forall\, \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{V}.$$

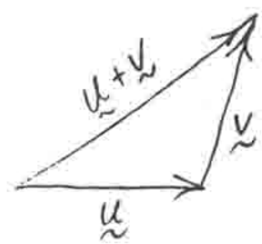
Proof.
$$\begin{aligned}
|\underset{\sim}{u} + \underset{\sim}{v}|^2 &= (\underset{\sim}{u} + \underset{\sim}{v}) \cdot (\underset{\sim}{u} + \underset{\sim}{v}) && \text{(definition of } |\cdot| \text{ in } \mathcal{V}) \\
&= |\underset{\sim}{u}|^2 + 2\underset{\sim}{u} \cdot \underset{\sim}{v} + |\underset{\sim}{v}|^2 && \text{(properties of inner product)} \\
&\le |\underset{\sim}{u}|^2 + 2|\underset{\sim}{u} \cdot \underset{\sim}{v}| + |\underset{\sim}{v}|^2 && \text{(property of } |\cdot| \text{ in } \mathbb{R}) \\
&\le |\underset{\sim}{u}|^2 + 2|\underset{\sim}{u}||\underset{\sim}{v}| + |\underset{\sim}{v}|^2 && \text{(inner product - magnitude ineq.)}
\end{aligned}$$

$$= \left(|\underset{\sim}{u}| + |\underset{\sim}{v}|\right)^2 \quad (\text{arithmetic})$$

$$\Rightarrow$$

$$|\underset{\sim}{u} + \underset{\sim}{v}| \leq |\underset{\sim}{u}| + |\underset{\sim}{v}| \quad (\text{monotonicity of } \sqrt{\cdot}\,). \ \square$$

On recalling the rule for adding <u>vectors</u> in the <u>Euclidean plane</u>,



we see why the triangle inequality is so-named. That it is instinctively known to all school-aged children is a fact readily attested to by any residential homeowner with a fenceless corner lot.

Now that we have the triangle inequality, we are at a good place to introduce

<u>Definition A2.3.3.</u> <u>A normed linear space</u> <u>is a real linear space</u>[1], <u>say</u> $\mathcal{V}$, <u>equipped with a function on</u> $\mathcal{V}$ <u>to</u> $\mathbb{R}^+$)[2], <u>denoted by</u>

$$\underset{\sim}{u} \longmapsto \|\underset{\sim}{u}\|$$

---

[1] Not necessarily an inner product space.

[2] $\mathbb{R}^+$ denotes the nonnegative reals; i.e., $\mathbb{R}^+ := \{x \in \mathbb{R} : x \geq 0\}$.

with the function value $\|\underset{\sim}{u}\|$ called the norm of $\underset{\sim}{u}$, which satisfies the following axioms:

(N1) (Positive) Homogeneity. $\forall \alpha \in \mathbb{R}$ and $\forall \underset{\sim}{u} \in \mathcal{V}$

$$\|\alpha \underset{\sim}{u}\| = |\alpha| \|\underset{\sim}{u}\| \, ;$$

(N2) Positive Definiteness. $\forall \underset{\sim}{u} \in \mathcal{V}$

$$\|\underset{\sim}{u}\| \geq 0 \text{ with } \|\underset{\sim}{u}\| = 0 \text{ iff } \underset{\sim}{u} = \underset{\sim}{0} \, ;$$

(N3) Triangle Inequality. $\forall \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{V}$

$$\|\underset{\sim}{u} + \underset{\sim}{v}\| \leq \|\underset{\sim}{u}\| + \|\underset{\sim}{v}\| \, .$$

An immediate consequence of this definition together with Theorems A2.3.6 and 8 is

Theorem A2.3.9. An inner product space is a normed linear space under the inner product norm

$$\|\underset{\sim}{u}\| := |\underset{\sim}{u}| = \sqrt{\underset{\sim}{u} \cdot \underset{\sim}{u}} \, .$$

$\boxed{\text{INSERT material on pp. A2.3.15a, b}}$

The inner product – magnitude inequality also can be used to define the "angle" between two elements of an inner product space. To see what to do here, we go back once again to the familiar example of vectors in the Euclidean plane.

INSERT on p. A2.3.15 just after Theorem A2.3.9.

The following generalization of the triangle inequality is often useful.

Theorem A2.3.9. Let $\mathcal{V}$ be a normed linear space. Then $\forall\ \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{V}$

$$\text{(i)} \quad \big|\, \|\underset{\sim}{u}\| - \|\underset{\sim}{v}\| \,\big| \le \|\underset{\sim}{u} + \underset{\sim}{v}\| \le \|\underset{\sim}{u}\| + \|\underset{\sim}{v}\|$$

and

$$\text{(ii)} \quad \big|\, \|\underset{\sim}{u}\| - \|\underset{\sim}{v}\| \,\big| \le \|\underset{\sim}{u} - \underset{\sim}{v}\| \le \|\underset{\sim}{u}\| + \|\underset{\sim}{v}\|.$$

Proof. The second inequality in (i) is exactly the triangle inequality. To get the first, consider

$$\begin{aligned}
\|\underset{\sim}{u}\| &= \|\underset{\sim}{u} + \underset{\sim}{v} - \underset{\sim}{v}\| && \text{(add and subtract } \underset{\sim}{v}\text{)} \\
&= \|(\underset{\sim}{u} + \underset{\sim}{v}) - \underset{\sim}{v}\| && \text{(associativity)} \\
&\le \|\underset{\sim}{u} + \underset{\sim}{v}\| + \|-\underset{\sim}{v}\| && \text{(triangle inequality)} \\
&= \|\underset{\sim}{u} + \underset{\sim}{v}\| + \|\underset{\sim}{v}\| && (\|-\underset{\sim}{v}\| = \|\underset{\sim}{v}\|)
\end{aligned}$$

$\Longrightarrow$

$$\|\underset{\sim}{u}\| - \|\underset{\sim}{v}\| \le \|\underset{\sim}{u} + \underset{\sim}{v}\|.$$

Interchanging the roles of $\underset{\sim}{u}$ and $\underset{\sim}{v}$, we obtain

$$\begin{aligned}
\|\underset{\sim}{v}\| - \|\underset{\sim}{u}\| &\le \|\underset{\sim}{v} + \underset{\sim}{u}\| \\
&= \|\underset{\sim}{u} + \underset{\sim}{v}\| && \text{(commutativity)}.
\end{aligned}$$

Thus, we have both

$$\|\underset{\sim}{u}\| - \|\underset{\sim}{v}\| \leq \|\underset{\sim}{u} + \underset{\sim}{v}\|$$

and

$$-\left(\|\underset{\sim}{u}\| - \|\underset{\sim}{v}\|\right) \leq \|\underset{\sim}{u} + \underset{\sim}{v}\| \quad ;$$

whence

$$\left| \|\underset{\sim}{u}\| - \|\underset{\sim}{v}\| \right| \leq \|\underset{\sim}{u} + \underset{\sim}{v}\| ,$$

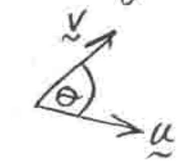
which is the first inequality in (i).

Inequalities (ii) are established by replacing $\underset{\sim}{v}$ with $-\underset{\sim}{v}$ in (i) and using $\|-\underset{\sim}{v}\| = \|\underset{\sim}{v}\|$. $\square$

Return to p. A2.3.15.

In that case, the <u>angle</u> θ between two vectors u and v was a primitive concept; and the inner product was defined as

$$u \cdot v = |u||v| \cos\theta .$$



For $u, v \neq 0$ (And, of course, the notion of angle makes no sense if one or both of the vectors is the zero vector.), we have

$$\cos\theta = \frac{u \cdot v}{|u||v|} .$$

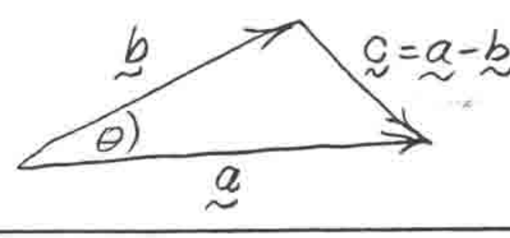
Now the r.h.s. makes sense in any inner product space; but to be the cosine of some angle, it must be in the interval $[-1, 1]$. This is guaranteed by the inner product-magnitude inequality. Thus, we are lead to

<u>Definition</u> A2.3.4. <u>Let u and v be two nonzero elements of an inner product space. Then the angle θ between them is defined by</u>

$$\cos\theta = \frac{u \cdot v}{|u||v|} , \quad 0 \le \theta \le \pi .$$

<u>Exercise</u> A2.3.10. Let a and b be nonzero elements of any inner product space. Let θ be the angle between them, and write $c = a - b$. Establish the <u>law of cosines</u>:

$$|c|^2 = |a|^2 + |b|^2 - 2|a||b|\cos\theta .$$

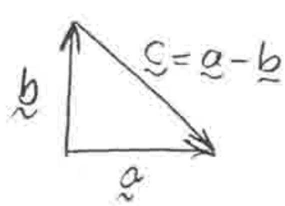


While it is amusing to be able to speak of the angle between two elements of spaces such as $C^0([a,b])$ where we usually do not think geometrically, it is not of much utility except for suggesting the definitions of "orthogonal" $(\theta = \pi/2)$, "parallel" $(\theta = 0 \text{ or } \pi)$, and "same direction" $(\theta = 0)$. The formal definitions below are not given directly in terms of θ, because this would restrict us to nonzero elements.

Definition A2.3.5. Two elements $\mathbf{u}$ and $\mathbf{v}$ of an inner product space are *orthogonal*[1] if $\mathbf{u}\cdot\mathbf{v} = 0$.

Exercise A2.3.11. Let $\mathbf{a}$ and $\mathbf{b}$ be orthogonal elements of any inner product space, and set $\mathbf{c} = \mathbf{a} - \mathbf{b}$. Prove the *Pythagorean Theorem*:

$$|\mathbf{c}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2.$$



Continuing in this same line of thought, parallel should mean $\theta = 0$ or π, which implies $\cos\theta = \pm 1$. Then the definition of angle gives $\mathbf{u}\cdot\mathbf{v} = \pm|\mathbf{u}||\mathbf{v}|$, which corresponds to equality in the inner product – magnitude inequality. But we can do better than this if we take a slightly different line of thought. For vectors in the Euclidean plane, parallel vectors lie along the same line; i.e., one is a scalar multiple

---

[1] Sometimes, *perpendicular*.

of the other. This leads us to the following definition which has the advantage of not being restricted to inner product spaces.

Definition A2.3.6. Two elements $\underset{\sim}{u}$ and $\underset{\sim}{v}$ of any linear space (not necessarily an inner product space) are *parallel*[1] if the set $\{\underset{\sim}{u}, \underset{\sim}{v}\}$ is linearly dependent.

In view of the preceding remarks, the following result is not surprising.

Theorem A2.3.11. Two elements $\underset{\sim}{u}$ and $\underset{\sim}{v}$ of an inner product space are parallel iff $\underset{\sim}{u} \cdot \underset{\sim}{v} = \pm |\underset{\sim}{u}||\underset{\sim}{v}|$.

Proof. *Exercise* A2.3.12. (Hint: Parallel $\Leftrightarrow$ linear dependence $\Leftrightarrow \exists\, \alpha, \beta$ not both zero $\ni \alpha \underset{\sim}{u} + \beta \underset{\sim}{v} = \underset{\sim}{0}$. In the difficult direction ($\underset{\sim}{u} \cdot \underset{\sim}{v} = \pm |\underset{\sim}{u}||\underset{\sim}{v}| \Rightarrow$ parallel), use the positive definiteness, $(\alpha \underset{\sim}{u} + \beta \underset{\sim}{v}) \cdot (\alpha \underset{\sim}{u} + \beta v) \geq 0$ with $= 0$ iff $\alpha \underset{\sim}{u} + \beta \underset{\sim}{v} = \underset{\sim}{0}$, together with an intelligent choice of α and β.) □

The next theorem is intuitively obvious, even if its proof is not.

Theorem A2.3.12. Let $\mathcal{V}$ be an inner product space. If $\underset{\sim}{u} \in \mathcal{V}$ is orthogonal to all $\underset{\sim}{v} \in \mathcal{V}$ which are orthogonal to $\underset{\sim}{w} \in \mathcal{V}$, then $\underset{\sim}{u}$ is parallel to $\underset{\sim}{w}$.

---

[1] Often, *collinear*.

Proof[1]: If $\underset{\sim}{w} = \underset{\sim}{0}$, then there is nothing to prove because, in view of Theorem A2.2.3, all elements of V are parallel to $\underset{\sim}{0}$. $\therefore$ we assume that $\underset{\sim}{w} \neq \underset{\sim}{0}$ for the remainder of the proof.

Let $\alpha, \beta \in \mathbb{R}$ and consider

$$(*) \qquad (\alpha \underset{\sim}{u} + \beta \underset{\sim}{w}) \cdot \underset{\sim}{v} = \alpha\, \underset{\sim}{u} \cdot \underset{\sim}{v} + \beta\, \underset{\sim}{w} \cdot \underset{\sim}{v} = 0$$

by hypothesis. With the particular choice

$$\alpha' = 1 \quad \text{and} \quad \beta' = -\frac{\underset{\sim}{u} \cdot \underset{\sim}{w}}{\underset{\sim}{w} \cdot \underset{\sim}{w}},$$

we also have

$$(\alpha' \underset{\sim}{u} + \beta' \underset{\sim}{w}) \cdot \underset{\sim}{w} = \left(\underset{\sim}{u} - \frac{\underset{\sim}{u} \cdot \underset{\sim}{w}}{\underset{\sim}{w} \cdot \underset{\sim}{w}} \underset{\sim}{w} \right) \cdot \underset{\sim}{w} = 0 .$$

Hence, $\alpha' \underset{\sim}{u} + \beta' \underset{\sim}{w}$ is orthogonal to $\underset{\sim}{w}$; i.e., it is one of the $\underset{\sim}{v}$'s. Now return to $(*)$ and take $\underset{\sim}{v} = \alpha' \underset{\sim}{u} + \beta' \underset{\sim}{w}$, $\alpha = \alpha'$, and $\beta = \beta'$ to get

$$(\alpha' \underset{\sim}{u} + \beta' \underset{\sim}{w}) \cdot (\alpha' \underset{\sim}{u} + \beta' \underset{\sim}{w}) = 0.$$

Then by the positive definiteness of the inner product,

$$\alpha' \underset{\sim}{u} + \beta' \underset{\sim}{w} = \underset{\sim}{0}.$$

$\therefore$ the subset $\{\underset{\sim}{u}, \underset{\sim}{w}\}$ is linearly dependent; i.e., $\underset{\sim}{u}$ and $\underset{\sim}{w}$ are parallel. $\square$

---

[1] This proof was shown to the author by Dr. Sang Lee.

Of course, for two elements $\underset{\sim}{u}$ and $\underset{\sim}{v}$ to have the same direction should mean that the angle θ between them be zero. Then

$$\cos\theta = \frac{\underset{\sim}{u}\cdot\underset{\sim}{v}}{|\underset{\sim}{u}||\underset{\sim}{v}|}$$

leads us to

Definition A2.3.7. Two elements $\underset{\sim}{u}$ and $\underset{\sim}{v}$ of an inner product space have the same direction if $\underset{\sim}{u}\cdot\underset{\sim}{v} = |\underset{\sim}{u}||\underset{\sim}{v}|$.

Obviously, two elements that have the same direction are necessarily parallel, and the zero element has the same direction as every other element of the underlying inner product space.

The following necessary and sufficient condition for equality of two elements of an inner product space agrees with the customary definition of equality between ordinary geometrical vectors.

Theorem A2.3.13. Two elements of an inner product space are equal iff they have the same magnitude and the same direction; i.e.,

$$\underset{\sim}{u} = \underset{\sim}{v} \iff |\underset{\sim}{u}| = |\underset{\sim}{v}| \text{ and } \underset{\sim}{u}\cdot\underset{\sim}{v} = |\underset{\sim}{u}||\underset{\sim}{v}|.$$

Proof. Suppose $\underset{\sim}{u} = \underset{\sim}{v}$. Then by substitution, $|\underset{\sim}{u}| = |\underset{\sim}{v}|$. Also by definition,

$$|\underset{\sim}{u}||\underset{\sim}{u}| = \underset{\sim}{u}\cdot\underset{\sim}{u} \;;$$

and then by substitution this can be expressed as

$$|\underset{\sim}{u}||\underset{\sim}{v}| = \underset{\sim}{u} \cdot \underset{\sim}{v}.$$

Conversely, suppose $|\underset{\sim}{u}| = |\underset{\sim}{v}|$ and $\underset{\sim}{u} \cdot \underset{\sim}{v} = |\underset{\sim}{u}||\underset{\sim}{v}|$. Consider

$$\begin{aligned}
|\underset{\sim}{u} - \underset{\sim}{v}|^2 &= (\underset{\sim}{u} - \underset{\sim}{v}) \cdot (\underset{\sim}{u} - \underset{\sim}{v}) \quad (\text{definition of } |\cdot|) \\
&= \underset{\sim}{u} \cdot \underset{\sim}{u} - 2\underset{\sim}{u} \cdot \underset{\sim}{v} + \underset{\sim}{v} \cdot \underset{\sim}{v} \quad (\text{properties of inner product}) \\
&= |\underset{\sim}{u}|^2 - 2\underset{\sim}{u} \cdot \underset{\sim}{v} + |\underset{\sim}{v}|^2 \quad (\text{definition of } |\cdot|) \\
&= |\underset{\sim}{u}|^2 - 2|\underset{\sim}{u}||\underset{\sim}{v}| + |\underset{\sim}{v}|^2 \quad (\underset{\sim}{u} \cdot \underset{\sim}{v} = |\underset{\sim}{u}||\underset{\sim}{v}|) \\
&= 0 \quad (|\underset{\sim}{u}| = |\underset{\sim}{v}|).
\end{aligned}$$

By positive definiteness, $|\underset{\sim}{u} - \underset{\sim}{v}| = 0 \Rightarrow$

$$\underset{\sim}{u} - \underset{\sim}{v} = \underset{\sim}{0};$$

from which we conclude (Carry out the details as Exercise A2.3.13.) that

$$\underset{\sim}{u} = \underset{\sim}{v}. \quad \square$$

An especially important application of the concept of orthogonal elements of an inner product space is given in

Definition A2.3.8. A subset $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$ of an inner product space is *orthonormal* if $\forall\, i,j = 1, 2, \cdots, k$

$$\underset{\sim}{u}_i \cdot \underset{\sim}{u}_j = \delta_{ij} := \begin{cases} 1 \text{ if } i=j \\ 0 \text{ if } i \neq j \end{cases} .$$ [1]

Thus, the elements of an orthonormal subset are mutually orthogonal, and they are normalized in the sense that they are each of unit magnitude.

Theorem A2.3.14. An orthonormal subset of an inner product space is linearly independent.

Proof. We use Theorem A2.2.2. Accordingly, suppose that

$$\alpha_1 \underset{\sim}{u}_1 + \alpha_2 \underset{\sim}{u}_2 + \cdots + \alpha_k \underset{\sim}{u}_k = \underset{\sim}{0} ,$$

where $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_k\}$ is the orthonormal subset under consideration. Take the inner product of this equation with $\underset{\sim}{u}_i$)[2] and conclude,

---

[1] The symbol δ_{ij} so defined is *Kronecker's delta*.

[2] This, of course, is permitted, cf. Theorem A2.3.5.

with the aid of the distributivity and homogeneity of the inner product and Theorem A2.3.3, that $\alpha_i = 0$, $i = 1, 2, \cdots, k$. □

Since computations involving linear combinations of orthonormal subsets are particularly simple, the following theorem and the alogrithm used to prove it are of considerable significance.

Theorem A2.3.15. Every finite-dimensional inner product space has an orthonormal basis.

Proof. Let V_n denote any n-dimensional inner product space. By definition, V_n has a basis, say $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$. We use the Gram-Schmidt procedure to construct an orthonormal basis from $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$.

First set $\underset{\sim}{a}_1 = \underset{\sim}{e}_1$, and define $\underset{\sim}{a}_2 = \underset{\sim}{e}_2 + \alpha \underset{\sim}{a}_1$, where $\alpha \in \mathbb{R}$. The elements $\underset{\sim}{a}_1$ and $\underset{\sim}{a}_2$ will be orthogonal if

$$\underset{\sim}{a}_1 \cdot \underset{\sim}{a}_2 = \underset{\sim}{a}_1 \cdot \underset{\sim}{e}_2 + \alpha\, \underset{\sim}{a}_1 \cdot \underset{\sim}{a}_1 = 0 .$$

Since $\underset{\sim}{a}_1 = \underset{\sim}{e}_1$, $\underset{\sim}{a}_1 \neq \underset{\sim}{0}$ by Theorem A2.2.3, and then $\underset{\sim}{a}_1 \cdot \underset{\sim}{a}_1 > 0$ by the positive definiteness of the inner product. ∴ we can solve this equation for α. The element $\underset{\sim}{a}_2$ so constructed will be $\neq \underset{\sim}{0}$ because $\underset{\sim}{a}_2 = \underset{\sim}{0} \Rightarrow \{\underset{\sim}{e}_1, \underset{\sim}{e}_2\}$ is linearly dependent which is ruled out by Theorem A2.2.4.

Next, define $\underset{\sim}{a}_3 = \underset{\sim}{e}_3 + \beta \underset{\sim}{a}_1 + \gamma \underset{\sim}{a}_2$.

$\underset{\sim}{a}_3$ will be orthogonal to $\underset{\sim}{a}_1$ and $\underset{\sim}{a}_2$ if

$$\underset{\sim}{a}_1 \cdot \underset{\sim}{a}_3 = \underset{\sim}{a}_1 \cdot \underset{\sim}{e}_3 + \beta\, \underset{\sim}{a}_1 \cdot \underset{\sim}{a}_1 = 0 ,$$

and

$$\underset{\sim}{a}_2 \cdot \underset{\sim}{a}_3 = \underset{\sim}{a}_2 \cdot \underset{\sim}{e}_3 + \gamma\, \underset{\sim}{a}_2 \cdot \underset{\sim}{a}_2 = 0 .$$

Since $\underset{\sim}{a}_1 \neq \underset{\sim}{0}$ and $\underset{\sim}{a}_2 \neq \underset{\sim}{0}$, we can solve these equations for β and γ. As above, the linear independence of $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3\} \Rightarrow \underset{\sim}{a}_3 \neq \underset{\sim}{0}$.

Continuing in this fashion,[1] we finally obtain a subset $\{\underset{\sim}{a}_1, \underset{\sim}{a}_2, \cdots, \underset{\sim}{a}_n\}$ of n mutually orthogonal nonzero elements of V_n. $\therefore$ the set

$$\left\{ \frac{\underset{\sim}{a}_1}{|\underset{\sim}{a}_1|}, \frac{\underset{\sim}{a}_2}{|\underset{\sim}{a}_2|}, \cdots, \frac{\underset{\sim}{a}_n}{|\underset{\sim}{a}_n|} \right\}$$

is orthonormal. By Theorem A2.3.14, this subset is linearly independent; and hence by Theorem A2.2.8, it is a basis for V_n. $\square$

<u>Exercise</u> A2.3.15. Consider the inner product space obtained by assigning $\mathbb{R}^n$ the standard inner product (See Theorem A2.3.1.). Show that the standard basis (See Exercise A2.2.6.) for $\mathbb{R}^n$ is orthonormal.

Generally, we have been denoting a typical basis of an n-dimensional linear space by $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$. Because

---

[1] Of course, a formal mathematical induction argument would be appropriate here. Work this out as <u>Exercise</u> A2.3.14.

of the peculiarities of orthonormal bases, the following variant of this notation will be helpful.

Definition A2.3.9. Given an n-dimensional inner product space, say $\mathcal{V}_n$,[1] $\{\underset{\sim}{e}_{<1>}, \underset{\sim}{e}_{<2>}, \cdots, \underset{\sim}{e}_{<n>}\}$ denotes an orthonormal basis for $\mathcal{V}_n$, and the corresponding components[2] of $\underset{\sim}{u} \in \mathcal{V}_n$ are denoted by $u^{<1>}, u^{<2>}, \cdots, u^{<n>}$; i.e.,

$$\underset{\sim}{u} = u^{<1>}\underset{\sim}{e}_{<1>} + u^{<2>}\underset{\sim}{e}_{<2>} + \cdots + u^{<n>}\underset{\sim}{e}_{<n>}.$$

The computational simplifications afforded by the utilization of orthonormal bases are a result of

Theorem A2.3.16. Let $\{\underset{\sim}{e}_{<1>}, \underset{\sim}{e}_{<2>}, \cdots, \underset{\sim}{e}_{<n>}\}$ be an orthonormal basis for an n-dimensional inner product space $\mathcal{V}_n$. For $\underset{\sim}{u} \in \mathcal{V}_n$, write

$$\underset{\sim}{u} = u^{<1>}\underset{\sim}{e}_{<1>} + u^{<2>}\underset{\sim}{e}_{<2>} + \cdots + u^{<n>}\underset{\sim}{e}_{<n>}.$$

Then

$$u^{<i>} = \underset{\sim}{e}_{<i>} \cdot \underset{\sim}{u} \quad , \quad i = 1, 2, \cdots, n$$

and

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = u^{<1>}v^{<1>} + u^{<2>}v^{<2>} + \cdots + u^{<n>}v^{<n>}.$$

Proof. Exercise A2.3.16. □

---

[1] Cf. Theorem A2.3.2.

[2] See Definition A2.2.4.

That the equation $u^{\langle i \rangle} = \underset{\sim}{e}_{\langle i \rangle} \cdot \underset{\sim}{u}$ matches superscripts and subscripts is a notational defect characteristic of the use of an orthonormal basis.

Note that the inner product when written in terms of components relative to an orthonormal basis has the same special form as the standard inner product for $\mathbb{R}^n$.[1] In view of the isomorphism between $\mathcal{V}_n$ and $\mathbb{R}^n$ given in the proof of Theorem A2.2.13 and Exercise A2.3.15, this is not surprising. If the inner product in terms of components has some other form, then the associated basis is not orthonormal.

---

[1] Cf. Theorem A2.3.1.

Supplementary Reading

GEL'FAND, *Lectures on Linear Algebra*

GREUB, *Linear Algebra*

HALMOS, *Finite-Dimensional Vector Spaces*

LICHNEROWICZ, *Linear Algebra and Analysis*

MARTIN and MIZEL, *Introduction to Linear Algebra*

MICHEL and HERGET, *Mathematical Foundations in Engineering and Science*

NAYLOR and SELL, *Linear Operator Theory in Engineering and Science*

NICKERSON, SPENCER, and STEENROD, *Advanced Calculus*

NOLL, *Finite-Dimensional Spaces*

NOMIZU, *Fundamentals of Linear Algebra*

ODEN, *Applied Functional Analysis*

A2.4. Linear Transformations

Functions on linear spaces into linear spaces which have the additional property that they preserve sums and scalar multiples occur so frequently that it is worthwhile to gather some of their general properties in one place. The present brief section is in no sense a complete treatment of this vast topic. Some special cases of linear transformations will be developed more fully in later sections.

Definition A2.4.1. Let $\mathcal{U}$ and $\mathcal{V}$ be real linear spaces. A function $\underset{\sim}{f}: \mathcal{U} \to \mathcal{V}$ is a *linear transformation*[1] if it preserves sums and scalar multiples; i.e., if it has the following properties:

(L1) *Additivity*. $\forall \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{U}$

$$\underset{\sim}{f}(\underset{\sim}{u}+\underset{\sim}{v}) = \underset{\sim}{f}(\underset{\sim}{u}) + \underset{\sim}{f}(\underset{\sim}{v}) \ ;$$

(L2) *Homogeneity*. $\forall \alpha \in \mathbb{R}$ and $\forall \underset{\sim}{u} \in \mathcal{U}$

$$\underset{\sim}{f}(\alpha\underset{\sim}{u}) = \alpha \underset{\sim}{f}(\underset{\sim}{u}) .$$

It is important to realize that in (L1) the addition operations on the left- and right-hand sides are those of $\mathcal{U}$ and $\mathcal{V}$, respectively; similarly, for the scalar multiplication operations in (L2).

[1] Often, *linear operator*, especially when $\mathcal{U} = \mathcal{V}$.

Actually, we have already seen some examples of linear transformations. The general isomorphism of Definition A2.2.5 and the particular isomorphism from $\mathcal{V}_n$ to $\mathbb{R}^n$ developed in the proof of Theorem A2.2.13 are obviously linear transformations. For a slightly more subtle example, let $\underset{\sim}{w}$ be a fixed element of an inner product space $\mathcal{V}$ and define $f: \mathcal{V} \to \mathbb{R}$ by

$$f(\underset{\sim}{u}) = \underset{\sim}{u} \cdot \underset{\sim}{w}.$$

Then

$$f(\underset{\sim}{u} + \underset{\sim}{v}) = (\underset{\sim}{u} + \underset{\sim}{v}) \cdot \underset{\sim}{w} = \underset{\sim}{u} \cdot \underset{\sim}{w} + \underset{\sim}{v} \cdot \underset{\sim}{w} = f(\underset{\sim}{u}) + f(\underset{\sim}{v})$$

and

$$f(\alpha \underset{\sim}{u}) = (\alpha \underset{\sim}{u}) \cdot \underset{\sim}{w} = \alpha(\underset{\sim}{u} \cdot \underset{\sim}{w}) = \alpha f(\underset{\sim}{u}),$$

so this f is a linear transformation on $\mathcal{V}$ to $\mathbb{R}$.)[1] Since the inner product is a function on $\mathcal{V} \times \mathcal{V}$ which is linear in either variable when the other is held fixed, it is often called a bilinear transformation.

$\boxed{\text{INSERT material on p. A2.4.2a}}$

A key property of a linear transformation is that it maps the zero element of the domain into the zero element of the codomain.

Theorem A2.4.1. Let $\underset{\sim}{f}$ be a linear transformation on a linear space $\mathcal{U}$ to a linear space $\mathcal{V}$. Then

---

[1] In this context, $\mathbb{R}$ is viewed as a linear space (cf. Exercise A2.1.2).

INSERT for p A2.4.2

Differentiation and integration provide more sophisticated examples. E.g., the mapping $f \mapsto f'$ is a linear transformation on $C^1([a,b])$ to $C^0([a,b])$. In the other direction, the mapping $g \mapsto h$ where h is defined by

$$h(x) = \int_a^x g(\xi)\,d\xi$$

is a linear transformation on $C^0([a,b])$ to $C^1([a,b])$.
Here,

$$C^1([a,b]) = \{f : f \text{ is a continuously differentiable func. on } [a,b] \text{ to } \mathbb{R}\}.$$

return to p. A2.4.2

$$\underset{\sim}{f}(\underset{\sim}{0}) = \underset{\sim}{0}\,. \qquad)^1$$

Proof. Exercise A2.4.1. (Hint: Make intelligent choices of the variables in either the additivity or the homogeneity property.) □

It is not difficult to see that the properties of additivity and homogeneity can be combined together into a single statement.

Theorem A2.4.2. Let $\underset{\sim}{f}$ be a function on a linear space $\mathcal{U}$ to a linear space $\mathcal{V}$. Then $\underset{\sim}{f}$ is a linear transformation iff

$$\underset{\sim}{f}(\alpha\underset{\sim}{u} + \beta\underset{\sim}{v}) = \alpha\underset{\sim}{f}(\underset{\sim}{u}) + \beta\underset{\sim}{f}(\underset{\sim}{v}) \quad \forall \alpha, \beta \in \mathbb{R} \text{ and } \forall \underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{U}.$$

Proof. Exercise A2.4.2. □

We need some notation to save on writing.

Definition A2.4.2. Let $\mathcal{U}$ and $\mathcal{V}$ be two real linear spaces. Then $Lin(\mathcal{U},\mathcal{V})$ $)^2$ is the set of all linear transformations on $\mathcal{U}$ to $\mathcal{V}$.

---

[1] Of course, the $\underset{\sim}{0}$ on the l.h.s. is the zero element of $\mathcal{U}$, while the $\underset{\sim}{0}$ on the r.h.s. is the zero element of $\mathcal{V}$.

[2] Sometimes, $L(\mathcal{U},\mathcal{V})$, $Hom(\mathcal{U},\mathcal{V})$, $H(\mathcal{U},\mathcal{V})$.

Definition A2.4.3. The sum of two linear transformations $\underset{\sim}{f}_1, \underset{\sim}{f}_2 \in Lin(\mathcal{U}, \mathcal{V})$ is the function

$$(\underset{\sim}{f}_1 + \underset{\sim}{f}_2) : \mathcal{U} \to \mathcal{V}$$

defined by

$$(\underset{\sim}{f}_1 + \underset{\sim}{f}_2)(\underset{\sim}{u}) = \underset{\sim}{f}_1(\underset{\sim}{u}) + \underset{\sim}{f}_2(\underset{\sim}{u}) \quad , \underset{\sim}{u} \in \mathcal{U}.$$

The scalar multiple of $\underset{\sim}{f} \in Lin(\mathcal{U}, \mathcal{V})$ by $\alpha \in \mathbb{R}$ is the function

$$(\alpha \underset{\sim}{f}) : \mathcal{U} \to \mathcal{V}$$

defined by

$$(\alpha \underset{\sim}{f})(\underset{\sim}{u}) = \alpha \underset{\sim}{f}(\underset{\sim}{u}), \quad \underset{\sim}{u} \in \mathcal{U}.$$

An easy but important result is that sums and scalar multiples of linear transformations are themselves linear transformations.

Theorem A2.4.3. Let $\underset{\sim}{f}_1, \underset{\sim}{f}_2 \in Lin(\mathcal{U}, \mathcal{V})$. Then

$$(\underset{\sim}{f}_1 + \underset{\sim}{f}_2) \in Lin(\mathcal{U}, \mathcal{V}).$$

Let $\underset{\sim}{f} \in Lin(\mathcal{U}, \mathcal{V})$ and $\alpha \in \mathbb{R}$. Then

$$(\alpha \underset{\sim}{f}) \in Lin(\mathcal{U}, \mathcal{V}).$$

Proof. Exercise A2.4.3. $\square$

Of course, two linear transformations are viewed as equal if they are the same function. Formally,)[1]

Definition A2.4.4. Two linear transformations $\underset{\sim}{f}_1, \underset{\sim}{f}_2 \in Lin(\mathcal{U},\mathcal{V})$ are *equal*, and we write $\underset{\sim}{f}_1 = \underset{\sim}{f}_2$, if

$$\underset{\sim}{f}_1(\underset{\sim}{u}) = \underset{\sim}{f}_2(\underset{\sim}{u}) \quad \forall\ \underset{\sim}{u} \in \mathcal{U}.$$

You probably sense that we are on the way to seeing that $Lin(\mathcal{U},\mathcal{V})$ is a linear space. We still need a zero element.

Theorem A2.4.4. Let $\mathcal{U}$ and $\mathcal{V}$ be linear spaces. Define the function $\underset{\sim}{\theta}: \mathcal{U} \to \mathcal{V}$ by

$$\underset{\sim}{\theta}(\underset{\sim}{u}) = \underset{\sim}{0} \quad , \quad \underset{\sim}{u} \in \mathcal{U}. \quad)^2$$

Then $\underset{\sim}{\theta} \in Lin(\mathcal{U},\mathcal{V})$. $\underset{\sim}{\theta}$ is called the *zero linear transformation*.[3]

Proof. *Exercise* A2.4.4. □

---

[2] Of course, the $\underset{\sim}{0}$ here is the zero element of $\mathcal{V}$.

[3] It follows from the definition of equality that there is but one zero linear transformation. Hence, it does make sense to speak of the zero linear transformation. We cannot invoke Theorem A2.1.2 here because we do not yet know that $Lin(\mathcal{U},\mathcal{V})$ is a linear space.

[1] Cf. Exercise A1.2.6.

Theorem A.2.4.5. With the above definitions for addition, scalar multiplication, equality, and zero, $\operatorname{Lin}(\mathcal{U},\mathcal{V})$ is a linear space.

Proof. Exercise A2.4.5. □

In view of this result, all of the general properties of linear spaces, such as those given in Theorems A2.1.2 – 10, apply to $\operatorname{Lin}(\mathcal{U},\mathcal{V})$. E.g., in the present context, Theorem A2.1.9 says

if $\underset{\sim}{f}, \underset{\sim}{g} \in \operatorname{Lin}(\mathcal{U},\mathcal{V})$ and $\alpha, \beta \in \mathbb{R}$, then
$$\alpha=\beta \text{ and } \underset{\sim}{f}=\underset{\sim}{g} \Rightarrow \alpha\underset{\sim}{f}=\beta\underset{\sim}{g}.$$

We shall regularly use such facts without additional comment.

Supplementary Reading

BOWEN and WANG, Introduction to Vectors and Tensors, Vol. 1, Linear and Multilinear Algebra

GEL'FAND, Lectures on Linear Algebra

HALMOS, Finite-Dimensional Vector Spaces

MARTIN and MIZEL, Introduction to Linear Algebra

MICHEL and HERGET, Mathematical Foundations in Engineering and Science

NAYLOR and SELL, Linear Operator Theory in Engineering and Science

NOLL, Finite-Dimensional Spaces

NOMIZU, Fundamentals of Linear Algebra

ODEN, Applied Functional Analysis

A2.5. Linear Functionals. Dual Bases. Covariant and Contravariant Components

Linear transformations on a linear space to the reals are of special importance.

Definition A2.5.1. Let V be a linear space. Then the linear space $Lin(V, \mathbb{R})$[1] is called the algebraic dual space[2] of V and is denoted by V^*. The elements of V^* are called linear functionals[3].

A very useful tool for us will be

Theorem A2.5.1. (Representation Theorem for Linear Functionals[4]) Let $f \in V_n^* = Lin(V_n, \mathbb{R})$ where V_n is an n-dimensional inner product space[5]. Then $\exists$ a unique element $\underset{\sim}{f} \in V_n \ni$

$$f(\underset{\sim}{u}) = \underset{\sim}{f} \cdot \underset{\sim}{u} \quad \forall\, \underset{\sim}{u} \in V_n .$$

---

[1] See Definition A2.4.2 and Theorem A2.4.5. Here, of course, we are viewing $\mathbb{R}$ as a linear space (See Exercise A2.1.2.).

[2] Often, simply, dual space. The topological dual space is defined the same way except that the linear functionals are required to be bounded (see Definition A3.4.1); this is automatic when V is finite-dimensional

[3] Sometimes, linear forms.

(cont.)

Proof. By Theorem A2.3.14, V_n has an orthonormal basis, say $\{\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle n\rangle}\}$. Using the notation introduced in Definition A2.3.9, for any $\underset{\sim}{u} \in V_n$, we have

$$f(\underset{\sim}{u}) = f(u^{\langle 1\rangle}\underset{\sim}{e}_{\langle 1\rangle} + u^{\langle 2\rangle}\underset{\sim}{e}_{\langle 2\rangle} + \cdots + u^{\langle n\rangle}\underset{\sim}{e}_{\langle n\rangle})$$

$$= u^{\langle 1\rangle} f(\underset{\sim}{e}_{\langle 1\rangle}) + u^{\langle 2\rangle} f(\underset{\sim}{e}_{\langle 2\rangle}) + \cdots + u^{\langle n\rangle} f(\underset{\sim}{e}_{\langle n\rangle}) \quad \text{(linearity)}$$

$$= \underset{\sim}{f} \cdot \underset{\sim}{u} \quad \text{(Theorem A2.3.15)},$$

where

$$\underset{\sim}{f} := f(\underset{\sim}{e}_{\langle 1\rangle})\underset{\sim}{e}_{\langle 1\rangle} + f(\underset{\sim}{e}_{\langle 2\rangle})\underset{\sim}{e}_{\langle 2\rangle} + \cdots + f(\underset{\sim}{e}_{\langle n\rangle})\underset{\sim}{e}_{\langle n\rangle}.$$

Here, we have taken advantage of the fact that no matter what the inner product is, it has the standard form in terms of components relative to an orthonormal basis.

Now suppose that $\underset{\sim}{f}'$ is another element of V_n with this property. Then

$$f(\underset{\sim}{u}) = \underset{\sim}{f} \cdot \underset{\sim}{u} = \underset{\sim}{f}' \cdot \underset{\sim}{u}$$

$$\Longrightarrow$$

$$(\underset{\sim}{f}' - \underset{\sim}{f}) \cdot \underset{\sim}{u} = 0.$$

(cont.) [4] The infinite-dimensional counterpart is the famous Riesz representation theorem for a bounded linear functional on a Hilbert space.

[5] Cf. Theorem A2.3.2.

Since this must hold $\forall\, \underset{\sim}{u} \in \mathcal{V}_n$, Theorem A2.3.4 $\Rightarrow \underset{\sim}{f}' = \underset{\sim}{f}$; i.e., the special element of $\mathcal{V}_n$ in the representation is unique. □

The dual space turns out to have the same dimension as the original space.

Theorem A2.5.2. (algebraic) Let $\mathcal{V}_n$ be an n-dimensional linear space. Then the dual space $\mathcal{V}_n^* := Lin(\mathcal{V}_n, \mathbb{R})$ is also n-dimensional. In fact, if $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ is a basis for $\mathcal{V}_n$, so that $\underset{\sim}{u} \in \mathcal{V}_n$ has the unique representation

$$\underset{\sim}{u} = u^1 \underset{\sim}{e}_1 + u^2 \underset{\sim}{e}_2 + \cdots + u^n \underset{\sim}{e}_n, \quad)^1$$

then the subset of linear functionals $\{f^1, f^2, \cdots, f^n\} \subset \mathcal{V}_n^*$ defined by

$$f^i(\underset{\sim}{u}) = u^i, \quad \underset{\sim}{u} \in \mathcal{V}_n, \quad i = 1, 2, \cdots, n$$

is a basis for $\mathcal{V}_n^*$.

Proof. By Definition A2.2.3, we merely need to show that $\{f^1, f^2, \cdots, f^n\}$ is a basis for $\mathcal{V}_n^*$. That the function $f^i : \mathcal{V}_n \to \mathbb{R}$ defined by

$$f^i(\underset{\sim}{u}) = u^i, \quad \underset{\sim}{u} \in \mathcal{V}_n$$

---

[1] Cf. Definition A2.2.4 and Theorem A2.2.11.

really is a linear functional is left as Exercise A2.5.1. Let f be an arbitrary element of $\mathcal{V}_n^*$. Then for any $\underset{\sim}{u} \in \mathcal{V}_n$

$$f(\underset{\sim}{u}) = f(u^1\underset{\sim}{e}_1 + u^2\underset{\sim}{e}_2 + \cdots + u^n\underset{\sim}{e}_n) \quad \text{(substitution)}$$

$$= u^1 f(\underset{\sim}{e}_1) + u^2 f(\underset{\sim}{e}_2) + \cdots + u^n f(\underset{\sim}{e}_n) \quad \text{(linearity)}$$

$$= f(\underset{\sim}{e}_1) f^1(\underset{\sim}{u}) + f(\underset{\sim}{e}_2) f^2(\underset{\sim}{u}) + \cdots + f(\underset{\sim}{e}_n) f^n(\underset{\sim}{u}) \quad \text{(substitution).}$$

Thus, the subset $\{f^1, f^2, \ldots, f^n\}$ spans $\mathcal{V}^*$.[1] To establish linear independence, we use Theorem A2.2.2. Accordingly, consider

$$\alpha_1 f^1(\underset{\sim}{u}) + \alpha_2 f^2(\underset{\sim}{u}) + \cdots + \alpha_n f^n(\underset{\sim}{u}) = 0.$$

When $\underset{\sim}{u} = \underset{\sim}{e}_k$, this becomes

$$\alpha_1 f^1(\underset{\sim}{e}_k) + \alpha_2 f^2(\underset{\sim}{e}_k) + \cdots + \alpha_k f^n(\underset{\sim}{e}_k) = 0.$$

But by the definition of f^i,

$$f^i(\underset{\sim}{e}_k) = (\underset{\sim}{e}_k)^i = \delta^i_k,$$ [2]

so that

$$\alpha_1 \delta^1_k + \alpha_2 \delta^2_k + \cdots + \alpha_n \delta^n_k = \alpha_k = 0.$$

$\therefore$ $\{f^1, f^2, \ldots, f^n\}$ is linearly independent; and by Definition A2.2.2 it is a basis for $\mathcal{V}_n^*$. □

[1] See Exercise A1.2.3 for remarks on the addition and scalar multiplication of functions.

[2] As in Definition 2.3.8, $\delta^i_k := \begin{cases} 1 \text{ if } i = k \\ 0 \text{ if } i \neq k \end{cases}$.

We had no need of an inner product in Theorem A2.5.2; but since $\mathcal{V}_n$ is finite-dimensional, we can always equip it with an inner product. Consider this done. Now, the elements of the basis $\{f^1, f^2, \ldots, f^n\}$ for the dual space $\mathcal{V}_n^*$ are linear functionals. $\therefore$ by the representation theorem[1] $\exists$ a unique $\underset{\sim}{e}^i \in \mathcal{V}_n \ni$

$$f^i(\underset{\sim}{u}) = \underset{\sim}{e}^i \cdot \underset{\sim}{u} \quad \forall\, \underset{\sim}{u} \in \mathcal{V}_n .$$

On remembering that f^i is defined by

$$f^i(\underset{\sim}{u}) = u^i ,$$

we are lead to

Theorem A2.5.3. Let $\mathcal{V}_n$ be an n-dimensional inner product space. Given a basis $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \ldots, \underset{\sim}{e}_n\}$ for $\mathcal{V}_n$, so that $\underset{\sim}{u} \in \mathcal{V}_n$ has the unique representation

$$\underset{\sim}{u} = u^1 \underset{\sim}{e}_1 + u^2 \underset{\sim}{e}_2 + \cdots + u^n \underset{\sim}{e}_n ,$$

$\exists$ a unique subset $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \ldots, \underset{\sim}{e}^n\} \subset \mathcal{V}_n \ni$

$$\underset{\sim}{e}^i \cdot \underset{\sim}{u} = u^i \quad \forall\, \underset{\sim}{u} \in \mathcal{V}_n .$$

The subset $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \ldots, \underset{\sim}{e}^n\}$ is a basis for $\mathcal{V}_n$. It is called

---

[1] Theorem A2.5.1.

<u>the dual basis</u>[1] <u>of the original basis</u> $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$.

Proof. All that remains to be established is that $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ is a basis. In view of Theorem A2.2.8, it suffices to show that $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ is linearly independent. As usual, we use Theorem A2.2.2. Accordingly, consider

$$\alpha_1 \underset{\sim}{e}^1 + \alpha_2 \underset{\sim}{e}^2 + \cdots + \alpha_n \underset{\sim}{e}^n = \underset{\sim}{0}.$$

Take the inner product of this equation with any $\underset{\sim}{u} \in \mathcal{V}_n$ and use the linearity of the inner product to get

$$\alpha_1 \underset{\sim}{e}^1 \cdot \underset{\sim}{u} + \alpha_2 \underset{\sim}{e}^2 \cdot \underset{\sim}{u} + \cdots + \alpha_n \underset{\sim}{e}^n \cdot \underset{\sim}{u} = \underset{\sim}{0} \cdot \underset{\sim}{u} = 0.$$

But $\underset{\sim}{e}^i \cdot \underset{\sim}{u} = u^i$, so this becomes

$$(*) \quad \alpha_1 u^1 + \alpha_2 u^2 + \cdots + \alpha_n u^n = 0.$$

Choose $\underset{\sim}{u} = \underset{\sim}{e}_k$, so that $u^i = (\underset{\sim}{e}_k)^i = \delta^i_k$. Then $(*) \Rightarrow$

$$\alpha_1 \delta^1_k + \alpha_2 \delta^2_k + \cdots + \alpha_n \delta^n_k = \alpha_k = 0,$$

and $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ is linearly independent. □

Let us review what we have done here. We started with a basis $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ for an n-dimensional inner product space $\mathcal{V}_n$, and used it to construct a basis $\{f^1, f^2, \cdots, f^n\}$ for the

---

[1] Sometimes, reciprocal basis.

(algebraic) dual space V_n^*. Then the representation theorem for linear functionals was applied to each of the f^i:

$$f^i(\underset{\sim}{u}) = \underset{\sim}{e}^i \cdot \underset{\sim}{u} \; , \quad \underset{\sim}{u} \in V_n .$$

The subset $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\} \subset V_n$ thus obtained is the dual basis of the primal basis $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$. Note that the dual basis is a basis for V_n — not for the dual space V_n^*.

The following theorem explains the terminology "reciprocal basis" which is sometimes used in place of dual basis.

Theorem A2.5.4. <u>Let</u> $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ <u>and</u> $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ <u>be a basis and its dual for an n-dimensional inner product space. Then</u>

$$\underset{\sim}{e}^i \cdot \underset{\sim}{e}_j = \delta^i_j \; , \quad i, j = 1, 2, \cdots, n .$$

Proof. By Theorem A2.5.3, $\forall\, \underset{\sim}{u}$ in the underlying space

$$\underset{\sim}{e}^i \cdot \underset{\sim}{u} = u^i .$$

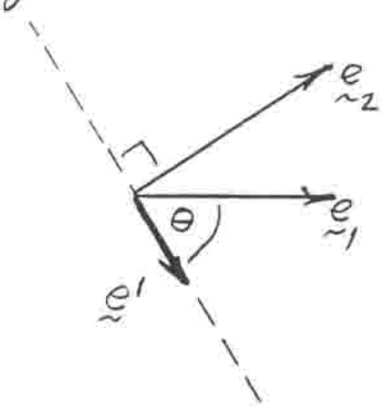
Choose $\underset{\sim}{u} = \underset{\sim}{e}_j$ to get

$$\underset{\sim}{e}^i \cdot \underset{\sim}{e}_j = (\underset{\sim}{e}_j)^i = \delta^i_j \; ,$$

in which the last step follows from

$$\underset{\sim}{e}_j = \delta^1_j \underset{\sim}{e}_1 + \delta^2_j \underset{\sim}{e}_2 + \cdots + \delta^n_j \underset{\sim}{e}_n \ . \quad \square$$

We can shed some light on $\underset{\sim}{e}^i \cdot \underset{\sim}{e}_j = \delta^i_j$ by considering the linear space of <u>vectors</u> associated with the <u>Euclidean plane</u>. Suppose that $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2\}$ is a given basis for this space. We wish to find the dual basis $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2\}$. Since $\underset{\sim}{e}^1 \cdot \underset{\sim}{e}_2 = \delta^1_2 = 0$, $\underset{\sim}{e}^1$ is orthogonal to $\underset{\sim}{e}_2$; thus, $\underset{\sim}{e}^1$ lies along the dotted line.



We also have $\underset{\sim}{e}^1 \cdot \underset{\sim}{e}_1 = \delta^1_1 = 1$. But $\underset{\sim}{e}^1 \cdot \underset{\sim}{e}_1 = |\underset{\sim}{e}^1||\underset{\sim}{e}_1|\cos\theta$, where θ is the angle between $\underset{\sim}{e}^1$ and $\underset{\sim}{e}_1$. $\therefore$ $\cos\theta > 0$, so θ must be the acute angle indicated in the figure; i.e., $\underset{\sim}{e}^1$ must point as shown. Now that θ is known, the length of $\underset{\sim}{e}^1$ is determined by $1 = |\underset{\sim}{e}^1||\underset{\sim}{e}_1|\cos\theta$. The vector $\underset{\sim}{e}^2$ can be obtained in the same manner.

Since no ambiguities showed up in the above construction, it suggests that the conditions $\underset{\sim}{e}^i \cdot \underset{\sim}{e}_j = \delta^i_j$ uniquely characterize the dual basis. This is, indeed, the case; formally, we have

Theorem A2.5.5. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ be a basis for an n-dimensional inner product space $\mathcal{V}_n$. If the subset $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\} \subset \mathcal{V}_n$ has the property that

$$\underset{\sim}{e}^i \cdot \underset{\sim}{e}_j = \delta^i_j \,, \quad i,j = 1,2,\cdots,n \,,$$

then $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ is the dual basis of $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$.

Proof. Let $\underset{\sim}{u} \in \mathcal{V}_n$ so that

$$\underset{\sim}{u} = u^1 \underset{\sim}{e}_1 + u^2 \underset{\sim}{e}_2 + \cdots + u^n \underset{\sim}{e}_n \,.$$

Take the inner product of this equation with $\underset{\sim}{e}^i$ and use the linearity of the inner product to get

$$\begin{aligned} \underset{\sim}{e}^i \cdot \underset{\sim}{u} &= u^1 \underset{\sim}{e}^i \cdot \underset{\sim}{e}_1 + u^2 \underset{\sim}{e}^i \cdot \underset{\sim}{e}_2 + \cdots + u^n \underset{\sim}{e}^i \cdot \underset{\sim}{e}_n \\ &= u^i \quad (\underset{\sim}{e}^i \cdot \underset{\sim}{e}_j = \delta^i_j) \,. \end{aligned}$$

Then by Theorem A2.5.3, $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ is the dual basis of $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$. $\square$

An immediate corollary is

Theorem A2.5.6. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ and $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ be a basis and its dual for an n-dimensional inner product space. Then the dual basis of the dual basis $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ is the original basis $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$.

A slightly more difficult corollary of Theorem A2.5.5 is

Theorem A2.5.7. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ and $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ be a basis and its dual for an n-dimensional inner product space. Then $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ is orthonormal[1] iff

$$\underset{\sim}{e}^i = \underset{\sim}{e}_i \quad , \quad i = 1, 2, \cdots, n.$$

Proof. Exercise A2.5.2. □

Now we are in position for more standard terminology; however, it would be inaccurate to describe our approach to this topic as standard.

Definition A2.5.2. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ and $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ be a basis and its dual for an n-dimensional inner product space $\mathcal{V}_n$, so that $\underset{\sim}{u} \in \mathcal{V}_n$ has the unique representations

$$\underset{\sim}{u} = u^1 \underset{\sim}{e}_1 + u^2 \underset{\sim}{e}_2 + \cdots + u^n \underset{\sim}{e}_n$$

and

$$\underset{\sim}{u} = u_1 \underset{\sim}{e}^1 + u_2 \underset{\sim}{e}^2 + \cdots + u_n \underset{\sim}{e}^n.$$

The scalars u^i (the components of $\underset{\sim}{u}$ w.r.t. to the original basis $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$) are the contravariant components of $\underset{\sim}{u}$, while the scalars u_i (the components of $\underset{\sim}{u}$ w.r.t. the dual basis $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$) are the covariant components of $\underset{\sim}{u}$.

---

[1] See Definition A2.3.10.

The mnemonic "co goes below" is useful in keeping this terminology straight. There is nothing special about the original basis other than that it is the one with which we happen to start.

Theorem A2.5.8. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ and $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ be a basis and its dual for an n-dimensional inner product space $\mathcal{V}_n$. Then for $\underset{\sim}{u} \in \mathcal{V}_n$

$$u^i = \underset{\sim}{e}^i \cdot \underset{\sim}{u}, \quad i = 1, 2, \cdots, n,$$

and

$$u_i = \underset{\sim}{e}_i \cdot \underset{\sim}{u}, \quad i = 1, 2, \cdots, n.$$

Proof. The statement $u^i = \underset{\sim}{e}^i \cdot \underset{\sim}{u}$ is the defining property of $\underset{\sim}{e}^i$. To establish the second statement, consider

$$\underset{\sim}{u} = u_1 \underset{\sim}{e}^1 + u_2 \underset{\sim}{e}^2 + \cdots + u_n \underset{\sim}{e}^n.$$

Take the inner product of this equation with $\underset{\sim}{e}_i$ to get

$$\underset{\sim}{e}_i \cdot \underset{\sim}{u} = u_1 \underset{\sim}{e}_i \cdot \underset{\sim}{e}^1 + u_2 \underset{\sim}{e}_i \cdot \underset{\sim}{e}^2 + \cdots + u_n \underset{\sim}{e}_i \cdot \underset{\sim}{e}^n \quad \text{(linearity of the inner product)}$$

$$= u_i \quad (\underset{\sim}{e}_i \cdot \underset{\sim}{e}^j = \delta_i^j). \quad \square$$

As an immediate consequence of the last two theorems, we have

Theorem A2.5.9. Let $\mathcal{V}_n$ be an n-dimensional inner product space. Then relative to an orthonormal basis for $\mathcal{V}_n$, the covariant and contravariant components of any element $\underset{\sim}{u} \in V_n$ coalesce; i.e.,

$$u_{\langle i \rangle} = u^{\langle i \rangle} \quad , \; i = 1, 2, \cdots, n.$$

That the equation above matches subscripts and superscripts is a notational effect that will manifest itself whenever an orthonormal basis is employed.

It will be convenient to adopt the

Summation Convention. If an unspecified[1] index appears exactly twice in the same monomial, once as a subscript and once as a superscript, then summation from 1 to n over this index is implied. n is the dimension of the underlying linear space.

For example,

$$u_i \underset{\sim}{e}^i = \sum_{i=1}^{n} u_i \underset{\sim}{e}^i = u_1 \underset{\sim}{e}^1 + u_2 \underset{\sim}{e}^2 + \cdots + u_n \underset{\sim}{e}^n .$$

The following notation will also be useful.

Definition A2.5.3, Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ and $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ be a basis and its dual for an n-dimensional inner product space V_n. Then the associated g-symbols are

[1] i, j, etc. as opposed to $1, 2$, etc.

$$g_{ij} = \underset{\sim}{e}_i \cdot \underset{\sim}{e}_j \,, \quad g^{ij} = \underset{\sim}{e}^i \cdot \underset{\sim}{e}^j \,, \quad g^i_j = \underset{\sim}{e}^i \cdot \underset{\sim}{e}_j \,, \quad i,j = 1,2,\cdots,n.$$

Some of the more important properties of the g-symbols are given by

Theorem A2.5.10. <u>Refer to the notation of Definitions A2.5.2 and 3. Then</u>

(i) $g_{ji} = g_{ij}$, $g^{ji} = g^{ij}$, $g^j_i = g^i_j$, $i,j = 1,2,\cdots,n$;

(ii) $g^i_j = \delta^i_j$, $i,j = 1,2,\cdots,n$;

(iii) <u>For any</u> $\underset{\sim}{u}, \underset{\sim}{v} \in \mathcal{V}_n$,

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = g_{ij} u^i v^j = g^{ij} u_i v_j = g^i_j u_i v^j = g^i_j u^j v_i = u_i v^i = u^i v_i \,;$$

(iv) <u>For any</u> $\underset{\sim}{u} \in \mathcal{V}_n$,

$$|\underset{\sim}{u}| = \sqrt{g_{ij} u^i u^j} = \sqrt{g^{ij} u_i u_j} = \sqrt{u_i u^i} \,;$$

(v) $g_{ik} g^{kj} = \delta^j_i$, $g^{ik} g_{kj} = \delta^i_j$, $i,j = 1,2,\cdots,n$;)[1]

(vi) <u>The matrices</u> $[g_{ij}]$ <u>and</u> $[g^{ij}]$ <u>are positive definite</u>;

(vii) $\det[g_{ij}] = \left(\det[g^{ij}]\right)^{-1}$.

---

[1] Thus, the matrix $[g^{ij}]$ is the inverse of the matrix $[g_{ij}]$. Here, $[g_{ij}]$ stands for the matrix whose ij-element is g_{ij}.

Proof. (i): $g_{ji} = \underset{\sim}{e}_j \cdot \underset{\sim}{e}_i$ (definition)

$= \underset{\sim}{e}_i \cdot \underset{\sim}{e}_j$ (commutativity of the inner product)

$= g_{ij}$ (definition).

Similarly, $g^{ji} = g^{ij}$.

$g^j_i = \underset{\sim}{e}^j \cdot \underset{\sim}{e}_i$ (definition)

$= \delta^j_i$ (Theorem A2.5.4)

$= \delta^i_j$ (property of Kronecker's delta)

$= \underset{\sim}{e}^i \cdot \underset{\sim}{e}_j$ (Theorem A2.5.4)

$= g^i_j$ (definition).

(ii): Contained in the proof of (i).

(iii): $\underset{\sim}{u} \cdot \underset{\sim}{v} = (u^i \underset{\sim}{e}_i) \cdot (v^j \underset{\sim}{e}_j)$ (component representation)

$= u^i v^j (\underset{\sim}{e}_i \cdot \underset{\sim}{e}_j)$ (linearity of inner product)

$= g_{ij} u^i v^j$ (definition).

Similarly, $\underset{\sim}{u} \cdot \underset{\sim}{v} = g^{ij} u_i v_j$, and

$$\underset{\sim}{u}\cdot\underset{\sim}{v} = g_i^j u^i v_j$$
$$= \delta_i^j u^i v_j \quad ((ii))$$
$$= u^j v_j \quad (\text{property of Kronecker's delta}).$$

Similarly, $\underset{\sim}{u}\cdot\underset{\sim}{v} = u_i v^i$.

(iv): $|\underset{\sim}{u}|^2 = \underset{\sim}{u}\cdot\underset{\sim}{u}$ (definition)

$$= g_{ij} u^i u^j = g^{ij} u_i u_j = u_i u^i \quad ((iii)).$$

(v): $g_{ik} g^{kj} = (\underset{\sim}{e}_i\cdot\underset{\sim}{e}_k)(\underset{\sim}{e}^k\cdot\underset{\sim}{e}^j)$ (definition)

$$= (\underset{\sim}{e}_i)_k (\underset{\sim}{e}^j)^k \quad (\text{Theorem A2.5.8})$$
$$= \underset{\sim}{e}_i\cdot\underset{\sim}{e}^j \quad ((iii))$$
$$= \delta_i^j \quad (\text{Theorem A2.5.4}).$$

Similarly, $g^{ik} g_{kj} = \delta_j^i$.

(vi): Consider $\underset{\sim}{u}\cdot\underset{\sim}{u} = g_{ij} u^i u^j$. By the positive definiteness of the inner product $\underset{\sim}{u}\cdot\underset{\sim}{u} \geq 0$ with $= 0$ iff $\underset{\sim}{u} = \underset{\sim}{0}$. By Theorem A2.2.12, $\underset{\sim}{u} = \underset{\sim}{0}$ iff $u^i = 0$, $i = 1, 2, \cdots, n$. Hence, the quadratic form $g_{ij} u^i u^j$ and the associated matrix $[g_{ij}]$ are positive definite by definition. Similarly, for $[g^{ij}]$.

(vii): $\delta^j_i = g_{ik} g^{kj}$ ((v))

$$\Rightarrow \det[\delta^j_i] = \det[g_{ik} g^{kj}]$$

$$= \det[g_{ik}] \det[g^{lm}]$$ (For square matrices A and B, $\det(AB) = (\det A)(\det B)$)

But $[\delta^j_i]$ is just the identity matrix

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 \cdots & 0 \\ & & \cdots & & \\ 0 & 0 & \cdots & & 0\ 1 \end{bmatrix},$$

so $\det[\delta^j_i] = 1$. $\therefore \det[g_{ik}] \det[g^{lm}] = 1$. Since the product on the l.h.s. equals 1, neither factor can vanish; and we can write

$$\det[g_{ik}] = (\det[g^{lm}])^{-1}. \quad \square$$

Our next result shows how the g-symbols can be used to convert between covariant and contravariant components.

Theorem A2.5.11. Refer to the notation of Definitions A2.5.2 and 3. Then for any $\underset{\sim}{u} \in \mathcal{V}_n$,

$$u^i = g^{ij} u_j \quad \text{and} \quad u_i = g_{ij} u^j, \quad i = 1, 2, \cdots, n.$$

Proof. $u^i = \underset{\sim}{e}^i \cdot \underset{\sim}{u}$ (Theorem A2.5.8)

$$= \underset{\sim}{e}^i \cdot (u_j \underset{\sim}{e}^j) \quad \text{(component representation)}$$

$$= u_j \, (\underset{\sim}{e}^i \cdot \underset{\sim}{e}^j) \quad \text{(linearity of the inner product)}$$

$$= g^{ij} u_j \quad (\text{definition of } g^{ij}).$$

Similarly, $u_i = g_{ij} u^j$. □

For obvious reasons the operations indicated in Theorem A2.5.11 are often referred to as "raising and lowering indices". The next result shows that we can do the same thing with the basis elements.

Theorem A2.5.12. Refer to the notation of Definition A2.5.3. Then

$$\underset{\sim}{e}^i = g^{ij} \underset{\sim}{e}_j \quad \text{and} \quad \underset{\sim}{e}_i = g_{ij} \underset{\sim}{e}^j, \quad i = 1, 2, \cdots, n.$$

Proof. *Exercise* A2.5.3. □

Supplementary Reading

Same as for § A2.3 plus

COLEMAN, MARKOVITZ, and NOLL, Viscometric Flows of Non-Newtonian Fluids (Appendix on Mathematical Concepts)

LICHNEROWICZ, Tensor Calculus

A2.6. Transformations of Bases

In this section, we investigate how the components of an element of a finite-dimensional linear space change when the basis is changed. We start by considering the relationships that must hold between any two bases.

Theorem A2.6.1. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$, $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ and $\{\bar{\underset{\sim}{e}}_1, \bar{\underset{\sim}{e}}_2, \cdots, \bar{\underset{\sim}{e}}_n\}$, $\{\bar{\underset{\sim}{e}}^1, \bar{\underset{\sim}{e}}^2, \cdots, \bar{\underset{\sim}{e}}^n\}$ be two bases and their duals[1] for an n-dimensional inner product space[2] $\mathcal{V}_{,}$. Define

$$\alpha^j_i = \bar{\underset{\sim}{e}}^j \cdot \underset{\sim}{e}_i \quad , \quad \beta^i_j = \bar{\underset{\sim}{e}}_j \cdot \underset{\sim}{e}^i \quad (i, j = 1, 2, \cdots, n).$$

Then

$$\underset{\sim}{e}_i = \alpha^j_i \, \bar{\underset{\sim}{e}}_j \quad , \quad \underset{\sim}{e}^i = \beta^i_j \, \bar{\underset{\sim}{e}}^j \quad (i = 1, 2, \cdots, n),$$

$$\bar{\underset{\sim}{e}}_i = \beta^j_i \, \underset{\sim}{e}_j \quad , \quad \bar{\underset{\sim}{e}}^i = \alpha^i_j \, \underset{\sim}{e}^j \quad (i = 1, 2, \cdots, n),$$

and

$$\alpha^\ell_i \beta^k_\ell = \delta^k_i \quad , \quad \alpha^i_\ell \beta^\ell_k = \delta^i_k \quad (i, k = 1, 2, \cdots, n).$$

Proof. By Definitions A2.2.2 and 3, $\underset{\sim}{e}_i$ can be written as a linear combination of the members of the basis $\{\bar{\underset{\sim}{e}}_1, \bar{\underset{\sim}{e}}_2, \cdots, \bar{\underset{\sim}{e}}_n\}$; i.e.,

$$\underset{\sim}{e}_i = \alpha^j_i \, \bar{\underset{\sim}{e}}_j .$$

---

[1] See Theorem A2.5.3.

[2] Cf. Theorem A2.3.2.

By Theorem A2.5.8, the components α_i^j are given by

$$\alpha_i^j = \underset{\sim}{e}_i \cdot \bar{\underset{\sim}{e}}^j$$

Similarly,

$$\underset{\sim}{e}^i = \beta_j^i \bar{\underset{\sim}{e}}^j \quad \text{with} \quad \beta_j^i = \underset{\sim}{e}^i \cdot \bar{\underset{\sim}{e}}_j .$$

With the above definitions for the α's and the β's,

$$\begin{aligned}\alpha_i^\ell \beta_\ell^k &= (\underset{\sim}{e}_i \cdot \bar{\underset{\sim}{e}}^\ell)(\underset{\sim}{e}^k \cdot \bar{e}_\ell) \\ &= (\underset{\sim}{e}_i)^{\bar{\ell}} (\underset{\sim}{e}^k)_{\bar{\ell}} \quad (\text{Theorem A2.5.8})^1 \\ &= \underset{\sim}{e}_i \cdot \underset{\sim}{e}^k \quad (\text{Theorem A2.5.10 (iii)}) \\ &= \delta_i^k \quad (\text{Theorem A2.5.4}).\end{aligned}$$

Similarly,

$$\alpha_\ell^i \beta_k^\ell = \delta_k^i .$$

Finally, we can use the above results to write

$$\beta_i^j \underset{\sim}{e}_j = \beta_i^j (\alpha_j^k \bar{\underset{\sim}{e}}_k) = \delta_i^k \bar{\underset{\sim}{e}}_k = \bar{\underset{\sim}{e}}_i .$$

---

[1] Here, $(\underset{\sim}{e}_i)^{\bar{\ell}}$ denotes the ℓ^{th} contravariant component of $\underset{\sim}{e}_i$ w.r.t. the barred basis. Marking the index (rather than the kernel letter as in Theorem A2.6.2) to indicate the underlying basis is occasionally advantageous (cf. LICHNEROWICZ's <u>Tensor Calculus</u>).

Similarly,

$$\alpha^i_j \underset{\sim}{e}^j = \bar{e}_i . \quad \square$$

As an important consequence of this theorem, we have

Theorem A2.6.2. Let the hypotheses of Theorem A2.6.1 hold. Let $\underset{\sim}{u} \in V_n$ so that

$$\underset{\sim}{u} = u^i \underset{\sim}{e}_i = u_i \underset{\sim}{e}^i = \bar{u}^i \bar{e}_i = \bar{u}_i \bar{\underset{\sim}{e}}^i .$$

Then, for $i = 1, 2, \cdots, n$,

$\bar{u}^i = \alpha^i_j u^j$ --- transformation rule for contravariant components,

$\bar{u}_i = \beta^j_i u_j$ --- transformation rule for covariant components,

and

$$\left.\begin{aligned} u^i &= \beta^i_j \bar{u}^j \\ u_i &= \alpha^j_i \bar{u}_j \end{aligned}\right\} \text{--- inverse transformation rules.}$$

Proof. We use the results of Theorem A2.6.1 together with the uniqueness of components[1] to write

$$\underset{\sim}{u} = u^i \underset{\sim}{e}_i = u^i \alpha^j_i \bar{e}_j = \bar{u}^j \bar{e}_j \Rightarrow \bar{u}^j = \alpha^j_i u^i ,$$

$$\underset{\sim}{u} = u_i \underset{\sim}{e}^i = u_i \beta^i_j \bar{e}^j = \bar{u}_j \bar{e}^j \Rightarrow \bar{u}_j = \beta^i_j u_i .$$

[1] Theorem A2.2.11.

Then

$$\beta^k_j \bar{u}^j = \beta^k_j (\alpha^j_i u^i) = \delta^k_i u^i = u^k ,$$

$$\alpha^j_k \bar{u}_j = \alpha^j_k (\beta^i_j u_i) = \delta^i_k u_i = u_k . \quad \square$$

Obviously, if we have a list of reals of length n, say $(u^1, u^2, \dots, u^n)$, we can generate an element, say $\underset{\sim}{u}$, of $\mathcal{V}_n$ by letting the u^i be the components of $\underset{\sim}{u}$ w.r.t. some basis, say $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \dots, \underset{\sim}{e}_n\}$:

$$\underset{\sim}{u} := u^i \underset{\sim}{e}_i .$$

A common occurence in the physical sciences is that by some natural scheme we have real lists associated with each basis, say

$$(u^1, u^2, \dots, u^n) \text{ with } \{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \dots, \underset{\sim}{e}_n\} ,$$

$$(\bar{u}^1, \bar{u}^2, \dots, \bar{u}^n) \text{ with } \{\underset{\sim}{\bar{e}}_1, \underset{\sim}{\bar{e}}_2, \dots, \underset{\sim}{\bar{e}}_n\} ,$$

etc.

This allows us to generate an infinity of elements of $\mathcal{V}_n$:

$$\underset{\sim}{u} = u^i \underset{\sim}{e}_i , \quad \underset{\sim}{\bar{u}} = \bar{u}^i \underset{\sim}{\bar{e}}_i , \quad \text{etc.}$$

By Theorem A2.6.3, if these elements of $\mathcal{V}_n$ are all the same (i.e., if $\underset{\sim}{u} = \underset{\sim}{\bar{u}}$, etc.), then the $u^i, \bar{u}^i$, etc. satisfy the transformation rules

$$\bar{u}^i = \alpha^i_j u^j \ , \quad u^i = \beta^i_j \bar{u}^j \ ,$$

etc.

This leads us to

Theorem A2.6.3. Let the hypotheses of Theorem A2.6.1 hold. Let there be real lists of length n associated with each basis for $\mathcal{V}_n$:

$$(u^1, u^2, \ldots, u^n) \ \underline{\text{with}} \ \{ \underset{\sim}{e}_1, \underset{\sim}{e}_2, \ldots, \underset{\sim}{e}_n \} \ ,$$

$$(\bar{u}^1, \bar{u}^2, \ldots, \bar{u}^n) \ \underline{\text{with}} \ \{ \underset{\sim}{\bar{e}}_1, \underset{\sim}{\bar{e}}_2, \ldots, \underset{\sim}{\bar{e}}_n \} \ ,$$

etc.

The entries in these lists are the contravariant[1] components of an element

$$\underset{\sim}{u} = u^i \underset{\sim}{e}_i = \bar{u}^i \underset{\sim}{\bar{e}}_i = \cdots$$

of $\mathcal{V}_n$ iff they satisfy the transformation rules

$$\bar{u}^i = \alpha^i_j u^j \ , \quad u^i = \beta^i_j \bar{u}^j \ ,$$

etc.

[1] Obviously, there is a strictly analogous result for covariant component.

Proof. Only the "if" portion remains to be proven. Accordingly, define

$$\underset{\sim}{u} = u^i \underset{\sim}{e}_i \ , \quad \underset{\sim}{\bar{u}} = \bar{u}^i \underset{\sim}{\bar{e}}_i \ , \quad \text{etc.},$$

where the components $u^i, \bar{u}^i$, etc. meet the stated transformation rules. Then

$$
\begin{aligned}
\underset{\sim}{\bar{u}} &= \bar{u}^i \underset{\sim}{\bar{e}}_i && (\text{construction}) \\
&= (\alpha^i_j u^j)\underset{\sim}{\bar{e}}_i && (\text{transformation rule}) \\
&= u^j (\alpha^i_j \underset{\sim}{\bar{e}}_i) && (\text{linear space axioms}) \\
&= u^j \underset{\sim}{e}_j && (\text{Theorem A2.6.1}) \\
&= \underset{\sim}{u} && (\text{construction}).
\end{aligned}
$$

Thus, if the transformation rules are satisfied, the natural constructions all generate the same element of $\mathcal{V}_n$. □

Usually in the physical sciences, elements of finite-dimensional linear spaces are defined through their components together with the requirement that their components transform according to the rules of Theorem A2.6.2 under a change of basis. We shall refer to such a scheme as the *component-transformation rule approach*.

The next theorem shows that the g-symbols obey similar transformation rules.

Theorem A2.6.4. *Let the hypotheses of Theorem A2.6.1 hold. Define*

$$g^{ij} = \underset{\sim}{e}^i \cdot \underset{\sim}{e}^j \,, \quad \bar{g}^{ij} = \bar{\underset{\sim}{e}}^i \cdot \bar{\underset{\sim}{e}}^j \,, \quad \textit{etc.}^{1} \quad (i,j = 1,2, \ldots, n).$$

Then, for $i,j = 1,2,\ldots,n$,

$$\bar{g}^{ij} = \alpha^i_k \alpha^j_l g^{kl} \,, \quad \bar{g}_{ij} = \beta^k_i \beta^l_j g_{kl}$$

$$\bar{g}^i_j = \alpha^i_k \beta^l_j g^k_l = \delta^i_j = g^i_j$$

and

$$g^{ij} = \beta^i_k \beta^j_l \bar{g}^{kl} \,, \quad g_{ij} = \alpha^k_i \alpha^l_j \bar{g}_{kl} \,,$$

---

[1] See Definition A2.5.3.

$$g^i_{\ j} = \beta^i_k \alpha^l_j \bar{g}^k_{\ l} = \delta^i_j = \bar{g}^i_{\ j} .$$

Proof. Exercise A2.6.1. □

Note that if both of the bases $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ and $\{\bar{\underset{\sim}{e}}_1, \bar{\underset{\sim}{e}}_2, \cdots, \bar{\underset{\sim}{e}}_n\}$ are orthonormal, then the above result yields orthogonality conditions such as

$$\delta^{ij} = \alpha^i_k \alpha^j_l \delta^{kl} = \sum_{k=1}^{n} \alpha^i_k \alpha^j_k$$

for the transformation coefficients, Such conditions are probably familiar to you from "Cartesian tensor algebra" (usually formulated via the component-transformation rule approach) where only orthonormal bases are employed.

Exercise A2.6.2. In the component-transformation rule approach, the inner product of two elements would be defined in terms of components as

$$\underset{\sim}{u} \cdot \underset{\sim}{v} := g^{ij} u_i v_j .$$

Of course, in this approach one needs to be concerned about the definition being tied fundamentally to the basis used in the definition. Use the transformation rules to show that

$$\bar{g}^{ij} \bar{u}_i \bar{v}_j = g^{ij} u_i v_j .$$

Of course, in our <u>basis-free</u> approach to the inner product such questions never arise; and

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = \bar{g}^{ij}\,\bar{u}_i\,\bar{v}_j\,.$$

follows directly from Theorem A2.5.10 (iii).

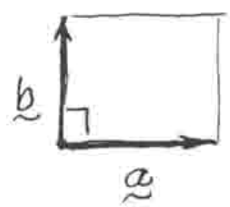
<u>Supplementary Reading</u>

Same as for § A2.5

A2.7. Volume Orientation Functionals. Oriented Spaces

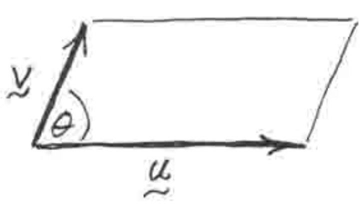
Here, we put in place the foundation for consideration of the "vector product" of two elements of a three-dimensional space in § A2.8 and the "determinant" of a "tensor" on an n-dimensional space in § A3,5. For motivation, we revert again to <u>vectors</u> in the <u>Euclidean plane</u>.

If $\underset{\sim}{a}$ and $\underset{\sim}{b}$ are two orthogonal vectors in the plane, they generate a rectangle.

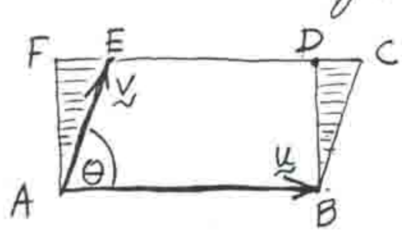


By definition, the area of a rectangle is the product of the base and the height; i.e., area $\{\underset{\sim}{a}, \underset{\sim}{b}\} := |\underset{\sim}{a}||\underset{\sim}{b}|$.

In the general case, two vectors $\underset{\sim}{u}$ and $\underset{\sim}{v}$ in the plane generate a parallelogram.



The area of a parallelogram, is still the product of the base and the height as is seen from the following figure.



The base times the height gives the area of the rectangle ABDF. This area includes the area of the triangle AFE, which is not part of the area of the parallelogram generated by $\underset{\sim}{u}$ and $\underset{\sim}{v}$; but it leaves out the area of the triangle BDC, which is included in the area of the parallelogram. However, the areas of these two triangles are equal. Thus, since the height is $|\underset{\sim}{v}|\sin\theta$,

$$\text{area}\,\{\underset{\sim}{u},\underset{\sim}{v}\} = |\underset{\sim}{u}||\underset{\sim}{v}|\sin\theta .$$

Since, by convention, the angle between two vectors is taken to be the smaller of the two possibilites, this general formula produces a nonnegative area. It also gives the correct result in the rectangular case.

The area of a parallelogram has several important properties. First, we note that if $\alpha > 0$ so that $\alpha\underset{\sim}{u}$ and $\underset{\sim}{u}$ point in the same direction, then

$$\text{area}\,\{\alpha\underset{\sim}{u},\underset{\sim}{v}\} = |\alpha\underset{\sim}{u}||\underset{\sim}{v}|\sin\theta = |\alpha||\underset{\sim}{u}||\underset{\sim}{v}|\sin\theta .$$

If $\alpha < 0$ so that $\alpha\underset{\sim}{u}$ and $\underset{\sim}{u}$ point in opposite directions, then

$$\text{area}\,\{\alpha\underset{\sim}{u},\underset{\sim}{v}\} = |\alpha\underset{\sim}{u}||\underset{\sim}{v}|\sin(\pi-\theta) = |\alpha||\underset{\sim}{u}|\sin\theta .$$

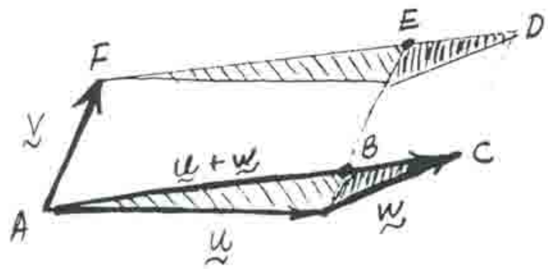
Thus, the area of a parallelogram is <u>positive homogeneous</u> in the sense that

$$\text{area}\,\{\alpha\underset{\sim}{u},\underset{\sim}{v}\} = |\alpha|\,\text{area}\,\{\underset{\sim}{u},\underset{\sim}{v}\} .$$

Similarly,

$$\text{area}\{\underset{\sim}{u}, \beta\underset{\sim}{v}\} = |\beta|\,\text{area}\{\underset{\sim}{u}, \underset{\sim}{v}\}.$$

Next, consider the area generated $\underset{\sim}{u}+\underset{\sim}{w}$ and $\underset{\sim}{v}$, where $\underset{\sim}{w}$ is another vector in the Euclidean plane. Let us examine the case where the angle between $\underset{\sim}{u}$ and $\underset{\sim}{w}$ is acute.



The area generated by $\underset{\sim}{u}+\underset{\sim}{w}$ and $\underset{\sim}{v}$ is the sum of the areas of the parallelograms ABEF and BCDE. But these areas are seen to be respectively equal to the areas generated by $\underset{\sim}{u}$ and $\underset{\sim}{v}$ and by $\underset{\sim}{w}$ and $\underset{\sim}{v}$. The case where the angle between $\underset{\sim}{u}$ and $\underset{\sim}{w}$ is obtuse can be handled in the same way, and thus the area of a parallelogram is <u>additive</u> in the sense that

$$\text{area}\{\underset{\sim}{u}+\underset{\sim}{w}, \underset{\sim}{v}\} = \text{area}\{\underset{\sim}{u}, \underset{\sim}{v}\} + \text{area}\{\underset{\sim}{w}, \underset{\sim}{v}\}.$$

Similarly,

$$\text{area}\{\underset{\sim}{u}, \underset{\sim}{v}+\underset{\sim}{w}\} = \text{area}\{\underset{\sim}{u}, \underset{\sim}{v}\} + \text{area}\{\underset{\sim}{u}, \underset{\sim}{w}\}.$$

Finally, we note that the area of the parallelogram generated by $\underset{\sim}{u}$ and $\underset{\sim}{v}$ is independent of the ordering; i.e.,

$$\text{area}\,\{\underset{\sim}{u}, \underset{\sim}{v}\} = \text{area}\,\{\underset{\sim}{v}, \underset{\sim}{u}\}.$$

Before generalizing the notion of area to higher dimensions, we need to review the concept of a permutation.[1]

Definition A2.7.1. If each of the first n integers

$$\mathbb{I}_n = \{1, 2, \cdots, n\}$$

appears once and only once in the list $(\sigma_1, \sigma_2, \cdots, \sigma_n)$, then $(\sigma_1, \sigma_2, \cdots, \sigma_n)$ is a permutation of $\mathbb{I}_n$. When in a permutation an integer precedes a smaller integer, the permutation is said to contain an inversion. A permutation is even or odd according as the total number of inversions it contains is even or odd.

Of course, the total number of inversions in a permutation can be found by counting the number of smaller integers following each integer of the permutation. E.g., the permutation $(6, 1, 4, 3, 2, 5)$ contains eight inversions since

6 is followed by 1, 4, 3, 2, and 5,
4 is followed by 3 and 2,
3 is followed by 2.

---

[1] See HOHN's Elementary Matrix Algebra for a comprehensive development along the line adopted here. See NOLL's Finite-Dimensional Spaces for a more abstract treatment.

Definition A2.7.2. The n-*dimensional alternating symbol* $\varepsilon_{\sigma_1\sigma_2\cdots\sigma_n}$ is defined by

$$\varepsilon_{\sigma_1\sigma_2\cdots\sigma_n} = \begin{cases} 1 & \text{if } (\sigma_1, \sigma_2, \cdots, \sigma_n) \text{ is an even permutation of } \mathbb{I}_n \\ -1 & \text{" " " odd " " "} \\ 0 & \text{" " " not a " " "} \end{cases}.$$

The next step in our generalization of parallelogram area is

Definition A2.7.3. A *volume-orientation functional* Δ for an n-dimensional inner product space[1] $\mathcal{V}_n$ is a function on $\underbrace{\mathcal{V}_n \times \mathcal{V}_n \times \cdots \times \mathcal{V}_n}_{n \text{ times}}$ to $\mathbb{R}$ which obeys the following axioms:

(O1) *Multilinearity*. Δ is linear in each argument; i.e., $\forall\ \underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n, \underset{\sim}{v} \in \mathcal{V}_n$ and $\forall \alpha \in \mathbb{R}$

$$\Delta(\underset{\sim}{u}_1, \cdots, \underset{\sim}{u}_{i-1}, \alpha\underset{\sim}{u}_i, \underset{\sim}{u}_{i+1}, \cdots, \underset{\sim}{u}_n)$$

$$= \alpha\Delta(\underset{\sim}{u}_1, \cdots, \underset{\sim}{u}_{i-1}, \underset{\sim}{u}_i, \underset{\sim}{u}_{i+1}, \cdots, \underset{\sim}{u}_n)$$

and

$$\Delta(\underset{\sim}{u}_1, \cdots, \underset{\sim}{u}_{i-1}, \underset{\sim}{u}_i + \underset{\sim}{v}, \underset{\sim}{u}_{i+1}, \cdots, \underset{\sim}{u}_n)$$

$$= \Delta(\underset{\sim}{u}_1, \cdots, \underset{\sim}{u}_{i-1}, \underset{\sim}{u}_i, \underset{\sim}{u}_{i+1}, \cdots, \underset{\sim}{u}_n) + \Delta(\underset{\sim}{u}_1, \cdots, \underset{\sim}{u}_{i-1}, \underset{\sim}{v}, \underset{\sim}{u}_{i+1}, \cdots, \underset{\sim}{u}_n)$$

for any $i \in \mathbb{I}_n$;

---

[1] Cf. Theorem A2.3.2.

(O2) Skew-Symmetry. Δ is skew-symmetric w.r.t. all arguments; i.e., $\forall\ \sigma_1, \sigma_2, \cdots, \sigma_n \in \mathbb{I}_n$

$$\Delta(\underset{\sim}{u}_{\sigma_1}, \underset{\sim}{u}_{\sigma_2}, \cdots, \underset{\sim}{u}_{\sigma_n}) = \varepsilon_{\sigma_1 \sigma_2 \cdots \sigma_n} \Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n);$$ [1]

(O3) If $\{\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle n\rangle}\}$ is an orthonormal basis for $\mathcal{V}_n$, then

$$|\Delta(\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle n\rangle})| = 1.$$

Next, we turn our attention to demonstrating that it is always possible to find a volume-orientation functional for $\mathcal{V}_n$; in fact, there will turn out to be exactly two, with one being the negative of the other.

Theorem A2.7.1. If $\exists$ a functional $\Delta: \mathcal{V}_n \times \mathcal{V}_n \times \cdots \times \mathcal{V}_n \to \mathbb{R}$ satisfying (O1) and (O2), it must have the form

$$\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = \varepsilon_{\sigma_1 \sigma_2 \cdots \sigma_n} u_1^{\sigma_1} u_2^{\sigma_2} \cdots u_n^{\sigma_n} \Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n),$$ [2]

where $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ is any basis for $\mathcal{V}_n$ and

$$\underset{\sim}{u}_i = u_i^{\,j} \underset{\sim}{e}_j, \quad i = 1, 2, \cdots, n.$$ [3]

---

[1] Here $(\sigma_1, \sigma_2, \cdots, \sigma_n)$ need not be a permutation of $\mathbb{I}_n$; e.g., $(\sigma_1, \sigma_2, \sigma_3) = (3, 1, 1)$ is allowed.

[2,3] As usual, the summation convention (p. A2.5.12) applies.

Proof. Since $\underset{\sim}{u}_i = \underset{\sim}{u}_i^j \underset{\sim}{e}_j$, repeated application of (O1) $\Longrightarrow$

$$\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = u_1^{\sigma_1} u_2^{\sigma_2} \cdots u_n^{\sigma_n} \Delta(\underset{\sim}{e}_{\sigma_1}, \underset{\sim}{e}_{\sigma_2}, \cdots, \underset{\sim}{e}_{\sigma_n}).$$

Then (O2) $\Longrightarrow$

$$\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = u_1^{\sigma_1} u_2^{\sigma_2} \cdots u_n^{\sigma_n} \varepsilon_{\sigma_1\sigma_2\cdots\sigma_n} \Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n). \square$$

We could have worked just as well with covariant components of the $\underset{\sim}{u}$'s in the above representation theorem. Then we would have obtained

$$\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = \varepsilon^{\sigma_1\sigma_2\cdots\sigma_n} u_{\sigma_1}^{(1)} u_{\sigma_2}^{(2)} \cdots u_{\sigma_n}^{(n)} \Delta(\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n),$$

where $\varepsilon^{\sigma_1\sigma_2\cdots\sigma_n}$ has the same meaning as $\varepsilon_{\sigma_1\sigma_2\cdots\sigma_n}$ and

$$\underset{\sim}{u}_i = u_j^{(i)} \underset{\sim}{e}^j, \quad (i = 1, 2, \cdots, n),$$

with $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ being the dual basis of $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$.

The above representations do not really define Δ because Δ evaluated on a basis appears on the r.h.s.'s. We can get around this by employing an orthonormal basis and utilizing Axiom (O3).

Theorem A2.7.2. $\exists$ <u>exactly two volume-orientation functionals for a given</u> n-<u>dimensional inner product space</u> V_n, <u>and one is</u>

the negative of the other. In fact, if $\{\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle n\rangle}\}$ is an orthonormal basis for $\mathcal{V}_n$ so that $\underset{\sim}{u}_i \in \mathcal{V}_n$ has the unique representation

$$\underset{\sim}{u}_i = u_i^{\langle j\rangle} \underset{\sim}{e}_{\langle j\rangle} \quad , \; i = 1, 2, \cdots, n \; ,$$

then

$$\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = \pm \varepsilon_{\sigma_1 \sigma_2 \cdots \sigma_n} u_1^{\langle \sigma_1\rangle} u_2^{\langle \sigma_2\rangle} \cdots u_n^{\langle \sigma_n\rangle} \; ,$$

with $+$ or $-$ according as $\Delta(\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle n\rangle}) = +1$ or -1.

Proof. If Δ exists, Theorem A2.7.1 $\Rightarrow$

$$\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = \varepsilon_{\sigma_1 \sigma_2 \cdots \sigma_n} u_1^{\langle \sigma_1\rangle} u_2^{\langle \sigma_2\rangle} \cdots u_n^{\langle \sigma_n\rangle} \Delta(\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle n\rangle}) \, .$$

But (O3) $\Rightarrow \Delta(\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle 3\rangle}) = \pm 1$. Hence, if Δ exists, it must be given by either

$$\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = +\varepsilon_{\sigma_1 \sigma_2 \cdots \sigma_n} u_1^{\langle \sigma_1\rangle} u_2^{\langle \sigma_2\rangle} \cdots u_n^{\langle \sigma_n\rangle}$$

or

$$\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = -\varepsilon_{\sigma_1 \sigma_2 \cdots \sigma_n} u_1^{\langle \sigma_1\rangle} u_2^{\langle \sigma_2\rangle} \cdots u_n^{\langle \sigma_n\rangle} \; ,$$

with the sign in accord with the value of Δ on the orthonormal basis employed. It is easy, and essential, to check that each of these tentative Δ's does, indeed, satisfy the axioms. The details are left as Exercise A2.7.1. [The identity $\sum_{\sigma} \varepsilon_{\sigma_1 \sigma_2 \cdots \sigma_n} a_{\tau_1 \sigma_1} a_{\tau_2 \sigma_2} \cdots a_{\tau_n \sigma_n} = \varepsilon_{\tau_1 \tau_2 \cdots \tau_n} \det[a_{ij}]$ (see HOHN's Elementary Matrix Algebra) will be useful in this regard.] $\square$

Note that even though the volume-orientation functional was introduced abstractly through axioms expressing its properties, we now know that it does exist; and we even have a formula for calculating it. A2.7.9

Definition A2.7.4. An n-dimensional inner product space $\mathcal{V}_n$ equipped with a volume-orientation functional Δ is said to be oriented by Δ.

According to Theorem A2.7.2, there are two, and only two, ways to orient $\mathcal{V}_n$. For the remainder of this section $\mathcal{V}_n$ is an oriented n-dimensional inner product space.

Now we are in position for the following generalization of the notion of area.

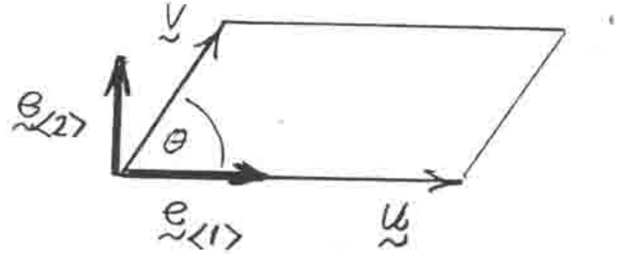
Definition A2.7.5. The volume spanned by a set of n elements $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n\} \subset \mathcal{V}_n$ is $|\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n)|$.

An easy but important result is

Theorem A2.7.3. The volume spanned by $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n\} \subset \mathcal{V}_n$ is independent of the choice of volume-orientation functional for $\mathcal{V}_n$, and also it is independent of the ordering of the n elements $\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n$.

Proof. Exercise A2.7.2. □

Exercise A2.7.3. Since $\mathcal{V}_2$ models the set of vectors in the Euclidean plane, the volume spanned by a pair of nonparallel elements $\underset{\sim}{u}$ and $\underset{\sim}{v}$ of $\mathcal{V}_2$ should be the area (in the usual geometrical sense discussed at the beginning of the present section) of the parallelogram that they generate. Show that this is, indeed, the case. (Hint: Employ an orthonormal basis as indicated in the figure below.)



Exercise A2.7.4. Consider the volume spanned by the singleton $\{\underset{\sim}{u}\} \subset \mathcal{V}_1$.

Of course, in Exercise A2.7.3 if $\underset{\sim}{u}$ and $\underset{\sim}{v}$ are parallel, then the parallelogram collapses and has zero area. In view of Definition A2.3.6 and the remarks that preceded it, this suggests

Theorem A2.7.4. <u>A set</u> $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n\} \subset \mathcal{V}_n$ <u>is linearly dependent iff</u> $\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = 0$; <u>i.e., iff the volume spanned by</u> $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n\}$ <u>is zero</u>.

Proof. Suppose that $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n\}$ is linearly dependent. Then one of the $\underset{\sim}{u}$'s is a linear combination of the rest. By relabeling, if necessary, we can assume that

$$\underset{\sim}{u}_1 = \sum_{i=2}^{n} \alpha_i \underset{\sim}{u}_i .$$

Then

$$\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = \Delta\Big(\sum_{i=2}^{n} \alpha_i \underset{\sim}{u}_i, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n\Big) \quad \text{(substitution,}$$

$$= \sum_{i=2}^{n} \alpha_i \, \Delta(\underset{\sim}{u}_i, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) \quad (\text{linearity})$$

$$= 0 \quad (\text{skew-symmetry}).$$

In the last step, we are using the fact that because of the skew-symmetry, Δ vanishes whenever any two of its arguments are equal. E.g.,

$$\Delta(\underset{\sim}{u}_2, \underset{\sim}{u}_2, \underset{\sim}{u}_3, \cdots, \underset{\sim}{u}_n) = \varepsilon_{223\cdots n} \, \Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n);$$

but $\varepsilon_{223\cdots n} = 0$ since $(2, 2, 3, \cdots, n)$ is not a permutation of $\mathrm{I\!I}_n$.

Conversely, suppose that $\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = 0$. Assume that $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n\}$ is linearly independent. Then by Theorem A2.2.8, $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n\}$ is a basis for $\mathcal{V}_n$. It then follows from the representation Theorem A2.7.1 that $\Delta(\underset{\sim}{v}_1, \underset{\sim}{v}_2, \cdots, \underset{\sim}{v}_n) = 0$ for any set of n elements of $\mathcal{V}_n$. This contradicts the axiom that $\Delta(\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle n\rangle}) = 1$. $\therefore$ $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n\}$ cannot be linearly independent; hence, it is linearly dependent. $\square$

By Theorem A2.7.4, if $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ is a basis for $\mathcal{V}_n$, then $\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n) \neq 0$. But $\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n)$ could be either positive or negative. This leads us to

Definition A2.7.6. <u>A basis</u> $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ <u>for</u> $\mathcal{V}_n$ <u>is</u> *positive* <u>or</u> *negative* <u>according as</u> $\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n) > 0$ <u>or</u>

< 0. Two bases $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ and $\{\underset{\sim}{\bar{e}}_1, \underset{\sim}{\bar{e}}_2, \cdots, \underset{\sim}{\bar{e}}_n\}$ are like-handed if

$$\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n)\Delta(\underset{\sim}{\bar{e}}_1, \underset{\sim}{\bar{e}}_2, \cdots, \underset{\sim}{\bar{e}}_n) > 0 .$$

As an immediate consequence of Theorem A2.7.2, we have

Theorem A2.7.5. The property of like-handedness is independent of the choice of volume-orientation functional.

We can get an interesting result by choosing the $\underset{\sim}{u}$'s in the representation Theorem A2.7.1 to be the elements of the dual basis.

$$\Delta(\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n) = \varepsilon_{\sigma_1\sigma_2\cdots\sigma_n}(\underset{\sim}{e}^1)^{\sigma_1}(\underset{\sim}{e}^2)^{\sigma_2}\cdots(\underset{\sim}{e}^n)^{\sigma_n}\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n).$$

By Theorem A2.5.8 and Definition A2.5.3,

$$(\underset{\sim}{e}^1)^{\sigma_1} = \underset{\sim}{e}^{\sigma_1} \cdot \underset{\sim}{e}^1 = g^{\sigma_1 1} \text{ , etc.}$$

$\therefore$

$$\Delta(\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n) = \underbrace{\varepsilon_{\sigma_1\sigma_2\cdots\sigma_n} g^{\sigma_1 1} g^{\sigma_2 2} \cdots g^{\sigma_n n}}_{\det[g^{ij}]} \Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n).$$

The last step is essentially the column expansion definition of the determinant[1] of a square matrix. By this and

[1] See, e.g., HOHN's Elementary Matrix Algebra.

Theorem A2.5.10 (vii), we are lead to

Theorem A2.7.6. Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n\}$ and $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n\}$ be a basis and its dual for an oriented n-dimensional inner product space $\mathcal{V}_n$. Then

(i) $$\frac{\Delta(\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n)}{\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n)} = \det[g^{ij}] = \frac{1}{\det[g_{ij}]};$$

(ii) $$\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n)\,\Delta(\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n) = 1;$$

(iii) $$[\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n)]^2 = \det[g_{ij}],\ [\Delta(\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n)]^2 = \det[g^{ij}].$$

Proof. (i): This part was established by the remarks preceding the statement of the theorem.

(ii)[1]: Let $\{\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle n\rangle}\}$ be an orthonormal basis for $\mathcal{V}_n$. Then by Theorems A2.7.1 and A2.3.14,

$$\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n)\,\Delta(\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n) =$$

$$= \varepsilon_{\sigma_1\sigma_2\cdots\sigma_n}\, a_1^{\ \sigma_1} a_2^{\ \sigma_2} \cdots a_n^{\ \sigma_n}\, \Delta(\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle n\rangle}) \times$$

$$\times\ \varepsilon_{\tau_1\tau_2\cdots\tau_n}\, b^{1\tau_1} b^{2\tau_2} \cdots b^{n\tau_n}\, \Delta(\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \cdots, \underset{\sim}{e}_{\langle n\rangle}),$$

where $a_i^{\ j} = \underset{\sim}{e}_{\langle j\rangle} \cdot \underset{\sim}{e}_i$ and $b^{ij} = \underset{\sim}{e}_{\langle j\rangle} \cdot \underset{\sim}{e}^i$. By Axiom (O3),

---

[1] This particular proof was shown to the author by Dr. Seyoung Im.

$$[\Delta(\underset{\sim}{e}_{<1>}, \underset{\sim}{e}_{<2>}, \cdots, \underset{\sim}{e}_{<n>})]^2 = 1 ;$$

so we are left with

$$\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \cdots, \underset{\sim}{e}_n)\, \Delta(\underset{\sim}{e}^1, \underset{\sim}{e}^2, \cdots, \underset{\sim}{e}^n) =$$

$$= \underbrace{\varepsilon_{\sigma_1 \sigma_2 \cdots \sigma_n}\, a_1{}^{\sigma_1} a_2{}^{\sigma_2} \cdots a_n{}^{\sigma_n}}_{\det [a_i{}^j]}\; \underbrace{\varepsilon_{\tau_1 \tau_2 \cdots \tau_n}\, b^{1\tau_1} b^{2\tau_2} \cdots b^{n\tau_n}}_{\det [b^{k\ell}]}$$

$$= \det A \det B$$

$$= \det A \det B^T$$

$$= \det (A B^T)$$

$$= \det \left[\sum_{j=1}^{n} a_i{}^j b^{kj} \right] ,$$

where have used some obvious matrix notation as well as some standard results about determinants of matrices.[1] Unraveling this notation, we get

$$\sum_{j=1}^{n} a_i{}^j\, b^{kj} = \sum_{j=1}^{n} (\underset{\sim}{e}_{<j>} \cdot \underset{\sim}{e}_i)(\underset{\sim}{e}_{<j>} \cdot \underset{\sim}{e}^k) \quad \text{(substitution)}$$

$$= \underset{\sim}{e}_i \cdot \underset{\sim}{e}^k \quad \text{(Theorem A2.3.16)}$$

$$= \delta_i^k \quad \text{(Theorem A2.5.10 (ii))}.$$

---

[1] See, e.g., HOHN's <u>Elementary Matrix Algebra</u>.

Since $\det[\delta_i^{\,k}] = 1$, we have the desired result.

(iii): These results follow directly from combining (i) and (ii). $\square$

The following observation is occasionally useful.

<u>Theorem</u> A2.7.7. <u>Let</u> $\{\underset{\sim}{e}_{<1>}, \underset{\sim}{e}_{<2>}, \cdots, \underset{\sim}{e}_{<n>}\}$ <u>be a positive orthonormal basis for</u> $\mathcal{V}_n$. <u>Then for</u> $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n\} \subset \mathcal{V}_n$,

$$\Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_n) = \det\left[u_i^{<j>}\right],$$

<u>where</u>

$$\underset{\sim}{u}_i = u_i^{<j>} \underset{\sim}{e}_{<j>}.$$

<u>Proof</u>. By Theorem A2.7.2 and Definition A2.7.6,

$$\begin{aligned} \Delta(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, u_n) &= \varepsilon_{\sigma_1\sigma_2\cdots\sigma_n} u_1^{<\sigma_1>} u_2^{<\sigma_2>} \cdots u_n^{<\sigma_n>} \\ &= \det\left[u_i^{<j>}\right] \quad (\text{definition}). \;\square \end{aligned}$$

We are now in a good position to establish the following useful relationships between alternating symbols and Kronecker deltas in the "three-dimensional" case. Note that despite our method of proof, the result stands by itself independent of the concept of a linear space.

Theorem A2.7.8. For $i, j, k, p, q, r \in \mathbb{I}_3$,

$$\text{(i)}\quad \varepsilon_{ijk}\,\varepsilon_{pqr} = \det\begin{bmatrix} \delta_{ip} & \delta_{iq} & \delta_{ir} \\ \delta_{jp} & \delta_{jq} & \delta_{jr} \\ \delta_{kp} & \delta_{kq} & \delta_{kr} \end{bmatrix} ;$$

$$\text{(ii)}\quad \sum_{i=1}^{3} \varepsilon_{ijk}\,\varepsilon_{iqr} = \delta_{jq}\delta_{kr} - \delta_{jr}\delta_{kq} ;$$

$$\text{(iii)}\quad \sum_{i=1}^{3}\sum_{j=1}^{3} \varepsilon_{ijk}\,\varepsilon_{ijr} = 2\,\delta_{kr} ;$$

$$\text{(iv)}\quad \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} \varepsilon_{ijk}\,\varepsilon_{ijk} = 6 .$$

Proof. (i): Let $\{\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \underset{\sim}{e}_{\langle 3\rangle}\}$ be a positive orthonormal basis for $\mathcal{V}_3$. Then

$$\begin{aligned}\Delta(\underset{\sim}{e}_{\langle i\rangle}, \underset{\sim}{e}_{\langle j\rangle}, \underset{\sim}{e}_{\langle k\rangle}) &= \varepsilon_{ijk}\,\Delta(\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \underset{\sim}{e}_{\langle 3\rangle}) \quad (\text{skew-symmetry of } \Delta) \\ &= \varepsilon_{ijk} \quad (\text{positiveness of } \{\underset{\sim}{e}_{\langle 1\rangle}, \underset{\sim}{e}_{\langle 2\rangle}, \underset{\sim}{e}_{\langle 3\rangle}\})\end{aligned}$$

Thus, by the previous theorem,

$$\varepsilon_{ijk} = \det\begin{bmatrix} (\underset{\sim}{e}_{\langle i\rangle})^{\langle 1\rangle} & (\underset{\sim}{e}_{\langle i\rangle})^{\langle 2\rangle} & (\underset{\sim}{e}_{\langle i\rangle})^{\langle 3\rangle} \\ (\underset{\sim}{e}_{\langle j\rangle})^{\langle 1\rangle} & (\underset{\sim}{e}_{\langle j\rangle})^{\langle 2\rangle} & (\underset{\sim}{e}_{\langle j\rangle})^{\langle 3\rangle} \\ (\underset{\sim}{e}_{\langle k\rangle})^{\langle 1\rangle} & (\underset{\sim}{e}_{\langle k\rangle})^{\langle 2\rangle} & (\underset{\sim}{e}_{\langle k\rangle})^{\langle 3\rangle} \end{bmatrix}$$

$$= \det \begin{bmatrix} \underset{\sim}{e}_{\langle i \rangle} \cdot \underset{\sim}{e}_{\langle 1 \rangle} & \underset{\sim}{e}_{\langle i \rangle} \cdot \underset{\sim}{e}_{\langle 2 \rangle} & \underset{\sim}{e}_{\langle i \rangle} \cdot \underset{\sim}{e}_{\langle 3 \rangle} \\ \underset{\sim}{e}_{\langle j \rangle} \cdot \underset{\sim}{e}_{\langle 1 \rangle} & \underset{\sim}{e}_{\langle j \rangle} \cdot \underset{\sim}{e}_{\langle 2 \rangle} & \underset{\sim}{e}_{\langle j \rangle} \cdot \underset{\sim}{e}_{\langle 3 \rangle} \\ \underset{\sim}{e}_{\langle k \rangle} \cdot \underset{\sim}{e}_{\langle 1 \rangle} & \underset{\sim}{e}_{\langle k \rangle} \cdot \underset{\sim}{e}_{\langle 2 \rangle} & \underset{\sim}{e}_{\langle k \rangle} \cdot \underset{\sim}{e}_{\langle 3 \rangle} \end{bmatrix} \quad \text{(Theorem A2.3.16)}$$

$$= \det \begin{bmatrix} \delta_{i1} & \delta_{i2} & \delta_{i3} \\ \delta_{j1} & \delta_{j2} & \delta_{j3} \\ \delta_{k1} & \delta_{k2} & \delta_{k3} \end{bmatrix} \quad \text{(Definition A2.3.8)}.$$

Then

$$\varepsilon_{ijk}\,\varepsilon_{pqr} = \det \begin{bmatrix} \delta_{i1} & \delta_{i2} & \delta_{i3} \\ \delta_{j1} & \delta_{j2} & \delta_{j3} \\ \delta_{k1} & \delta_{k2} & \delta_{k3} \end{bmatrix} \det \begin{bmatrix} \delta_{p1} & \delta_{p2} & \delta_{p3} \\ \delta_{q1} & \delta_{q2} & \delta_{q3} \\ \delta_{r1} & \delta_{r2} & \delta_{r3} \end{bmatrix}$$

$$= \det \begin{bmatrix} \delta_{i1} & \delta_{i2} & \delta_{i3} \\ \delta_{j1} & \delta_{j2} & \delta_{j3} \\ \delta_{k1} & \delta_{k2} & \delta_{k3} \end{bmatrix} \det \begin{bmatrix} \delta_{p1} & \delta_{q1} & \delta_{r1} \\ \delta_{p2} & \delta_{q2} & \delta_{r2} \\ \delta_{p3} & \delta_{q3} & \delta_{r3} \end{bmatrix} \quad (\det A = \det A^T)$$

$$= \det \left(\begin{bmatrix} \delta_{i1} & \delta_{i2} & \delta_{i3} \\ \delta_{j1} & \delta_{j2} & \delta_{j3} \\ \delta_{k1} & \delta_{k2} & \delta_{k3} \end{bmatrix} \begin{bmatrix} \delta_{p1} & \delta_{q1} & \delta_{r1} \\ \delta_{p2} & \delta_{q2} & \delta_{r2} \\ \delta_{p3} & \delta_{q3} & \delta_{r3} \end{bmatrix} \right) \quad (\det(AB) = (\det A)(\det B))$$

$$= \det \begin{bmatrix} \left(\sum_{m=1}^{3} \delta_{im}\delta_{pm}\right) & \left(\sum_{m=1}^{3} \delta_{im}\delta_{qm}\right) & \left(\sum_{m=1}^{3} \delta_{im}\delta_{rm}\right) \\ \left(\sum_{m=1}^{3} \delta_{jm}\delta_{pm}\right) & - & - \\ - & - & - \end{bmatrix} \quad \text{(matrix multiplication)}$$

$$= \det \begin{bmatrix} \delta_{ip} & \delta_{iq} & \delta_{ir} \\ \delta_{jp} & \delta_{jq} & \delta_{jr} \\ \delta_{kp} & \delta_{kq} & \delta_{kr} \end{bmatrix} \quad \left(\text{Substitution property of Kronecker delta}\right)$$

(ii): Expanding the above determinant, we have

$$\varepsilon_{ijk}\varepsilon_{pqr} = \delta_{ip}(\delta_{jq}\delta_{kr} - \delta_{kq}\delta_{jr}) - \delta_{iq}(\delta_{jp}\delta_{kr} - \delta_{kp}\delta_{jr})$$
$$+ \delta_{ir}(\delta_{jp}\delta_{kq} - \delta_{kp}\delta_{jq})\ ;$$

and then

$$\sum_{i=1}^{3} \varepsilon_{ijk}\varepsilon_{iqr} = \sum_{i=1}^{3}\Big[\delta_{ii}(\delta_{jq}\delta_{kr} - \delta_{kq}\delta_{jr}) - \delta_{iq}(\delta_{ji}\delta_{kr} - \delta_{ki}\delta_{jr})$$
$$+ \delta_{ir}(\delta_{ji}\delta_{kq} - \delta_{ki}\delta_{jq})\Big]$$

$$= \sum_{i=1}^{3}\Big[\delta_{ii}(\delta_{jq}\delta_{kr} - \delta_{kq}\delta_{jr}) - (\delta_{iq}\delta_{ji})\delta_{kr} + (\delta_{iq}\delta_{ki})\delta_{jr}$$
$$+ (\delta_{ir}\delta_{ji})\delta_{kq} - (\delta_{ir}\delta_{ki})\delta_{jq}\Big] \quad (\text{calculation})$$

$$= 3(\delta_{jq}\delta_{kr} - \delta_{kq}\delta_{jr}) - \delta_{jq}\delta_{kr} + \delta_{kq}\delta_{jr} \quad \left(\text{properties of Kronecker delta}\right)$$
$$+ \delta_{jr}\delta_{kq} - \delta_{kr}\delta_{jq}$$

$$= \delta_{jq}\delta_{kr} - \delta_{kq}\delta_{jr} \quad (\text{calculation}).$$

(iii), (iv) : *Exercise* A2.7.5. $\square$

Supplementary Reading

BORISENKO and TARAPOV, Vector and Tensor Analysis with Applications

GREUB, Linear Algebra

MARTIN and MIZEL, Introduction to Linear Algebra

NICKERSON, SPENCER, and STEENROD, Advanced Calculus

A2.8. The Vector Product

Almost certainly you have had some experience with the vector or cross product of two vectors from three-dimensional Euclidean space. Elementary treatments of the vector product often have the defect that somewhere in the development loose or imprecise notions such as "right-hand rule" or "like-handed triads" (usually explained with sketches of the author's extremities) are snuck in to give the direction of the vector product. In this section, we shall use the concept of volume orientation functionals to introduce the vector product in a mathematically sound way.

The vector product of two vectors is strictly a three-dimensional concept, and accordingly for the whole of this section $\mathcal{V}_3$ denotes an oriented[1] three-dimensional inner product space[2]. The associated volume orientation functional is Δ.

We get the cross product from the volume orientation functional as follows. Let $\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{w} \in \mathcal{V}_3$. Then for fixed $\underset{\sim}{u}$ and $\underset{\sim}{v}$, $\Delta(\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{w})$ is a linear functional in $\underset{\sim}{w}$. More precisely, $\Delta(\underset{\sim}{u}, \underset{\sim}{v}, \cdot)$ is a linear function on $\mathcal{V}_3$ to $\mathbb{R}$. By the representation theorem for linear functionals[3], $\exists$ a unique vector, which we

---

[1] Cf. Theorem A2.7.2 and Definition A2.7.4.

[2] Cf. Theorem A2.3.2,

[3] Theorem A2.5.1.

denote by $\underset{\sim}{u}\times\underset{\sim}{v}$ since it can be expected to depend on $\underset{\sim}{u}$ and $\underset{\sim}{v}$, $\ni$ $\Delta(\underset{\sim}{u},\underset{\sim}{v},\underset{\sim}{w}) = (\underset{\sim}{u}\times\underset{\sim}{v})\cdot\underset{\sim}{w}$ $\forall\, \underset{\sim}{w}\in\mathcal{V}_3$. Hence, we have established

Theorem A2.8.1. To each ordered pair $(\underset{\sim}{u},\underset{\sim}{v})$ of elements from $\mathcal{V}_3$ there corresponds a unique element $\underset{\sim}{u}\times\underset{\sim}{v}\in\mathcal{V}_3$ $\ni$

$$\Delta(\underset{\sim}{u},\underset{\sim}{v},\underset{\sim}{w}) = (\underset{\sim}{u}\times\underset{\sim}{v})\cdot\underset{\sim}{w} \quad \forall\, \underset{\sim}{w}\in\mathcal{V}_3 .$$

$\underset{\sim}{u}\times\underset{\sim}{v}$ is called the *vector product*[1] of $\underset{\sim}{u}$ and $\underset{\sim}{v}$.

Obviously, the vector product depends on the choice of volume orientation functional. By Theorem A2.7.2, this amounts to changing the vector product by a scalar factor of -1 (i.e., reversing its "sense") if the alternative orientation for $\mathcal{V}_3$ is chosen. This is where "right-hand rules", etc. enter into the elementary geometrical treatment.

Insert material on p. A2.8.2a

The following result provides component representations for the vector product.

Theorem A2.8.2. Let $\{\underset{\sim}{e}_1,\underset{\sim}{e}_2,\underset{\sim}{e}_3\}$ and $\{\underset{\sim}{e}^1,\underset{\sim}{e}^2,\underset{\sim}{e}^3\}$ be a basis and its dual for $\mathcal{V}_3$. Then $\forall\, \underset{\sim}{u},\underset{\sim}{v}\in\mathcal{V}_3$,

$$\underset{\sim}{u}\times\underset{\sim}{v} = \Delta(\underset{\sim}{e}_1,\underset{\sim}{e}_2,\underset{\sim}{e}_3)\,\varepsilon_{ijk}\, u^j v^k\, \underset{\sim}{e}^i$$

---

[1] Often, *cross product*. Sometimes the notation $\underset{\sim}{u}\wedge\underset{\sim}{v}$ is used, but we shall save this for the "skew product" of §A3.6.

INSERT for p. A2.8.2

Exercise A2.8.0. Prove that in terms of components relative to a positive orthonormal basis

$$(\underset{\sim}{u} \times \underset{\sim}{v}) \cdot \underset{\sim}{w} = \begin{vmatrix} u^{\langle 1\rangle} & u^{\langle 2\rangle} & u^{\langle 3\rangle} \\ v^{\langle 1\rangle} & v^{\langle 2\rangle} & v^{\langle 3\rangle} \\ w^{\langle 1\rangle} & w^{\langle 2\rangle} & w^{\langle 3\rangle} \end{vmatrix} .$$

---

return to p. A2.8.2

$$= \Delta(\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3)\, \varepsilon^{ijk} u_j v_k \mathbf{e}_i \, . \ )$$[1]

<u>In particular</u>, <u>in terms of a positive orthonormal basis</u> $\{\mathbf{e}_{\langle 1\rangle}, \mathbf{e}_{\langle 2\rangle}, \mathbf{e}_{\langle 3\rangle}\} = \{\mathbf{e}^{\langle 1\rangle}, \mathbf{e}^{\langle 2\rangle}, \mathbf{e}^{\langle 3\rangle}\}$,

$$\mathbf{u} \times \mathbf{v} = \varepsilon_{ijk}\, u^{\langle j\rangle} v^{\langle k\rangle} \mathbf{e}^{\langle i\rangle} .$$

Proof. By Theorem A2.7.1,

$$\Delta(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \varepsilon_{ijk}\, u^i v^j w^k \Delta(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3).$$

On defining $x_k = \Delta(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)\, \varepsilon_{ijk} u^i v^j$, we have

$$\Delta(\mathbf{u}, \mathbf{v}, \mathbf{w}) = x_k w^k .$$

$\therefore$ by Theorem A2.5.10 (iii),

$$\Delta(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \mathbf{x} \cdot \mathbf{w} ,$$

where $\mathbf{x} := x_k \mathbf{e}^k$. Then by Theorem A2.8.1, $\mathbf{x} = \mathbf{u} \times \mathbf{v}$. To get exact agreement with the assertion of the theorem, we note that by the skew-symmetry of the alternating symbol[2]

---

[1] Here, ε^{ijk} is simply the three-dimensional alternating symbol written with superscripts for the sake of the summation convention.

[2] See Definition A2.7.2.

$$x_k = \Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3)\, \varepsilon_{ijk}\, u^i v^j = -\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3)\, \varepsilon_{ikj}\, u^i v^j$$

$$= + \Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3)\, \varepsilon_{kij}\, u^i v^j .$$

To get the formula for the contravariant components of $\underset{\sim}{u} \times \underset{\sim}{v}$ start with

$$\Delta(\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{w}) = \varepsilon^{ijk} u_i v_j w_k\, \Delta(\underset{\sim}{e}^1, \underset{\sim}{e}^2, \underset{\sim}{e}^3)$$

and follow the same steps. □

We can use the above component representation of the vector product to get a geometrical interpretation which is the customary elementary definition. First we turn to some geometrical preliminaries.

Definition A2.8.1. Let $\underset{\sim}{u}$ and $\underset{\sim}{v}$ be any two nonparallel[1] elements of $\mathcal{V}_3$. Then the *plane* of $\underset{\sim}{u}$ and $\underset{\sim}{v}$ is $Lsp\{\underset{\sim}{u}, \underset{\sim}{v}\}$. A *unit normal* to the plane of $\underset{\sim}{u}$ and $\underset{\sim}{v}$ is an element $\underset{\sim}{n} \in \mathcal{V}_3$ with the properties that

$$\underset{\sim}{n} \cdot \underset{\sim}{w} = 0 \;\; \forall\, \underset{\sim}{w} \in Lsp\{\underset{\sim}{u}, \underset{\sim}{v}\} \;\; \text{and} \;\; |\underset{\sim}{n}| = 1 .$$

The following result is intuitively evident; but, as is usually the case in such matters, its proof is rather

[1] In particular, neither $\underset{\sim}{u}$ nor $\underset{\sim}{v}$ can be the zero element of $\mathcal{V}_3$; cf. Definition A2.3.6 and Theorem A2.2.3.

involved.

<u>Theorem A2.8.3. Let $\underset{\sim}{u}$ and $\underset{\sim}{v}$ be two nonparallel elements of $\mathcal{V}_3$. Then the plane of $\underset{\sim}{u}$ and $\underset{\sim}{v}$</u>

(i) <u>is a two-dimensional linear subspace of $\mathcal{V}_3$</u>

<u>and</u>

(ii) <u>has exactly two unit normals, with one the negative of the other.</u>

Proof. (i): (Since $\underset{\sim}{u}$ and $\underset{\sim}{v}$ are nonparallel, the set $\{\underset{\sim}{u}, \underset{\sim}{v}\}$ is linearly independent by definition. Also by construction, $\{\underset{\sim}{u}, \underset{\sim}{v}\}$ spans $Lsp\{\underset{\sim}{u}, \underset{\sim}{v}\}$. Hence, $\{\underset{\sim}{u}, \underset{\sim}{v}\}$ is a basis for $Lsp\{\underset{\sim}{u}, \underset{\sim}{v}\}$, and the dimension of $Lsp\{\underset{\sim}{u}, \underset{\sim}{v}\}$ is 2.)[1]

($Lsp\{\underset{\sim}{u}, \underset{\sim}{v}\}$ is a linear subspace by Theorem A2.1.12.)

(ii): For notational convenience, set $\underset{\sim}{u} = \underset{\sim}{e}_1$, $\underset{\sim}{v} = \underset{\sim}{e}_2$. By Theorem A2.2.9, $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2\}$ can be extended to a basis for $\mathcal{V}_3$; i.e., $\exists\ \underset{\sim}{e}_3 \in \mathcal{V}_3 \ni \{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3\}$ is a basis for $\mathcal{V}_3$.

Now we want to find an $\underset{\sim}{n} \in \mathcal{V}_3 \ni$

$$\underset{\sim}{n} \cdot \underset{\sim}{w} = 0 \quad \forall\ \underset{\sim}{w} \in Lsp\{\underset{\sim}{e}_1, \underset{\sim}{e}_2\}.$$

(It is (necessary and) sufficient to have $\underset{\sim}{n} \cdot \underset{\sim}{e}_1 = \underset{\sim}{n} \cdot \underset{\sim}{e}_2 = 0$.)[2] Then by Theorem A2.5.8 and Definition A2.5.2,

---

[1] See Definition A2.2.2.

[2] **Prove this** as Exercise A2.8.1.

$$\left.\begin{aligned} n_1 = \underset{\sim}{e}_1 \cdot \underset{\sim}{n} = 0 \\ n_2 = \underset{\sim}{e}_2 \cdot \underset{\sim}{n} = 0 \end{aligned}\right\} \Rightarrow \underset{\sim}{n} = n_3 \underset{\sim}{e}^3 ,$$

where $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \underset{\sim}{e}^3\}$ is the dual basis of $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3\}$. The condition $|\underset{\sim}{n}| = 1 \Rightarrow |n_3| = (|\underset{\sim}{e}^3|)^{-1}$. Thus,

$$\underset{\sim}{n} = \pm \frac{1}{|\underset{\sim}{e}^3|} \underset{\sim}{e}^3 .$$

At this point, we have two $\underset{\sim}{n}$'s, and one is the negative of the other. However, there may well be other pairs of unit normals because of the nonuniqueness of the element $\underset{\sim}{e}_3$ chosen to extend $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2\}$ to a basis. To investigate this, suppose that the basis is completed by appending a different element, say $\overset{*}{\underset{\sim}{e}}_3$. The dual basis of $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \overset{*}{\underset{\sim}{e}}_3\}$ is denoted by $\{\overset{*}{\underset{\sim}{e}}{}^1, \overset{*}{\underset{\sim}{e}}{}^2, \overset{*}{\underset{\sim}{e}}{}^3\}$. In terms of the original dual basis $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \underset{\sim}{e}^3\}$, the element $\overset{*}{\underset{\sim}{e}}{}^3$ has the representation

$$\overset{*}{\underset{\sim}{e}}{}^3 = a_1 \underset{\sim}{e}^1 + a_2 \underset{\sim}{e}^2 + a_3 \underset{\sim}{e}^3 .$$

By Theorems A2.5.8 and A2.5.4,

$$a_1 = \underset{\sim}{e}_1 \cdot \overset{*}{\underset{\sim}{e}}{}^3 = 0, \quad a_2 = \underset{\sim}{e}_2 \cdot \overset{*}{\underset{\sim}{e}}{}^3 = 0 ;$$

and $\therefore$

$$\overset{*}{\underset{\sim}{e}}{}^3 = a_3 \underset{\sim}{e}^3 .$$

Thus,

$$\frac{1}{|\overset{*}{\underset{\sim}{e}}{}^3|} \overset{*}{\underset{\sim}{e}}{}^3 = \frac{a_3}{|a_3|} \frac{1}{|\underset{\sim}{e}^3|} \underset{\sim}{e}^3 = \pm \frac{1}{|\underset{\sim}{e}^3|} \underset{\sim}{e}^3 ;$$

and we see that no matter how $\{e_1, e_2\}$ is extended to a basis, the resulting pair of n's is always the same. □

Now we are in a position to state and prove the following theorem which corresponds to the elementary geometrical definition of the vector product.

Theorem A2.8.4. Let u and v be arbitrary elements of $\mathcal{V}_3$ and let θ be the angle between them. Then

$$\text{(i)} \quad |u \times v| = |u||v| \sin\theta \qquad)^{1}$$

and

$$\text{(ii)} \quad u \times v = |u \times v| n,$$

where n is the unique unit normal to the plane[2] of u and v $\ni$ $\Delta(u, v, n) > 0$. Furthermore,

$$\text{(iii)} \quad |u \times v| = \Delta(u, v, n).$$

Proof. (i): This assertion follows easily from Theorem A2.8.2 with the aid of the identity

---

[1] Of course, if either $u = 0$ or $v = 0$, then the angle θ is not defined; but in this case Theorem A2.8.2 $\Rightarrow u \times v = 0$.

[2] If u and v are parallel, they do not span a plane; but in this case again Theorem A2.8.2 $\Longrightarrow u \times v = 0$. Work out the details of the argument as Exercise A2.8.2. The result also follows from $(u \times v) \cdot w = \Delta(u, v, w) = 0 \ \forall w$ since $\{u, v\}$ is linearly dependent.

$$\sum_{i=1}^{3} \varepsilon_{ijk}\,\varepsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} \quad (\text{Theorem A2.7.8 (ii)}).$$

The details are left as Exercise A2.8.3. Alternatively, one can get this result by combining the representation of Theorem A2.8.2 with the special choice of basis used in Exercise A2.7.3, Work out such a proof as Exercise A2.8.4.

(iii): If $\underset{\sim}{u}$ and $\underset{\sim}{v}$ are parallel, (i) $\Rightarrow \underset{\sim}{u} \times \underset{\sim}{v} = \underset{\sim}{0}$; and (iii) holds trivially. Accordingly, for the remainder of the proof we assume that $\underset{\sim}{u}$ and $\underset{\sim}{v}$ are not parallel. It follows from Theorem A2.8.3 that the plane of $\underset{\sim}{u}$ and $\underset{\sim}{v}$ has exactly two unit normals, one being the negative of the other. For either choice, it is not difficult to see that the set $\{\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{n}\}$ is linearly independent. The proof of this fact is left as Exercise A2.8.5. (Hint: Recall the construction of $\underset{\sim}{n}$ in the proof of Theorem A2.8.3.) Then by Theorem A2.7.4, $\Delta(\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{n}) \neq 0$; consequently, in view of the multilinearity of Δ, the requirement $\Delta(\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{n}) > 0$ does select $\underset{\sim}{n}$ uniquely. For suppose the two unit normals are $\underset{\sim}{n}_1$ and $\underset{\sim}{n}_2$, then

$$\Delta(\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{n}_2) = \Delta(\underset{\sim}{u}, \underset{\sim}{v}, -\underset{\sim}{n}_1) \quad (\text{substitution})$$

$$= -\Delta(\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{n}_1); \quad (\text{multilinearity})$$

and only one of $\Delta(\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{n}_1)$ and $\Delta(\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{n}_2)$ is positive.

By Theorem A2.2.8, we can take $\{\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{n}\}$ to be a basis for $\mathcal{V}_3$. For convenience, we write

$$\underset{\sim}{e}_1 := \underset{\sim}{u} \Rightarrow u^1 = 1,\ u^2 = u^3 = 0;$$

$$\underset{\sim}{e}_2 := \underset{\sim}{v} \Rightarrow v^1 = 0,\ v^2 = 1,\ v^3 = 0;$$

$$\underset{\sim}{e}_3 := \underset{\sim}{n} \Rightarrow n^1 = n^2 = 0,\ n^3 = 1.$$

Substitution of this information into the component formula of Theorem A2.8.2 yields

$$\underset{\sim}{u} \times \underset{\sim}{v} = \Delta(\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{n})\, \underset{\sim}{e}^3.$$

By Theorem A2.5.5, $\underset{\sim}{e}^3$ is uniquely determined by the conditions

$$\underset{\sim}{e}^3 \cdot \underset{\sim}{e}_1 = \underset{\sim}{e}^3 \cdot \underset{\sim}{u} = 0,$$

$$\underset{\sim}{e}^3 \cdot \underset{\sim}{e}_2 = \underset{\sim}{e}^3 \cdot \underset{\sim}{v} = 0,$$

$$\underset{\sim}{e}^3 \cdot \underset{\sim}{e}_3 = \underset{\sim}{e}^3 \cdot \underset{\sim}{n} = 1.$$

Clearly, $\underset{\sim}{e}^3 = \underset{\sim}{n}$ satisfies these conditions; and by the uniqueness, nothing else will. Hence, we have

$$\underset{\sim}{u} \times \underset{\sim}{v} = \Delta(\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{n})\, \underset{\sim}{n}.$$

Then

$$|\underset{\sim}{u}\times\underset{\sim}{v}| = |\Delta(\underset{\sim}{u},\underset{\sim}{v},n)|\,|\underset{\sim}{n}| \quad (\text{Theorem A2.3.6 (i)})$$

$$= \Delta(\underset{\sim}{u},\underset{\sim}{v},\underset{\sim}{n}) \quad (\Delta(\underset{\sim}{u},\underset{\sim}{v},\underset{\sim}{n})>0 \text{ and } |\underset{\sim}{n}|=1),$$

and ∴

$$\underset{\sim}{u}\times\underset{\sim}{v} = |\underset{\sim}{u}\times\underset{\sim}{v}|\,\underset{\sim}{n}.$$

(iii): Established above. □

Exercise A2.8.6. Use Theorem A2.8.4 to show that $|\Delta(\underset{\sim}{u},\underset{\sim}{v},\underset{\sim}{w})| = |(\underset{\sim}{u}\times\underset{\sim}{v})\cdot\underset{\sim}{w}|$ is the volume (in the usual geometrical sense) of the parallelepiped generated by $\underset{\sim}{u}$, $\underset{\sim}{v}$, and $\underset{\sim}{w}$. Do not hesitate to draw pictures.

The next theorem gathers some of the rules for manipulating the vector product. Each rule is accessible from Theorem A2.8.1, our abstract definition of the vector product in terms of the volume orientation functional, or Theorem A2.8.2, the representation of the vector product in terms of components, or Theorem A2.8.4, the geometrical interpretation of the vector product, but most of them are established most readily via the component representation.

Theorem A2.8.5. <u>The vector product has the following properties:</u>

(i) Skew-Symmetry. $\forall\ \underset{\sim}{u},\underset{\sim}{v}\in\mathcal{V}_3$

$$\underset{\sim}{v}\times\underset{\sim}{u} = -(\underset{\sim}{u}\times\underset{\sim}{v});$$

(ii) Homogeneity. $\forall \mathbf{u}, \mathbf{v} \in \mathcal{V}_3$ and $\forall \alpha, \beta \in \mathbb{R}$

$$(\alpha \mathbf{u}) \times (\beta \mathbf{v}) = (\alpha\beta)(\mathbf{u} \times \mathbf{v});$$

(iii) Distributivity w.r.t. addition. $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}_3$

$$\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w},$$

$$(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w};$$

(iv) $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}_3$

$$(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{v} \cdot \mathbf{w})\mathbf{u};$$ [1]

(v) For $\mathbf{u}, \mathbf{v} \in \mathcal{V}_3$

$$\mathbf{u} \times \mathbf{v} = \mathbf{0} \text{ iff } \mathbf{u} \text{ and } \mathbf{v} \text{ are parallel};$$

(vi) For $\mathbf{u} \in \mathcal{V}_3$

$$\mathbf{u} \times \mathbf{v} = \mathbf{0} \;\; \forall \mathbf{v} \in \mathcal{V}_3 \Rightarrow \mathbf{u} = \mathbf{0}.$$

Proof. Exercises A2.8.7–12. □

We end this section with an interesting application of the

---

[1] Thus, in general, $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} \neq \mathbf{u} \times (\mathbf{v} \times \mathbf{w})$; i.e., the vector product is nonassociative.

the component representation formula of Theorem A2.8.2. To wit,

$$\underset{\sim}{e}_2 \times \underset{\sim}{e}_3 = \Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3)\, \underbrace{\varepsilon_{ijk} \underbrace{(\underset{\sim}{e}_2)^j}_{\delta^j_2} \underbrace{(\underset{\sim}{e}_3)^k}_{\delta^k_3}}_{\varepsilon_{i23}} \underset{\sim}{e}^i = \Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3)\underset{\sim}{e}^1 ,$$

which provides a formula for calculating $\underset{\sim}{e}^1$. Formally, we have established

Theorem A2.8.6. <u>Let $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3\}$ and $\{\underset{\sim}{e}^1, \underset{\sim}{e}^2, \underset{\sim}{e}^3\}$ be a basis and its dual for $\mathcal{V}_3$. Then</u>

$$\underset{\sim}{e}^1 = \frac{\underset{\sim}{e}_2 \times \underset{\sim}{e}_3}{\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3)} , \quad \underset{\sim}{e}^2 = \frac{\underset{\sim}{e}_3 \times \underset{\sim}{e}_1}{\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3)} , \quad \underset{\sim}{e}^3 = \frac{\underset{\sim}{e}_1 \times \underset{\sim}{e}_2}{\Delta(\underset{\sim}{e}_1, \underset{\sim}{e}_2, \underset{\sim}{e}_3)} .$$

Supplementary Reading

Same as for § A2.7