Vectors: Additional Topics from a New Perspective

Matt McCutchen

March 23, 2003 – February 6, 2005

Contents

Preface

This document is about vectors. They are the arrows that mathematicians use to represent motion and numerous other things that have to do with a coordinate space. A typical math course on vectors defines vectors themselves, sums, scalar multiples, and magnitudes. One uses dot and cross products and investigates their applications. From there, it's time to go on to the more exciting world of vector-valued functions and, eventually, vector fields with their "div, grad, curl". However, there are many more topics in vector analysis than most courses or textbooks cover.

I have two goals in writing this document. The first is to treat the basics of vectors in a more formal, general, and powerful way. The second is to present several vector operations that I defined and investigated myself. These include a generalization of the cross product and a new operation I call the "hash product". Sections 1 through 3 address the first goal, while the remaining sections address the second. As I stumbled upon more and more interesting ideas related to vectors, I added them to the document. Thus, it has become a diverse compendium of definitions, theorems, and scattered thoughts with the unifying theme of vectors and their properties.

This document cannot take the place of a standard course (or book) on vectors. I leave out many important topics that are covered well in such a course. In addition, my coverage of vector basics in the first three sections is more difficult than the one in a typical course. Therefore, this paper is aimed principally at those who have already had some experience with vectors and would like to come back for more.

Rigor has its place (and plenty will be found in the first part), but I greatly prefer intuition as a way of learning mathematics. If one can gain an intuitive understanding of vectors and what the different operations mean, many of the results here will seem obvious.

Conventions of This Document

This document has two levels of sections: those such as 2 ("sections") and those such as 2.3 ("subsections").

This document uses one-based counting because most mathematicians are (unfortunately) accustomed to it.

Statements that are to be taken as true (definitions, theorems, etc.) are numbered in the format *$x.y$, where x is the major section number and y is the statement index, which starts from 1 at the beginning of every major section. Statements are referenced simply as *$x.y$, not Theorem *$x.y$ or Definition *$x.y$.

One reason for this convention is that whether a statement is a definition or theorem may depend on one's interpretation of the connection between vectors and geometry (see Section 1.3). Some statements contain substatements in numbered or lettered lists. Substatements are referenced by their number and/or letter within their statement. Elsewhere, the reference is preceded by the statement number, e.g. *2.1.1a. The symbol $\Box$ concludes a statement, while $\triangle$ concludes a proof.

A symbol $[x]$ refers to reference number x; see the References section at the end.

Later on (beginning in *1.1), we will see exactly what a "vector" is. But for now, it should be said that bold letters like $\mathbf{a}$ denote vector variables or functions that return vectors.

Acknowledgements

I would like to thank the faculty of the magnet math department at Montgomery Blair High School, especially Mr. Eric Walstein, and Dr. Brian Hunt of the University of Maryland for teaching me the basic ideas I needed in order to work on a project such as this one.

I would also like to thank my parents, Rosemary and Rick A. McCutchen, for their help in proofreading, guidance, and support.

For those readers who did not surmise as much after seeing the font, this document was typeset with LaTeX. See `http://www.latex-project.org/` for more information.

1 Mathematical Preliminaries

1.1 List-vectors

***1.1** *Definition: List-vector.* Let F be a field (see *1.12 in Section 1.4). A **list-vector** over F of dimension n is an ordered list of n elements of F (called its **components**). $\mathbb{LV}_{F,n}$ denotes the set of list-vectors over F of dimension n. A list-vector can be written as a list of its components in angle brackets; for example, $\langle 1, 2, 3 \rangle$ represents the list-vector over $\mathbb{R}$ of dimension 3 whose components are 1, 2, and 3.

If $(expr)$ is an expression that returns a vector, then $(expr)_i$ denotes its ith component. However, if $\mathbf{a}$ is a vector variable, its ith component is written a_i and is considered to be a single variable in F.

Two list-vectors of the same dimension are considered equal if and only if each component of the first vector equals the corresponding component of the second. $\Box$

***1.2** *Definition: Real vector.* A **real vector** is a list-vector over $\mathbb{R}$, the set of real numbers. $\Box$

We will be concerned mainly with real vectors, but it is best to be general.

1.2 Builder notation

***1.3** *Builder notation.* Suppose a composite value of some type can be built up from a list of zero or more pieces of type T by enclosing the list in delimiters $[\![\ldots]\!]$. Then

$$[\![_{i=1}^{n} f(i)]\!] \,,$$

for a function $f : \{1 \ldots n\} \to T$ and a dummy integer variable i, denotes the composite value with n pieces whose ith piece is $f(i)$. That is, $[\![_{i=1}^{n} f(i)]\!] = [\![f(1), f(2), \ldots, f(n)]\!]$. $\Box$

This abstract definition needs some explanation. In this document, it will be used mainly for list-vectors. Below is a specialized version of *1.3 for use with list-vectors:

***1.4** *Vector-builder notation.*

$$\left\langle_{i=1}^{n} f(i)\right\rangle = \left\langle f(1), f(2), \ldots, f(n)\right\rangle .$$

□

Note the similarity between $\sum$ or $\prod$ notation and this use of $\langle \ldots \rangle$. An expression is evaluated for each value of a dummy variable, producing a list of numbers. $\sum$ and $\prod$ perform an operation on these numbers, but $\langle \ldots \rangle$ builds a list-vector containing them (hence the name "builder notation").

As defined above, the dummy variable in vector-builder notation must start at 1. If this restriction were not made, one could write an expression for a list-vector $\mathbf{a}$ in which a function is evaluated at 2 to get a_1, 3 to get a_2, and so on. If one wanted the ith component of this list-vector, one would have to remember to set the dummy variable to $i+1$; confusion would result. It is much better to link the dummy variable to the component number.

Vector-builder notation has two properties that should be obvious and will henceforth be used without explicit reference:

- $\left\langle_{i=1}^{n} f(i)_j\right\rangle = f(j)$. That is, to pick the jth component out of a list-vector generated by vector-builder notation, one need only evaluate the vector-builder function at j.
- $\mathbf{a} = \left\langle_i\, a_i\right\rangle$. That is, a list-vector equals the list-vector built from its components.

We will have occasion to use a definition similar to *1.3 for matrices:

***1.5** *Matrix-builder notation.*

$$\left[_{1\le i\le m, 1\le j\le n}\ a_{i,j}\right] = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

□

One could generalize *1.3 to composites containing multidimensional arrays, but this generalization is difficult to state readably (as I discovered by trying it) and we will not need it in this document.

1.3 Geometric interpretations and their role

A real vector of dimension n can be thought to represent a motion or displacement in a coordinate space with n dimensions; each component gives the signed distance to move along one coordinate axis.

A directed line segment indicates a real vector: the displacement needed to carry its initial point to its final point. If real vectors are considered to *be* directed line segments, one must be careful to stipulate that two real vectors are equal if and only if their directed line segments have the same length and direction, although the segments could be in different places. Or, one can state that a real vector is the set of all directed line segments having a certain length and direction.

A vector in physics often has a unit of measure attached to it. If this unit is not one of length, the interpretation of a real vector as a motion is inconvenient. For this reason, it is often said that a real vector consists of a magnitude (a nonnegative real number, perhaps with a unit of measure) and a direction in a coordinate space of some dimension.

Geometric interpretations like these should help the reader to develop intuition about real vectors. However, none of them is formally necessary for the algebraic vector analysis presented here. The reader may choose to consider vector analysis to exist inside ordinary geometry (in which case vector theorems are considered valid in geometry and vice versa), or she/he may think of it as a new development that is analogous in many ways to ordinary geometry but bears no formal relationship. Though this distinction is pedantic, I feel that it is important. A third, and perhaps even more interesting, approach is to develop geometry inside of vector analysis by defining points, lines, circles, etc. in terms of real vectors and proving Euclid's axioms of geometry.

1.4 Some abstract algebra

***1.6** *Definition: Group.* A **group** is a tuple

$$G = (S, +, 0, (-)),$$

where S is a set, $+ : S \times S \to S$[1] and $(-) : S \to S$ are operators[2], $0 \in S$[3], and the following statements hold:

1. *Associative property.* $(a + b) + c = a + (b + c)$ for all $a, b, c \in S$.
2. *Identity property.* $0 + a = a + 0 = a$ for all $a \in S$.
3. *Inverse property.* $a + (-a) = 0$ for all $a \in S$.

   G is said to be **Abelian** if and only if the following statement also holds:

4. *Commutative property.* $a + b = b + a$ for all $a, b \in S$.

   G can be used to mean S in a context where a set is required. □

***1.7** *Theorem.* $-(-a) = a$ for any a in some group. □

Proof:

$$\begin{aligned}
-(-a) &= -(-a) + 0 && \text{(identity property of } +) \\
&= -(-a) + (-a + a) && \text{(inverse property of } +) \\
&= (-(-a) + (-a)) + a && \text{(associativity of } +) \\
&= a + 0 && \text{(inverse property of } +\text{, applied to } -a) \\
&= a && \text{(identity property of } +)
\end{aligned}$$

△

***1.8** *Theorem: Cancellation.* Let a, b, and c be elements of a set S on which an associative operation $+$, having an identity 0, has been defined. If $a + b = 0$ and $c + a = 0$, then $b = c$. □

Proof:

$$\begin{aligned}
a + b = 0 &\Rightarrow c + (a + b) = c + 0 && \text{(add } c \text{ to both sides)} \\
&\Rightarrow (c + a) + b = c + 0 && \text{(associativity of } +) \\
&\Rightarrow 0 + b = c + 0 && \text{(because } c + a = 0) \\
&\Rightarrow b = c && \text{(identity property of } +\text{, twice)}
\end{aligned}$$

[1] If $A_1, A_2, \ldots, A_n$ are sets, then $A_1 \times A_2 \times \ldots \times A_n$ denotes the set of n-tuples whose ith component is in A_i. So the notation above just means that the $+$ operation works on a pair of elements of S and produces a single element of S.

[2] $(-)$ denotes the negation operation, as on TI graphing calculators. Similarly, $(^{-1})$ denotes the multiplicative inverse operation; other such symbols will appear later. When any such operator is actually applied to a value, the parentheses are dropped.

[3] I feel that important aspects of a group, such as its 0 value and $(-)$ operation, should be part of its specification (the tuple). We must then require that the 0 and $(-)$ that come with the group work as they should, but once this is done, we can safely speak of "the" identity and "the" inverse. Almost all other authors state that a group is a tuple $G = (S, +)$, *there exists* an element 0 that behaves as an identity, and for each element a *there exists* an element b such that $a + b = 0$. Either way, it's necessary to prove uniqueness, which in my case is the statement that any 0 or $(-)$ that works must equal the official one, and in others' case is the statement that any two working 0s or $(-)$s are equal. However, I prefer the concreteness afforded by my technique.

△

***1.9** *Theorem: Only one* 0. If $G = (S, +, 0, (-))$ is a group, $a \in S$, and $a + b = b$ for all $b \in S$, then $a = 0$. Similarly, if $b + a = b$ for all $b \in S$, then $a = 0$. (In other words, 0 is the *only* identity.) □

Proof: $a + 0 = a$ by the identity property of $+$, but we assumed $a + 0 = 0$, so $a = 0$. A similar argument works for the other case. △

***1.10** *Theorem: Only one* $(-a)$. Let a and b be elements of a group G. If $a + b = 0$ or $b + a = 0$, then $b = -a$. (In other words, $-a$ is the *only* inverse of a.) □

Proof: If $a + b = 0$: Apply *1.8 with $c = -a$. We know $a + b = 0$ and $-a + a = 0$, so $b = -a$.
If $b + a = 0$: Apply *1.8 with $b \to -a, c \to b$. We know $a + (-a) = 0$ and $b + a = 0$, so $b = -a$. △

***1.11** *Definition: Subtraction.* $a - b = a + (-b)$ for all a, b in some group. □

***1.12** *Definition: Field.* A **field** is a tuple

$$F = (S, +, 0, (-), \cdot, 1, (^{-1})),$$

where $(S, +, 0, (-))$ is an Abelian group, $\cdot : S \times S \to S$ and $(^{-1}) : (S - \{0\}) \to S$ are operators, and the following statements hold:

1. *Commutative property of* $\cdot$. $ab = ba$ for all $a, b \in S$.
2. *Associative property of* $\cdot$. $(ab)c = a(bc)$ for all $a, b, c \in S$.
3. *Identity property of* $\cdot$. $1a = a(1) = 1$ for all $a \in S$.
4. *Inverse property of* $\cdot$. $a(a^{-1}) = 0$ for all $a \in S$ $(a \neq 0)$.
5. *Right distributive property.* $a(b + c) = (ab) + (ac)$ for all $a, b, c \in S$.

F can be used to mean S in a context where a set is required. □

***1.13** *Theorem: Left distributive property.* $(a + b)c = (ac) + (bc)$ for all a, b, c in a field F. □

Proof:

$$\begin{aligned} (a+b)c &= c(a+b) && \text{(*1.12.1)} \\ &= (ca) + (cb) && \text{(*1.12.5)} \\ &= (ac) + (bc) && \text{(*1.12.1 twice)} \end{aligned}$$

△

***1.14** *Theorem.* $0a = 0$ for any a in some field. □

Proof:

$$\begin{aligned} 0a &= (0+0)a && \text{(identity property of } +) \\ &= (0a) + (0a) && \text{(distributive property)} \\ &\Rightarrow (0a) + (-(0a)) = ((0a) + (0a)) + (-(0a)) && \text{(add } -(0a) \text{ to both sides)} \\ &\Rightarrow (0a) + (-(0a)) = (0a) + ((0a) + (-(0a))) && \text{(associativity of } +) \\ &\Rightarrow 0 = (0a) + 0 && \text{(inverse property of } + \text{ applied twice to } 0a) \\ &\Rightarrow 0 = 0a && \text{(identity property of } +) \end{aligned}$$

△

***1.15** *Theorem.* $(-1)a = -a$ for any a in some field. □

Proof:

$$\begin{aligned}
(-1)a &= (-1)a + (a + (-a)) && \text{(inverse property of } +\text{)} \\
&= ((-1)a + a) + (-a) && \text{(associativity of } +\text{)} \\
&= ((-1)a + 1a) + (-a) && \text{(identity property of } \cdot\text{)} \\
&= (-1 + 1)a + (-a) && \text{(distributive property)} \\
&= 0a + (-a) && \text{(inverse property of } +\text{)} \\
&= 0 + (-a) && \text{(*1.14)} \\
&= -a && \text{(identity property of } +\text{)}
\end{aligned}$$

△

***1.16** *Theorem.* $(-a)b = a(-b) = -(ab)$ for any a and b in some field. □

Proof: By *1.15, $(-a)b = ((-1)a)b$, $a(-b) = a((-1)b)$, and $-(ab) = (-1)(ab)$. By *1.12.1 and *1.12.2, these three are all equal. △

***1.17** *Theorem: Only one multiplicative inverse.* Let a and b be elements of a field. If $a \cdot b = 1$, then $b = a^{-1}$. □

Proof: Since $a \cdot b \neq 0$, $a \neq 0$ by *1.14, so a^{-1} exists. Apply *1.8 with $+ \to \cdot, c = a^{-1}$. We know $a \cdot b = 1$ and $a^{-1} \cdot a = 1$ (*1.12.4 and *1.12.1), so $b = a^{-1}$. △

***1.18** *Definition: Division.* $a/b = a \cdot (b^{-1})$ for any a, b in some field. □

***1.19** *Definition: Ordered field.* An **ordered field** is a tuple

$$OF = (S, +, 0, (-), \cdot, 1, (^{-1}), <),$$

where $F = (S, +, 0, (-), \cdot, 1, (^{-1}))$ is a field, $<: S \times S \to \mathbb{B}$[4] is an operator, and the following properties hold:

1. *Irreflexive property of $<$.* $a \not< a$ for all $a \in S$.
2. *Transitive property of $<$.* For all $a, b, c \in S$, if $a < b$ and $b < c$, then $a < c$.
3. *Trichotonomy property.* Given $a, b \in S$ with $a \neq b$, $a < b$ or $b < a$.
4. $a < b \Rightarrow a + c < b + c$ for all $a, b, c \in S$.
5. $a < b \Rightarrow ac < bc$ for all $a, b, c \in S$ with $0 < c$.
6. $0 < 1$.

□

***1.20** *Strengthened trichotonomy.* Let a and b be elements of an ordered field. Then exactly one of the following statements is true: $a = b$, $a < b$, $b < a$. □

Proof: If $a = b$, then $a \not< b$ and $b \not< a$ by *1.19.1, and our statement is true.

[4] $\mathbb{B}$ is the set of the two Boolean values. Most authors say $\mathbb{B} = \{\text{true}, \text{false}\}$, which is unfortunate because all mathematical values should be nouns. *Principia Mathematica* uses $\{\text{truth}, \text{falsity}\}$; I advocate the much simpler $\{\text{yes}, \text{no}\}$.

Otherwise $a \neq b$. By *1.19.3, $a < b$ or $b < a$, so we need only show that these statements can't both be true. If they were, we'd have $a < b < a \Rightarrow a < a$ by *1.19.2, which contradicts *1.19.1. △

***1.21** *Relational operators.* In an ordered field, we say that $a > b \Leftrightarrow b < a$, $a \leq b \Leftrightarrow a < b \vee a = b$, and $a \geq b \Leftrightarrow a > b \vee a = b$. By *1.20, $a \leq b \Leftrightarrow a \not> b$ and $a \geq b \Leftrightarrow a \not< b$. □

***1.22** *Theorem: Negation and* $<$. Let a be an element of a field. If $a > 0$, then $-a < 0$, and if $a < 0$, then $-a > 0$. □

Proof: If $a > 0$, then $a + (-a) > 0 + (-a)$ by *1.19.4. By the identity and inverse properties of $+$, $-a < 0$. The argument is analogous for the case $a < 0$. △

***1.23** *Theorem.* Let a, b, c be elements of an ordered field with $c < 0$. Then $a < b \Rightarrow bc < ac$. □

Proof: By *1.22, $0 < -c$. By *1.19.5, $a(-c) < b(-c)$, and by *1.16, $-(ac) < -(bc)$. Add $(ac) + (bc)$ to both sides to get $bc < ac$. △

***1.24** *Theorem: Squares.* Let a be an element of a field. Then $a^2 \geq 0$ (of course, a^2 denotes aa), and $a^2 = 0$ if and only if $a = 0$. □

Proof: If $a = 0$, then $a^2 = 0$ by *1.14.

If $a \neq 0$, then by *1.19.3, either $a > 0$ or $a < 0$. If $a > 0$, then $a^2 > 0a$ by *1.19.5, so $a^2 > 0$ by *1.14. If $a < 0$, then $a^2 > 0a$ by *1.23, so $a^2 > 0$ by *1.14.

If $a^2 = 0$, then it cannot be the case that $a > 0$ or $a < 0$, because in both of those cases, $0 < a^2$. Therefore, $a = 0$ by *1.19.3. △

***1.25** *Theorem.* Let $a, b \in OF$ for an ordered field OF, and suppose that $a \geq 0$ and $b \geq 0$. Then $a + b \geq 0$, and $a + b = 0$ if and only if $a = 0$ and $b = 0$. □

Proof: Suppose at least one of a and b is 0; without loss of generality, let it be a. If $b = 0$ as well, then $a+b = 0$ and the theorem is true. Otherwise $b > 0$, we have $a + b = b > 0$, and the theorem is true.

If this case does not apply, then $a > 0$ and $b > 0$. *1.19.4 tells us $a + b > 0 + b = b$. *1.19.2 then gives $a + b > 0$, and the theorem is true. △

***1.26** *Theorem.* Let the a_i be elements of an ordered field OF for $i = 1, \ldots, n$, and suppose that $a_i \geq 0$ for all i. Then $\sum_i a_i \geq 0$, and furthermore, $\sum_i a_i = 0$ if and only if $a_i = 0$ for all i. □

Proof: We will use induction on n. Our base case, $n = 0$, is trivial: the sum of zero things is 0, which equals 0 if and only if all the things are zero (which is vacuously true).

Suppose the theorem holds for $n - 1$; we will prove it for n. We know $\sum_{i=1}^{n} a_i = (\sum_{i=1}^{n-1} a_i) + a_n$. By the inductive hypothesis, $\sum_{i=1}^{n-1} a_i \geq 0$. By *1.25, $\sum_{i=1}^{n} a_i \geq 0$. By the inductive hypothesis, $\sum_{i=1}^{n-1} a_i = 0$ if and only if $a_i = 0$ for all $i \in 1, \ldots, n-1$. Applying the other part of *1.25, $\sum_{i=1}^{n} a_i = 0$ if and only if $\sum_{i=1}^{n-1} a_i = 0$ and $a_n = 0$, that is, if and only if $a_i = 0$ for all i. This completes the proof. △

***1.27** *Definition and theorem: Square roots.* Let OF be an ordered field, and let $a \in OF$ with $a \geq 0$. Then there exists at most one $b \in OF$ such that $b \geq 0$ and $b^2 = a$.

$\sqrt{a}$ denotes b if such a b exists. (Thus $(\sqrt{})$ is an operator from some subset of $\mathbb{P}_0$ to $\mathbb{P}_0$, where $\mathbb{P}_0 = \{a \in OF : a \geq 0\}$.) If $\sqrt{a}$ exists for all $a \geq 0$, then OF **has square roots**. □

Proof: We must prove that b is unique. First we will show that $b_1 < b_2 \Rightarrow {b_1}^2 < {b_2}^2$ for all $b_1, b_2 \geq 0$. In the case $b_1 = 0$, we know that $0 < b_2 \Rightarrow 0 = 0b_2 < {b_2}^2$ from *1.19.5. Since ${b_1}^2 = 0$, ${b_1}^2 < {b_2}^2$ and the lemma is true. If $b_1 \neq 0$, then we have $b_1 < b_2$, which implies $b_1 b_1 < b_1 b_2$ and $b_1 b_2 < b_2 b_2$ (using *1.19.5 twice). Thus, ${b_1}^2 < {b_2}^2$ by *1.19.2, and the lemma is again true.

Now, suppose there were two unequal elements of OF whose squares are both a. Let b_1 be the smaller and b_2 be the larger; *1.19.3 ensures that we can compare them this way. By the lemma, ${b_1}^2 < {b_2}^2$, so $a < a$, contradicting *1.19.1. △

The rational numbers ($\mathbb{Q}$) and the real numbers are the two best-known ordered fields. There are many others that consist of real numbers expressible in a certain form, such as $\{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$.

2 Vector Spaces

List-vectors can be added, negated, and multiplied by real numbers, with five key properties. From these, much about list-vectors can be deduced. However, there are many other kinds of vectors that have these five properties; these include complex vectors, polynomials, and even quantum states. In the spirit of generality, we'll let a **vector space** be any set of elements ("honorary vectors", in the words of Dr. Hunt) that can be added and multiplied by numbers with the five key properties. We thus have two tasks: to investigate the properties of vector spaces in general and to prove that list-vectors form a vector space. That way, anything we prove about vector spaces will apply to list-vectors.

As we define operations on list-vectors and prove that they form a vector space, we will take full advantage of the formality afforded by vector-builder notation (*1.4) as opposed to ellipses. The substitutions $\mathbf{a} = \left\langle_i a_i \right\rangle$ and $\mathbf{b} = \left\langle_i b_i \right\rangle$ may make some of the definitions and theorems easier to understand. In fact, after these substitutions, the statements below look a lot like the $\sum$ identities (e.g. $\sum_i (a_i + b_i) = \sum_i a_i + \sum_i b_i$). Though both the $\sum$ identities and vector definitions tell one how to add two $\sum$s or vectors, there is an important difference. The $\sum$ identities are theorems of algebra, proved by induction on the number of terms involved and using corresponding properties of the $+$ operator. However, we are *defining* addition and scalar multiplication on list-vectors; these operations could be defined differently (even though that would break everything we like about list-vectors).

***2.1** *Definition: Vector space.* Let $F = (S, +, 0, (-), \cdot, 1, (^{-1}))$ be a field.

A **vector space over** F is a tuple

$$VS = (V, +, \mathbf{0}, (-), \cdot),$$

where

- V is a set,
- $+ : V \times V \to V$ is an operation,
- $\mathbf{0} \in V$,
- $(-) : V \to V$ is an operation,
- $\cdot : S \times V \to V$ is an operation,

and the following statements are true:

1. $(V, +, \mathbf{0}, (-))$ is an Abelian group. This means:
   (a) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ for all $\mathbf{a}, \mathbf{b} \in V$.
   (b) $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$.

(c) $\mathbf{a}+\mathbf{0}=\mathbf{a}$ for all $\mathbf{a}\in V$.

(d) $\mathbf{a}+(-\mathbf{a})=\mathbf{0}$ and $(-\mathbf{a})+\mathbf{a}=\mathbf{0}$[5] for all $\mathbf{a}\in V$.

2. $1\mathbf{a}=\mathbf{a}$ for all $\mathbf{a}\in V$.
3. $r(s\mathbf{a})=(rs)\mathbf{a}$ for all $r,s\in S$ and $\mathbf{a}\in V$.
4. $(r+s)\mathbf{a}=(r\mathbf{a})+(s\mathbf{a})$ for all $r,s\in S$ and $\mathbf{a}\in V$.
5. $r(\mathbf{a}+\mathbf{b})=(r\mathbf{a})+(r\mathbf{b})$ for all $r\in S$ and $\mathbf{a},\mathbf{b}\in V$.

Elements of the set V are called **vectors** over F. VS can be used to mean V in a context where a set is required. □

Our eventual goal is to prove that list-vectors fit the mold of a vector space. To achieve this, we will define $+$, $\mathbf{0}$, $(-)$, and $\cdot$ on list-vectors. Then we will prove that $(\mathbb{LV}_{F,n},+,\mathbf{0},(-),\cdot)$ is a vector space over F. For this task, the notation of *1.4 will be invaluable because it provides a way to write statements involving list-vectors of arbitrary dimension without the regrettable informality of ellipses.

(In the previous paragraph, as in the next few sections, "for all fields F and all nonnegative integers n" is understood in statements about list-vectors, and when a statement speaks of several vectors, it is understood that they all belong to any single vector space.)

I feel that it is more logical to discuss addition followed by scalar multiplication rather than to prove that $(\mathbb{LV}_{F,n},+,\mathbf{0},(-),\cdot)$ is a vector space and then investigate properties of all vector spaces. Therefore, the next two sections will interleave results about list-vectors with results about all vector spaces.

2.1 Vector addition

***2.2** *Definition: + for list-vectors.* $\mathbf{a}+\mathbf{b}=\left\langle_i\, a_i+b_i\right\rangle$ for any $\mathbf{a},\mathbf{b}\in\mathbb{LV}_{F,n}$. □

***2.3** *Definition: $\mathbf{0}$ for list-vectors.* In any $\mathbb{LV}_{F,n}$, $\mathbf{0}=\left\langle_i\, 0\right\rangle$. That is, $\mathbf{0}$ denotes the vector all of whose components are 0 (the additive identity of F). □

The following definition is specific to list-vectors. Many vector spaces have no equivalent of it.

***2.4** *Definition of one vector.* In any $\mathbb{LV}_{F,n}$, $\mathbf{1}_i=\left\langle_{j=1}^{n}\,\delta_{j,i}\right\rangle$. (Here $\delta_{j,i}$ is the Kronecker delta notation: $\delta_{x,y}$ is 1 if $x=y$ and 0 otherwise.) In other words, $\mathbf{1}_i$ is the vector containing a one for its ith component but zeros for all others; its dimension is determined by the context.[6] □

***2.5** *Definition: $(-)$ for list-vectors.* $-\mathbf{a}=\left\langle_i\, -a_i\right\rangle$ for any $\mathbf{a}\in\mathbb{LV}_{F,n}$. □

***2.6** *Theorem.* $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$ for any $\mathbf{a},\mathbf{b}\in\mathbb{LV}_{F,n}$. □

Proof:

$$\begin{aligned}\mathbf{a}+\mathbf{b}&=\left\langle_i\, a_i+b_i\right\rangle && \text{(*2.2)}\\ &=\left\langle_i\, b_i+a_i\right\rangle && \text{(commutativity of + in } F\text{)}\\ &=\mathbf{b}+\mathbf{a} && \text{(*2.2)}\end{aligned}$$

[5] A group must satisfy *2.1.1b, *2.1.1c, and *both* parts of *2.1.1d, though it need not satisfy *2.1.1a. In the special case of an Abelian group (one that also satisfies *2.1.1a), the second part of *2.1.1d is unnecessary because it can be deduced from the first and *2.1.1a.

[6] Most textbooks can get away with using a different bold letter for each such vector because they only consider 2 or 3 dimensions. [2] uses $\mathbf{e}_i$. I chose $\mathbf{1}_i$ because its meaning is impossible to miss.

△

***2.7** *Theorem.* $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$ for any $\mathbf{a},\mathbf{b},\mathbf{c}\in\mathbb{LV}_{F,n}$. □

Proof:

$$\begin{aligned}(\mathbf{a}+\mathbf{b})+\mathbf{c} &= \left\langle_i\, (\mathbf{a}+\mathbf{b})_i + c_i\right\rangle && \text{(*2.2)}\\ &= \left\langle_i\, (a_i+b_i)+c_i\right\rangle && \text{(*2.2)}\\ &= \left\langle_i\, a_i+(b_i+c_i)\right\rangle && \text{(associativity of + in } F\text{)}\\ &= \left\langle_i\, a_i+(\mathbf{b}+\mathbf{c})_i\right\rangle && \text{(*2.2)}\\ &= \mathbf{a}+(\mathbf{b}+\mathbf{c}) && \text{(*2.2)}\end{aligned}$$

△

***2.8** *Theorem.* $\mathbf{a}+\mathbf{0}=\mathbf{a}$ for any $\mathbf{a}\in\mathbb{LV}_{F,n}$. □

Proof:

$$\begin{aligned}\mathbf{a}+\mathbf{0} &= \left\langle_i\, a_i+\mathbf{0}_i\right\rangle && \text{(*2.2)}\\ &= \left\langle_i\, a_i+0\right\rangle && \text{(*2.3)}\\ &= \left\langle_i\, a_i\right\rangle && \text{(identity property of + in F)}\\ &= \mathbf{a}\end{aligned}$$

△

***2.9** *Theorem.* $\mathbf{a}+(-\mathbf{a})=\mathbf{0}$ for any vector $\mathbf{a}$. □

Proof:

$$\begin{aligned}\mathbf{a}+(-\mathbf{a}) &= \left\langle_i\, a_i+(-\mathbf{a})_i\right\rangle && \text{(*2.2)}\\ &= \left\langle_i\, a_i-a_i\right\rangle && \text{(*2.5)}\\ &= \left\langle_i\, 0\right\rangle && \text{(inverse property of + in F)}\\ &= \mathbf{0} && \text{(*2.3)}\end{aligned}$$

△

***2.10** *Theorem.* $(\mathbb{LV}_{F,n},+,\mathbf{0},(-))$ is an Abelian group, where $+$, $\mathbf{0}$, and $(-)$ are defined as in *2.2, *2.3, and *2.5. □

Proof: $(\mathbb{LV}_{F,n},+,\mathbf{0},(-))$ has the four properties of an Abelian group, as shown in *2.6, *2.7, *2.8, and *2.9. △

By virtue of *2.1.1, *1.7 and *1.10 apply to vector addition. *1.11 will also be used to give meaning to $\mathbf{a}-\mathbf{b}$.

2.2 The scalar product

("Scalar" is the traditional term for a single number; i.e., not a vector. Here, scalars are real numbers.)

***2.11** *Definition: $\cdot$ for list-vectors.* $r\mathbf{a}=\left\langle_i\, ra_i\right\rangle$ for any $r\in F$ and $\mathbf{a}\in\mathbb{LV}_{F,n}$. □

***2.12** *Theorem.* $1\mathbf{a}=\mathbf{a}$ for any $\mathbf{a}\in\mathbb{LV}_{F,n}$. □

Proof:

$$\begin{aligned}1\mathbf{a} &= \left\langle_i\, 1a_i\right\rangle && \text{(*2.11)}\\ &= \left\langle_i\, a_i\right\rangle && \text{(identity property of } \cdot \text{ in } F\text{)}\\ &= \mathbf{a}\end{aligned}$$

△

***2.13** *Theorem.* $r(s\mathbf{a}) = (rs)\mathbf{a}$ for any $r, s \in F$ and $\mathbf{a} \in \mathbb{LV}_{F,n}$. □

Proof:

$$\begin{aligned} r(s\mathbf{a}) &= r\left\langle_i\, sa_i\right\rangle && \text{(*2.11)} \\ &= \left\langle_j\, r\left\langle_i\, sa_i\right\rangle_j\right\rangle && \text{(*2.11)} \\ &= \left\langle_j\, r(sa_j)\right\rangle && \\ &= \left\langle_j\, (rs)a_j\right\rangle && \text{(associativity of $\cdot$ in F)} \\ &= (rs)\mathbf{a} && \text{(*2.11)} \end{aligned}$$

△

***2.14** *Theorem.* $r(s\mathbf{a}) = (rs)\mathbf{a} = s(r\mathbf{a})$ for any $r, s \in F$ and any vector $\mathbf{a}$ over F. □

Proof: The first part of the theorem is just *2.1.3. To get the second, swap r and s in *2.1.3 to get $(sr)\mathbf{a} = s(r\mathbf{a})$, and apply the commutativity of $\cdot$ in F to get $(rs)\mathbf{a} = s(r\mathbf{a})$. △

***2.15** *Theorem.* $(r + s)\mathbf{a} = r\mathbf{a} + s\mathbf{a}$ for any $r, s \in F$ and $\mathbf{a} \in \mathbb{LV}_{F,n}$. □

Proof:

$$\begin{aligned} (r+s)\mathbf{a} &= \left\langle_i\, (r+s)a_i\right\rangle && \text{(*2.11)} \\ &= \left\langle_i\, ra_i + sa_i\right\rangle && \text{(distributive property of $\cdot$ over $+$ in F)} \\ &= \left\langle_i\, ra_i\right\rangle + \left\langle_i\, sa_i\right\rangle && \text{(*2.2)} \\ &= r\mathbf{a} + s\mathbf{b} && \text{(*2.11 twice)} \end{aligned}$$

△

***2.16** *Theorem.* $r(\mathbf{a} + \mathbf{b}) = r\mathbf{a} + r\mathbf{b}$ for any $r \in F$ and $\mathbf{a}, \mathbf{b} \in \mathbb{LV}_{F,n}$. □

Proof:

$$\begin{aligned} r(\mathbf{a}+\mathbf{b}) &= r(\left\langle_i\, a_i + b_i\right\rangle) && \text{(*2.2)} \\ &= \left\langle_i\, r(a_i + b_i)\right\rangle && \text{(*2.11)} \\ &= \left\langle_i\, ra_i + rb_i\right\rangle && \text{(distributive property of $\cdot$ over $+$ in F)} \\ &= \left\langle_i\, ra_i\right\rangle + \left\langle_i\, rb_i\right\rangle && \text{(*2.2)} \\ &= r\mathbf{a} + r\mathbf{b} && \text{(*2.11 twice)} \end{aligned}$$

△

***2.17** *Theorem.* $0\mathbf{a} = \mathbf{0}$ for any vector $\mathbf{a}$. □

Proof:

$$\begin{aligned} 0\mathbf{a} &= 0\mathbf{a} + (0\mathbf{a} + (-(0\mathbf{a}))) && \text{(*2.1.1d, applied to $0\mathbf{a}$)} \\ &= (0\mathbf{a} + 0\mathbf{a}) + (-(0\mathbf{a})) && \text{(*2.1.1b)} \\ &= (0+0)\mathbf{a} + (-(0\mathbf{a})) && \text{(*2.1.4)} \\ &= 0\mathbf{a} + (-(0\mathbf{a})) && \text{(identity property of $+$ in F)} \\ &= \mathbf{0} && \text{(*2.1.1d)} \end{aligned}$$

△

***2.18** *Theorem.* $(-1)\mathbf{a} = -\mathbf{a}$ for any vector $\mathbf{a}$. □

Proof:

$$\begin{aligned}
(-1)\mathbf{a} &= (-1)\mathbf{a} + (\mathbf{a} + (-\mathbf{a})) && \text{(*2.1.1d)}\\
&= ((-1)\mathbf{a} + \mathbf{a}) + (-\mathbf{a}) && \text{(*2.1.1b)}\\
&= ((-1)\mathbf{a} + 1\mathbf{a}) + (-\mathbf{a}) && \text{(*2.1.2)}\\
&= (-1+1)\mathbf{a} + (-\mathbf{a}) && \text{(*2.1.4)}\\
&= 0\mathbf{a} + (-\mathbf{a}) && \text{(inverse property of } + \text{ in } F)\\
&= \mathbf{0} + (-\mathbf{a}) && \text{(*2.17)}\\
&= -\mathbf{a} && \text{(*2.1.1c)}
\end{aligned}$$

△

This is a notational convenience very similar to *1.11:

***2.19** *Definition: Scalar quotient.* $\mathbf{a}/r = (1/r)\mathbf{a}$ for any $r \in F$ $(r \neq 0)$ and vector $\mathbf{a}$. □

Here is the theorem the reader has probably been waiting for:

***2.20** *Theorem.* $(\mathbb{LV}_{F,n}, +, \mathbf{0}, (-), \cdot)$ is a vector space over F, where $+$, $\mathbf{0}$, $(-)$, and $\cdot$ are as defined in *2.2, *2.3, *2.5, and *2.11. We will use $\mathbb{LV}_{F,n}$ to refer to this vector space as well as the set of vectors. □

Proof: $(\mathbb{LV}_{F,n}, +, \mathbf{0}, (-), \cdot)$ has the five properties of a vector space listed in *2.1:

- *2.10 proves *2.1.1.
- *2.12 proves *2.1.2.
- *2.13 proves *2.1.3.
- *2.15 proves *2.1.4.
- *2.16 proves *2.1.5.

△

***2.21** *Definition: Same direction.* Two vectors $\mathbf{a}$ and $\mathbf{b}$ over a field F, neither of which is $\mathbf{0}$, are said to be **parallel** if there exists an $r \in F$ such that $\mathbf{a} = r\mathbf{b}$. The order of the vectors is unimportant because, if $\mathbf{a} = r\mathbf{b}$, then $\mathbf{b} = (1/r)\mathbf{a}$; $r \neq 0$ because this would mean $\mathbf{a} = 0$. □

***2.22** *Theorem.* In any vector space, parallelism is an equivalence relation on the set of vectors excluding $\mathbf{0}$. (That is, it is reflexive, symmetric, and transitive.) □

Proof: Parallelism is reflexive because we can let $r = 1$, and $\mathbf{a} = 1\mathbf{a}$ by *2.1.2.

Parallelism is symmetric because, if $\mathbf{a} = r\mathbf{b}$, then $\mathbf{b} = (1/r)\mathbf{a}$ by *2.1.3 and *1.12.4; since $\mathbf{a} \neq \mathbf{0}$, $r \neq 0$.

Parallelism is transitive because, if $\mathbf{a} = r\mathbf{b}$ and $\mathbf{b} = s\mathbf{c}$, then $\mathbf{a} = (rs)\mathbf{c}$ by *2.1.3. △

***2.23** *Definition: Same and opposite direction.* Two vectors $\mathbf{a}$ and $\mathbf{b}$, neither of which is $\mathbf{0}$, are said to have the **same direction** if they are parallel with $r > 0$ and **opposite directions** if they are parallel with $r < 0$. □

3 The Dot Product

Some vector spaces have, in addition to the operations $+$, $(-)$, and $\cdot$, another kind of product: the **dot product**. This operation takes two vectors and returns an element of the underlying field. We will take the same approach to the dot product as we did to vector addition and multiplication by scalars: we will define a **dot product space** as a vector space with a dot product operator having certain desirable properties.

On list-vectors, the definition of dot product is very simple: the sum of the products of corresponding elements. (We'll get back to this and treat it more formally.) The dot product of a list-vector with itself is the sum of the squares of its elements, so it gives some measure of the "size" of the vector. This is well-behaved for real vectors. However, if we take the size of the complex vector $\langle i \rangle$, we get -1, which is very bad. We would, therefore, like to define the dot product in a sensible manner for complex numbers. The operation of **conjugation** will be the key, and its development requires much abstract algebra. The reader might wish to review Section 1.4 before proceeding.

3.1 Conjugation

***3.1** *Definition: Conjugation field.* A **conjugation field** is a tuple

$$CF = (S, +, 0, (-), \cdot, 1, (^{-1}), (^*)),$$

where $F = (S, +, 0, (-), \cdot, 1, (^{-1}))$ is a field, $(^*) : S \to S$ is an operator (called **conjugation**), and the following is true:

1. $(a^*)^* = a$ for all $a \in S$.
2. $a^* + b^* = (a + b)^*$ for all $a, b \in S$.
3. $a^* \cdot b^* = (a \cdot b)^*$ for all $a, b \in S$.

CF can be used to mean F or S depending on the context. □

Some elements of a conjugation field are their own conjugates. We call these **core elements**. It happens that 0 and 1 are core elements, and so is the result of any of the operations $+$, $(-)$, $\cdot$, $(^{-1})$ on core elements.

***3.2** *Definition: Core element.* A **core element** of a conjugation field CF is an $a \in CF$ such that $a^* = a$. □

***3.3** *Theorem.* The elements 0 and 1 of any conjugation field CF are core elements of CF. □

Proof: Let $a = 0$ and $b = 0^*$ in *3.1.2, which becomes $0^* + (0^*)^* = (0 + 0^*)^*$. By the identity property of $+$ in CF, $0^* + (0^*)^* = (0^*)^*$. Applying *3.1.1 twice, we have $0^* + 0 = 0$. By the identity property of $+$ in CF, $0^* = 0$, so 0 is a core element of CF. An analogous argument shows that $1^* = 1$: simply use 1 instead of 0, *3.1.3 instead of *3.1.2, and the identity property of $\cdot$ instead of $+$. △

***3.4** *Theorem.* Let a and b be core elements of some conjugation field CF. Then $a + b$ and $a \cdot b$ are also core elements of CF. □

Proof: We have $a^* = a$ and $b^* = b$. By *3.1.2, $(a + b)^* = a^* + b^* = a + b$. Since $(a + b)^* = a + b$, $a + b$ is a core element of CF. An analogous argument shows that $a \cdot b$ is a core element of CF. △

***3.5** *Theorem.* $-(a^*) = (-a)^*$ for any a in some conjugation field. □

Proof: Let $b = -a$ in *3.1.2: $a^* + (-a)^* = (a + (-a))^*$. By the inverse property of $+$, $a^* + (-a)^* = 0^*$. By *3.3, $0^* = 0$, so $a^* + (-a)^* = 0$. By *1.10 with $a \to a^*$ and $b \to (-a)^*$, $(-a)^* = -(a^*)$. △

***3.6** *Theorem.* $(a^*)^{-1} = (a^{-1})^*$ for any a in some conjugation field ($a \neq 0$). □

Proof: Let $b = a^{-1}$ in *3.1.3: $(a^*)((a^{-1})^*) = (a + (-a))^*$. By the inverse property of $\cdot$, $(a^*)((a^{-1})^*) = 1^*$. By *3.3, $1^* = 1$, so $(a^*)((a^{-1})^*) = 1$. By *1.17 with $a \to a^*$ and $b \to (-a)^*$, $(-a)^* = -(a^*)$. △

***3.7** *Theorem.* If a is a core element of a conjugation field CF, then $-a$ is a core element of CF. Furthermore, if $a \neq 0$, then a^{-1} is a core element of CF. □

Proof: We have $a^* = a$. Substituting this in *3.5 gives $-a = (-a)^*$, so $-a$ is a core element of CF. If $a \neq 0$, then substituting $a^* = a$ gives $a^{-1} = (a^{-1})^*$, so a^{-1} is a core element of CF. △

These theorems lead to the interesting observation that a conjugation field's set of core elements forms another field:

***3.8** *Theorem.* Let $CF = (S, +, 0, (-), \cdot, 1, (^{-1}), (^*))$ be a conjugation field. Then $(S', +, 0, (-), \cdot, 1, (^{-1}))$ is a field, where S' is the set of core elements of CF. This field will be denoted $\operatorname{core} CF$. □

Proof: By *3.4 and *3.7, $+$, $(-)$, $\cdot$, and $(^{-1})$ produce results in S' when all inputs are in S'. Because of this, it makes sense to think of $+$, $(-)$, $\cdot$, and $(^{-1})$ as defined on S'. In addition, $0 \in S'$ and $1 \in S'$ by *3.3. The field properties of $\operatorname{core} CF$ are special cases of those of CF, so we are done. △

We now know that any conjugation field has another field as its core. We are most interested in cases where the core is an ordered field (*1.19).

***3.9** *Definition: Ordered-core conjugation field.* An **ordered-core conjugation field** is a tuple

$$OCCF = (S, +, 0, (-), \cdot, 1, (^{-1}), (^*), <),$$

where:

1. $CF = (S, +, 0, (-), \cdot, 1, (^{-1}), (^*))$ is a conjugation field.
2. $<: \operatorname{core} CF \times \operatorname{core} CF$ is an operator.
3. $\operatorname{core} CF$ with $<$ is an ordered field.
4. For all $a \in S$ the following is true: $aa^* \in \operatorname{core} CF$ and $aa^* \geq 0$.

$OCCF$ can be used to mean CF in a context where an ordinary conjugation field is required. □

***3.10** *Definition: Size.* $\|a\| = aa^*$ for all a in an ordered-core conjugation field $OCCF$. By *3.9.4, this is a nonnegative element of $\operatorname{core} OCCF$. □

***3.11** *Theorem.* $\|ab\| = \|a\|\|b\|$ for all $a, b \in OCCF$. □

Proof:

$$\begin{aligned}
\|ab\| &= (ab)(ab)^* && \text{(*3.10)} \\
&= aba^*b^* && \text{(*3.1.3)} \\
&= (aa^*)(bb^*) && \text{(commutative and associative properties)} \\
&= \|a\|\|b\| && \text{(*3.10 twice)}
\end{aligned}$$

△

Though its meaning is rather difficult to discern, the theorem below will yield the Triangle Inequality (*3.15). Its proof was inspired by the proof of Schwarz's Inequality in [2].

***3.12** *Triangle inequality lemma.* $4\|a\|\|b\| \geq (ab^* + a^*b)^2$ for all $a, b \in OCCF$. □

Proof: If $a = 0$ or $b = 0$, then the equation above is $0 \geq 0$, which is true. Thus we may assume both a and b are nonzero.

Let $t \in \operatorname{core} OCCF$; we will choose a value for t later. Consider the size of $a + tb$:

$$\begin{aligned}
\|a+tb\| &= (a+tb)(a+tb)^* && \text{(*3.10)} \\
&= (a+tb)(a^*+tb^*) && \text{(*3.1.2, *3.1.3, } t^*=t\text{)} \\
&= aa^* + a(tb^*) + (tb)a^* + (tb)(tb^*) && \text{(distributive property)} \\
&= \|a\| + t(ab^*+ba^*) + t^2\|b\| && \text{(*3.10)}
\end{aligned}$$

If we let $t = 1$ in this equation, then we get $ab^* + ba^* = \|a+tb\| - \|a\| - \|b\|$. The right side of this equation is in $\operatorname{core} OCCF$, so $ab^* + ba^* \in \operatorname{core} OCCF$. Let $k = ab^* + ba^*$ for convenience.

We actually want $t = -k/(2\|b\|)$. (We needed to know $k \in \operatorname{core} OCCF$ before we could make this assignment. In addition, $b \neq 0$, so $\|b\| \neq 0$ by *3.9.4.) When we put this in the equation above we get:

$$\begin{aligned}
\|a+tb\| &= \|a\| + t(ab^*+ba^*) + t^2\|b\| \\
&= \|a\| + tk + t^2\|b\| \\
&= \|a\| + (-k/(2\|b\|))k + (k^2/(4\|b\|^2))\|b\| \\
&= \|a\| - k^2/(2\|b\|) + k^2/(4\|b\|) \\
&= \|a\| - k^2/(4\|b\|)
\end{aligned}$$

Here comes the "punch line", as the Hubbards would say in [2]. $\|a+tb\| \geq 0$ by *3.9.4. Thus $\|a\| - k^2/(4\|b\|) \geq 0$, and $4\|a\|\|b\| \geq k^2 = (ab^*+ba^*)^2$. △

***3.13** *Definition: Absolute value.* Let a be an element of an ordered-core conjugation field $OCCF$. $|a| = \sqrt{\|a\|}$ if this expression is defined. (If $\operatorname{core} OCCF$ has square roots, $|a|$ is defined for all a.) □

Square roots allow us to define absolute value in general. Actually, in ordered fields (or simple ordered-core conjugation fields), we can define absolute value without them: we can simply set $|x| = x$ if $x \geq 0$ or $-x$ if $x \leq 0$, which is equivalent to the definition above.

***3.14** *Theorem.* $|ab| = |a||b|$ for all $a, b \in OCCF$ for which both sides of the equation are defined. □

Proof: By *3.11, $\|ab\| = \|a\|\|b\| = |a|^2|b|^2 = (|a||b|)^2$. $|a||b|$ is a square root of $\|ab\|$, so it is the only one, and $|ab| = |a||b|$. △

***3.15** *Theorem: Triangle inequality.* $|a+b| \leq |a| + |b|$ for all $a, b \in OCCF$ for which both sides of the inequality are defined. □

Proof: Assume for the sake of contradiction that $|a+b| > |a| + |b|$. We can square both sides and the inequality will remain true; see the proof of *1.27. Thus $\|a+b\| > \|a\| + \|b\| + 2|a||b|$. Rearrange this and use the definition of size (*3.10) to get $(a+b)(a+b)^* - aa^* - bb^* > 2|a||b|$. Expand the leftmost term and cancel to get $ab^* + a^*b > 2|a||b|$. Now square both sides. The result is $(ab^* + a^*b)^2 > 4\|a\|\|b\|$, which contradicts *3.12. △

The complex numbers, of course, form an ordered-core conjugation field, using the rule $(a+bi)^* = a - bi$. (The reader may easily verify the statements in *3.1 and *3.9.) Their core is the ordered field of real numbers.

But the real numbers also form an ordered-core conjugation field if each element is made its own conjugate. We call this kind of conjugation field **simple**:

***3.16** *Definition: Simple conjugation field.* A conjugation field CF is said to be **simple** if and only if, for all $a \in CF$, $a^* = a$ (in which case $\operatorname{core} CF$ is CF itself). □

The following theorem states that any (ordered) field can be made into a simple (ordered-core) conjugation field by specifying that each element is its own conjugate.

***3.17** *Theorem: Making simple conjugation fields.* Any field $F = (S, +, 0, (-), \cdot, 1, (^{-1}))$ can be made into a simple conjugation field $CF = (S, +, 0, (-), \cdot, 1, (^{-1}), (^*))$ by letting $a^* = a$ for all $a \in S$.

Similarly, any ordered field $OF = (S, +, 0, (-), \cdot, 1, (^{-1}), <)$ can be made into an ordered-core conjugation field $OCCF = (S, +, 0, (-), \cdot, 1, (^{-1}), <, (^*))$ by letting $a^* = a$ for all $a \in S$.

If an (ordered) field is used in a context where an (ordered-core) conjugation field is required, this construction is to be used. □

Proof: We must verify the properties of a conjugation field as stated in *3.1. We have assumed that F is a field. Because $a^* = a$ for all $a \in S$, *3.1.1, *3.1.2, and *3.1.3 become trivial. This shows that CF is a conjugation field.

For the second part, *3.9.1 follows from the first part, *3.9.3 is assumed, and *1.24 takes care of *3.9.4. △

There is one remaining conjugation-related concept to present: that of a conjugation vector space. Essentially, a conjugation vector space must satisfy the three properties in *3.1, but for vector operations.

***3.18** *Definition: Conjugation vector space.* Let $F = (S, +, 0, (-), \cdot, 1, (^{-1}))$ be a field.

A **conjugation vector space over** F is a tuple $VS = (V, +, \mathbf{0}, (-), \cdot, (^*))$, where $VS = (V, +, \mathbf{0}, (-), \cdot)$ is a vector space,

- V is a set,
- $+ : V \times V \to V$ is an operation,
- $\mathbf{0} \in V$,
- $(-) : V \to V$ is an operation, and
- $\cdot : S \times V \to V$ is an operation,

and the following is true:

1. $(\mathbf{a}^*)^* = \mathbf{a}$ for all $a \in V$.
2. $\mathbf{a}^* + \mathbf{b}^* = (\mathbf{a} + \mathbf{b})^*$ for all $a, b \in V$.
3. $r^*\mathbf{a}^* = (r\mathbf{a})^*$ for all $r \in S$ and $a \in V$.

□

We could go on and adapt the concept of a core for conjugation vector spaces, but we won't need this. However, it is useful to specify the meaning of a simple conjugation vector space:

***3.19** *Definition: Simple conjugation vector space.* A conjugation vector space CVS over a conjugation field CF is said to be **simple** if and only if CF is simple and $\mathbf{a} = \mathbf{a}^*$ for all vectors $\mathbf{a} \in CVS$. □

Now we will show that list-vectors over an arbitrary conjugation field form a conjugation vector space.

***3.20** *Definition: Conjugation for list-vectors.* $\mathbf{a}^* = \left\langle_i a_i{}^* \right\rangle$ for all $\mathbf{a} \in \mathbb{LV}_{CF,n}$, for all conjugation fields CF and for all nonnegative integers n. □

***3.21** *Theorem.* $(\mathbf{a}^*)^* = \mathbf{a}$ for all $a \in \mathbb{LVCVS}_{CF,n}$. □

Proof:

$$\begin{aligned} (\mathbf{a}^*)^* &= \left\langle_i (a_i{}^*)^* \right\rangle && \text{(*3.20 twice)} \\ &= \left\langle_i a_i \right\rangle && \text{(*3.1.1)} \\ &= \mathbf{a} \end{aligned}$$

△

***3.22** *Theorem.* $\mathbf{a}^* + \mathbf{b}^* = (\mathbf{a}+\mathbf{b})^*$ for all $a, b \in \mathbb{LV}_{CF,n}$. □

Proof:

$$\begin{aligned}
\mathbf{a}^* + \mathbf{b}^* &= \left\langle_i a_i{}^* \right\rangle + \left\langle_i b_i{}^* \right\rangle && (*3.20)\\
&= \left\langle_i a_i{}^* + b_i{}^* \right\rangle && (*2.2)\\
&= \left\langle_i (a_i + b_i)^* \right\rangle && (*3.1.2)\\
&= \left\langle_i a_i + b_i \right\rangle^* && (*3.20)\\
&= (\mathbf{a}+\mathbf{b})^* && (*2.2)
\end{aligned}$$

△

***3.23** *Theorem.* $r^*\mathbf{a}^* = (r\mathbf{a})^*$ for all $r \in CF$ and $a \in \mathbb{LV}_{CF,n}$. □

Proof:

$$\begin{aligned}
r^*\mathbf{a}^* &= r^* \left\langle_i a_i{}^* \right\rangle && (*3.20)\\
&= \left\langle_i r^* a_i{}^* \right\rangle && (*2.11)\\
&= \left\langle_i (r a_i)^* \right\rangle && (*3.1.3)\\
&= \left\langle_i r a_i \right\rangle^* && (*3.20)\\
&= (r\mathbf{a})^* && (*2.11)
\end{aligned}$$

△

***3.24** *Theorem.* $(\mathbb{LV}_{CF,n}, +, \mathbf{0}, (-), \cdot, (^*))$ is a conjugation vector space over CF, where $+$, $\mathbf{0}$, $(-)$, $\cdot$, and $(^*)$ are as defined in *2.2, *2.3, *2.5, *2.11, and *3.20. $\mathbb{LV}_{CF,n}$ will denote this conjugation vector space as well as its set of vectors. □

Proof: We already know that $(\mathbb{LV}_{CF,n}, +, \mathbf{0}, (-), \cdot)$ is a vector space by *2.20. *3.21, *3.22, and *3.23 satisfy the other requirements in *3.18. △

There are many other interesting topics related to conjugation fields and their properties. Chief among these is the **complex construction**. Given a set S with operations $+, (-), \cdot, (^{-1}), (^*)$, let $\|x\| = xx^*$ for all $x \in S$, and let $S' = \{⦅x, y⦆ : x, y \in S\}$. Let's define these same five operations on elements of S':

- $⦅x_1, y_1⦆ + ⦅x_2, y_2⦆ = ⦅x_1 + x_2, y_1 + y_2⦆$
- $-⦅x, y⦆ = ⦅-x, -y⦆$
- $⦅x_1, y_1⦆ \cdot ⦅x_2, y_2⦆ = ⦅x_1x_2 - y_1(y_2{}^*), x_1y_2 + x_2(y_1{}^*)⦆$
- $⦅x, y⦆^{-1} = ⦅x, y⦆^*/\|⦅x, y⦆\|$
- $⦅x, y⦆^* = ⦅x^*, -y⦆$

Suppose we start with S being an arbitrary ordered field OF (taken as a simple ordered-core conjugation field: $x^* = x$ for all $x \in OF$). Then, after applying this procedure once, we get an ordered-core conjugation field that is not simple. One can continue to recursively apply this process, and the results will always be vector spaces over OF. However, the resulting systems quickly lose most of the other algebraic properties. For example, multiplication in general becomes noncommutative and then nonassociative. Applying this procedure repeatedly to the real numbers, in particular, yields the complex numbers, the quaternions, the octonions, and even crazier systems. I will not prove any of these claims. The curious reader should consult [3] for much more information.

3.2 Dot product spaces

We are finally ready to define the dot product through the idea of a **dot product space**.

***3.25** *Definition: Dot product space.* Let $OCCF = (S, +, 0, (-), \cdot, 1, (^{-1}), (^*), <)$ be an ordered-core conjugation field.

A **dot product space over** $OCCF$ is a tuple

$$DPS = (V, +, \mathbf{0}, (-), \cdot_{\text{scalar}}, (^*), \cdot_{\text{dot}}),$$

where $CVS = (V, +, \mathbf{0}, (-), \cdot_{\text{scalar}}, (^*))$ is a conjugation vector space over $OCCF$, $\cdot_{\text{dot}} : V \times V \to OCCF$ is an operator (called the **dot product**), and the following statements hold:

1. $\mathbf{a}^* \cdot \mathbf{b}^* = (\mathbf{a} \cdot \mathbf{b})^* = \mathbf{b} \cdot \mathbf{a}$ for all $\mathbf{a}, \mathbf{b} \in V$.
2. $(r\mathbf{a}) \cdot \mathbf{b} = r(\mathbf{a} \cdot \mathbf{b})$ for all $r \in OCCF$ and $\mathbf{a}, \mathbf{b} \in V$.
3. $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = (\mathbf{a} \cdot \mathbf{c}) + (\mathbf{b} \cdot \mathbf{c})$ for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$.
4. $\mathbf{a} \cdot \mathbf{a} \in \operatorname{core} OCCF$ and $\mathbf{a} \cdot \mathbf{a} \geq 0$ for all $\mathbf{a} \in V$, and $\mathbf{a} \cdot \mathbf{a} = 0$ if and only if $\mathbf{a} = \mathbf{0}$.

DPS can be used to mean CVS in a context where a conjugation vector space is required. □

It is important to note that in a simple dot product space we can throw out all the * symbols. In this case, *3.25.1 reduces to a statement that *the dot product is commutative*. Most of the time one works with simple dot product spaces, but I felt that the complex-number version of the dot product merited generalization of the entire framework of the dot product.

Our goal now is to show that list-vectors form a dot product space. However, this time we cannot use list-vectors over an arbitrary field: the field must be an ordered-core conjugation field (or an ordered field, in which case we use the construction in *3.17).

***3.26** *Definition:* $\cdot_{\text{dot}}$ *for list-vectors.* $\mathbf{a} \cdot \mathbf{b} = \sum_i (a_i)(b_i{}^*)$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{LV}_{OCCF,n}$. □

***3.27** *Theorem.* $\mathbf{a}^* \cdot \mathbf{b}^* = (\mathbf{a} \cdot \mathbf{b})^* = \mathbf{b} \cdot \mathbf{a}$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{LV}_{OCCF,n}$. □

Proof: The following proves that the first quantity equals the second:

$$\begin{aligned}
\mathbf{a}^* \cdot \mathbf{b}^* &= \textstyle\sum_i (\mathbf{a}^*)_i (\mathbf{b}^*)_i{}^* && \text{(*3.26)} \\
&= \textstyle\sum_i a_i{}^* (b_i{}^*)^* && \text{(*3.20 twice)} \quad (*) \\
&= \textstyle\sum_i (a_i b_i{}^*)^* && \text{(*3.1.3)} \\
&= (\textstyle\sum_i a_i b_i{}^*)^* && \text{(repeated use of *3.1.2}^7\text{)} \\
&= (\mathbf{a} \cdot \mathbf{b})^* && \text{(*3.26)}
\end{aligned}$$

We will resume from $(*)$ to prove that the first quantity equals the third:

$$\begin{aligned}
\textstyle\sum_i a_i{}^* (b_i{}^*)^* &= \textstyle\sum_i a_i{}^* b_i && \text{(*3.1.1)} \\
&= \textstyle\sum_i b_i a_i{}^* && \text{(*1.12.1)} \\
&= \mathbf{b} \cdot \mathbf{a} && \text{(*3.26)}
\end{aligned}$$

△

[7] In general, "repeated use" indicates that a theorem about a binary operator (such as $+$) is to be used to prove the corresponding theorem for the corresponding list operator (such as $\sum$) by induction, and the latter theorem is to be applied. For example, see *1.25 and *1.26. In this case, we would use $a^* + b^* = (a + b)^*$ to prove that $\sum_i x^* = (\sum_i x)^*$.

***3.28** *Theorem.* $(r\mathbf{a})\cdot\mathbf{b} = r(\mathbf{a}\cdot\mathbf{b})$ for all $r \in OCCF$ and $\mathbf{a},\mathbf{b} \in \mathbb{LV}_{OCCF,n}$. □

Proof:

$$\begin{aligned}(r\mathbf{a})\cdot\mathbf{b} &= \textstyle\sum_i (r\mathbf{a})_i b_i{}^* && \text{(*3.26)}\\ &= \textstyle\sum_i (ra_i) b_i{}^* && \text{(*2.11)}\\ &= \textstyle\sum_i r(a_i b_i{}^*) && \text{(associativity of $\cdot$ in $OCCF$)}\\ &= r\textstyle\sum_i a_i b_i{}^* && \text{(repeated use of distributive property in $OCCF$)}\\ &= r(\mathbf{a}\cdot\mathbf{b}) && \text{(*3.26)}\end{aligned}$$

△

***3.29** *Theorem.* $\mathbf{a}\cdot(r\mathbf{b}) = r^*(\mathbf{a}\cdot\mathbf{b})$ for all $a, b \in DPS$. □

Proof:

$$\begin{aligned}\mathbf{a}\cdot(r\mathbf{b}) &= ((r\mathbf{b})\cdot\mathbf{a})^* && \text{(*3.25.1)}\\ &= (r(\mathbf{b}\cdot\mathbf{a}))^* && \text{(*3.25.2)}\\ &= r^*(\mathbf{b}\cdot\mathbf{a})^* && \text{(*3.1.3)}\\ &= r^*(\mathbf{a}\cdot\mathbf{b}) && \text{(*3.25.1)}\end{aligned}$$

△

***3.30** *Theorem.* $(\mathbf{a}+\mathbf{b})\cdot\mathbf{c} = \mathbf{a}\cdot\mathbf{c} + \mathbf{b}\cdot\mathbf{c}$ for all $\mathbf{a},\mathbf{b},\mathbf{c} \in \mathbb{LV}_{OCCF,n}$. □

Proof:

$$\begin{aligned}(\mathbf{a}+\mathbf{b})\cdot\mathbf{c} &= \langle_i\, a_i + b_i\rangle\cdot\mathbf{c} && \text{(*2.2)}\\ &= \textstyle\sum_i (a_i+b_i)c_i && \text{(*3.26)}\\ &= \textstyle\sum_i a_ic_i + b_ic_i && \text{(distributive property in $OCCF$)}\\ &= \textstyle\sum_i a_ic_i + \sum_i b_ic_i && \text{(splitting the $\textstyle\sum_i$[8])}\\ &= \mathbf{a}\cdot\mathbf{c} + \mathbf{b}\cdot\mathbf{c} && \text{(*3.26 twice)}\end{aligned}$$

△

***3.31** *Theorem.* $\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) = \mathbf{a}\cdot\mathbf{b} + \mathbf{a}\cdot\mathbf{c}$ for all $\mathbf{a},\mathbf{b},\mathbf{c} \in DPS$. □

Proof:

$$\begin{aligned}\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) &= ((\mathbf{b}+\mathbf{c})\cdot\mathbf{a})^* && \text{(*3.25.1)}\\ &= (\mathbf{b}\cdot\mathbf{a} + \mathbf{c}\cdot\mathbf{a})^* && \text{(*3.25.3)}\\ &= ((\mathbf{a}\cdot\mathbf{b})^* + (\mathbf{a}\cdot\mathbf{c})^*)^* && \text{(*3.25.1 twice)}\\ &= (\mathbf{a}\cdot\mathbf{b} + \mathbf{a}\cdot\mathbf{c})^{**} && \text{(*3.1.2)}\\ &= \mathbf{a}\cdot\mathbf{b} + \mathbf{a}\cdot\mathbf{c} && \text{(*3.1.1)}\end{aligned}$$

△

***3.32** *Theorem.* $\mathbf{a}\cdot\mathbf{a} \in \operatorname{core} OCCF$ and $\mathbf{a}\cdot\mathbf{a} \geq 0$ for all $\mathbf{a} \in \mathbb{LV}_{OCCF,n}$, and $\mathbf{a}\cdot\mathbf{a} = 0$ if and only if $\mathbf{a} = \mathbf{0}$. □

Proof: According to *3.9.4, $xx^* \in \operatorname{core} OCCF$ and $xx^* \geq 0$ for all $x \in OCCF$. In addition, $\mathbf{a}\cdot\mathbf{a} = \sum_i a_i a_i{}^*$ by *3.26. When we combine these two, a straightforward application of *1.26 (with $\operatorname{core} OCCF$ as the ordered field) proves the theorem. △

***3.33** *Theorem.* $(\mathbb{LV}_{OCCF,n}, +, \mathbf{0}, (-), \cdot, (^*), \cdot_{\text{dot}})$ is a dot product space over $OCCF$, where $+$, $\mathbf{0}$, $(-)$, $\cdot$, $(^*)$, and $\cdot_{\text{dot}}$ are as defined in *2.2, *2.3, *2.5, *2.11, *3.20, and *3.26. Again, this dot product space will be denoted $\mathbb{LV}_{OCCF,n}$, the same as its set of vectors. □

[8] In general, splitting a $\sum$ involves replacing an expression $(x_1+y_1)+(x_2+y_2)+\ldots+(x_n+y_n)$ by $(x_1+x_2+\ldots+x_n)+(y_1+y_2+\ldots+y_n)$. The validity of this rearrangement can be proven by induction using the commutativity and associativity of $+$, in similar spirit to "repeated use".

3.3 Orthogonality and size

***3.34** *Definition: Orthogonal.* Two vectors $\mathbf{a}$ and $\mathbf{b}$ are said to be **orthogonal** if and only if $\mathbf{a} \cdot \mathbf{b} = 0$. (The order of the vectors is unimportant because, by *3.25.1, $\mathbf{a} \cdot \mathbf{b} = 0$ if and only if $\mathbf{b} \cdot \mathbf{a} = 0^* = 0$.) □

***3.35** *Theorem.* In any dot product space, all vectors are orthogonal to $\mathbf{0}$. □

Proof: Consider any vector $\mathbf{a}$. By *2.17, $0(\mathbf{0}) = \mathbf{0}$. Thus:

$$\begin{aligned} \mathbf{0} \cdot \mathbf{a} &= (0(\mathbf{0})) \cdot \mathbf{a} && \text{(see above)} \\ &= 0(\mathbf{0} \cdot \mathbf{a}) && \text{(*3.25.2)} \\ &= 0 && \text{(*1.14)} \end{aligned}$$

Thus, by *3.34, $\mathbf{a}$ is orthogonal to $\mathbf{0}$. △

***3.36** *Definition: Size for vectors.* $\|\mathbf{a}\| = \mathbf{a} \cdot \mathbf{a}$ for all vectors $\mathbf{a} \in DPS$. $\|\mathbf{a}\|$ is called the **size** or **magsquare** of a. □

***3.37** *Theorem.* Let DPS be a dot product space over an ordered-core conjugation field $OCCF$. Then $\|\mathbf{a}\| \in \operatorname{core} OCCF$ and $\|\mathbf{a}\| \geq 0$ for all vectors $\mathbf{a}$, and $\|\mathbf{a}\| = 0$ if and only if $\mathbf{a} = \mathbf{0}$. □

Proof: This is really just a restatement of *3.25.4. △

In analogy to *3.13, we define:

***3.38** *Definition: Absolute value for vectors.* $|\mathbf{a}| = \sqrt{\|\mathbf{a}\|}$ if this is defined, for all vectors $\mathbf{a} \in DPS$. $|\mathbf{a}|$ is called the **absolute value**, **magnitude**, or **norm** of $\mathbf{a}$. □

***3.39** *Theorem.* Let DPS be a dot product space over an ordered-core conjugation field $OCCF$. Then $\|r\mathbf{a}\| = \|r\|\|\mathbf{a}\|$ for all $r \in OCCF$ and $\mathbf{a} \in DPS$. □

Proof:

$$\begin{aligned} \|r\mathbf{a}\| &= (r\mathbf{a}) \cdot (r\mathbf{a}) && \text{(*3.36)} \\ &= r(\mathbf{a} \cdot (r\mathbf{a})) && \text{(*3.25.2)} \\ &= rr^*(\mathbf{a} \cdot \mathbf{a}) && \text{(*3.29)} \\ &= \|r\|\|a\| && \text{(*3.10 and *3.36)} \end{aligned}$$

△

***3.40** *Theorem.* Let DPS be a dot product space over an ordered-core conjugation field $OCCF$. Then $|r\mathbf{a}| = |r||\mathbf{a}|$ for all $r \in OCCF$ and $\mathbf{a} \in DPS$ for which both sides of the equation are defined. □

Proof: Analogous to the proof of *3.14. △

The triangle inequality (*3.15) also applies to vectors. In fact, the proofs of the vector triangle inequality and its lemma are nearly identical to those of the scalar versions. a and b become vectors while t and k remain scalars. Some multiplication is replaced by the dot product, and many conjugation symbols must be dropped because the dot product includes some conjugation of its own. Certain uses of theorems about ordered-core conjugation fields are replaced by their dot product space analogues. The reader is encouraged to go back through both proofs and verify that each step still holds.

3.4 Unit vectors

A unit vector is one with a very special length: 1. Unit vectors are occasionally useful. In this section we will be concerned only with dot product spaces whose ordered-core conjugation fields have square roots.

***3.41** *Definition: Unit vector.* A vector $\mathbf{a}$ is said to be a **unit vector** if $\|\mathbf{a}\| = 1$. □

***3.42** *Definition: Unit vector operator.* $\mathrm{unitV}\,\mathbf{a} = \mathbf{a}/|a|$ for any vector $\mathbf{a}$ other than $\mathbf{0}$. □

***3.43** SameDirectionSameUnitVector $\mathrm{unitV}\,\mathbf{a} = \mathrm{unitV}\,\mathbf{b}$ if and only if $\mathbf{a}$ and $\mathbf{b}$ have the same direction (see *2.23). This applies to all vectors $\mathbf{a}$ and $\mathbf{b}$, neither of which is $\mathbf{0}$. □

Proof: $\Rightarrow$**:** We have $\mathrm{unitV}\,\mathbf{a} = \mathrm{unitV}\,\mathbf{b} \Rightarrow \mathbf{a}/|a| = \mathbf{b}/|b| \Rightarrow \mathbf{a} = (|a|/|b|)\mathbf{b}$. Take $r = |a|/|b|$.

$\Leftarrow$**:**

$$\begin{aligned} \mathrm{unitV}\,\mathbf{a} &= \mathrm{unitV}\,(r\mathbf{b}) && \text{(because } \mathbf{a} = r\mathbf{b}\text{)} \\ &= \tfrac{r\mathbf{b}}{|r\mathbf{b}|} && \text{(*3.42)} \\ &= \tfrac{r\mathbf{b}}{|r||\mathbf{b}|} && \text{(*3.40)} \\ &= \tfrac{r\mathbf{b}}{r|\mathbf{b}|} && (r > 0 \Rightarrow |r| = r) \\ &= \tfrac{\mathbf{b}}{|\mathbf{b}|} && \text{(canceling)} \\ &= \mathrm{unitV}\,\mathbf{b} && \text{(*3.42)} \end{aligned}$$

△

3.5 Additional theorems and geometry

Here are two useful theorems about vector sizes:

***3.44** *Theorem.* $\|\mathbf{a}+\mathbf{b}\| = \|\mathbf{a}\| + \|\mathbf{b}\| + \mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{a}$ for any vectors $\mathbf{a}$ and $\mathbf{b}$. (In a simple dot product space, this reduces to $\|\mathbf{a}+\mathbf{b}\| = \|\mathbf{a}\| + \|\mathbf{b}\| + 2(\mathbf{a}\cdot\mathbf{b})$.) □

Proof:

$$\begin{aligned} \|\mathbf{a}+\mathbf{b}\| &= (\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b}) && \text{(*3.36)} \\ &= \mathbf{a}\cdot\mathbf{a} + \mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{a} + \mathbf{b}\cdot\mathbf{b} && \text{(*3.25.3 and *3.31)} \\ &= \|\mathbf{a}\| + \|\mathbf{b}\| + \mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{a} && \text{(*3.36 twice)} \end{aligned}$$

△

***3.45** *Theorem.* $\|\mathbf{a}-\mathbf{b}\| = \|\mathbf{a}\| + \|\mathbf{b}\| - \mathbf{a}\cdot\mathbf{b} - \mathbf{b}\cdot\mathbf{a}$ for any vectors $\mathbf{a}$ and $\mathbf{b}$. (In a simple dot product space, this reduces to $\|\mathbf{a}-\mathbf{b}\| = \|\mathbf{a}\| + \|\mathbf{b}\| - 2(\mathbf{a}\cdot\mathbf{b})$.) □

Proof: Put $b \mapsto -b = (-1)b$ in *3.44, and use *3.25.2 and *3.29 to move the negations to the outside. △

*3.45 looks suspiciously like the Law of Cosines; let us investigate this potential connection to geometry by writing the Law of Cosines using real vectors.

Consider a triangle ABC with two sides $\mathbf{a}$ and $\mathbf{b}$; then $\mathbf{c} = \mathbf{a} - \mathbf{b}$ is the third side. The Law of Cosines, in vector form, would state:

***3.46** *Vector Law of Cosines.* $\|\mathbf{a}-\mathbf{b}\| = \|\mathbf{a}\| + \|\mathbf{b}\| - 2|\mathbf{a}||\mathbf{b}|\cos\theta(\mathbf{a},\mathbf{b})$ □

($\theta(\mathbf{a},\mathbf{b})$ denotes the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$.) This suggests the following:

***3.47** $\cos\theta(\mathbf{a},\mathbf{b}) = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$. □

We have two options at this point; recall Section 1.3.

The purely algebraic approach is to take *3.47 as a definition of $\theta(\mathbf{a},\mathbf{b})$ for any simple dot product space over the real numbers. *3.46 can then be deduced from *3.47 and *3.45. No formal connection would be made between $\theta(\mathbf{a}, \mathbf{b})$ and ordinary geometric angles, nor would a connection be made between *3.46 and the Law of Cosines in geometry. It simply happens that vectors have similar properties.

The geometric approach is to define $\theta(\mathbf{a}, \mathbf{b})$ to be the ordinary geometric angle between $\mathbf{a}$ and $\mathbf{b}$. *3.46 is considered to be another form of the geometric Law of Cosines and is therefore a theorem. *3.47 can then be deduced from *3.46 and *3.45. Some textbooks take this approach but don't present *3.45 as a standalone theorem of algebraic vector theory. They establish it quickly for the proof of *3.47 and then move on.

Either way, *3.46 and *3.47 are henceforth to be considered true statements.

Just as *3.45 shows the equivalence of *3.47 and *3.46, *3.44 shows the equivalence of *3.47 and this statement:

***3.48** *Vector Law of Cosines, additive form.* $\|\mathbf{a}+\mathbf{b}\| = \|\mathbf{a}\| + \|\mathbf{b}\| + 2|\mathbf{a}||\mathbf{b}|\cos\theta(\mathbf{a}, \mathbf{b})$ □

This corresponds to a theorem of geometry that is easily seen to be equivalent to the Law of Cosines:

***3.49** *Geometric Law of Cosines, additive form.* For any parallelogram $ABCD$,

$$AC^2 = AB^2 + AD^2 + 2(AB)(AD)\cos\angle BAD.$$

□

To reduce this to the ordinary Law of Cosines, replace AD by BC (they are equal). Then, since $\angle BAD$ and $\angle CBA$ are supplementary, $\cos\angle BAD = -\cos\angle CBA$. Thus, the Law of Cosines and *3.49 are equivalent within geometry. Since *3.46 and *3.48 are both equivalent to *3.47, they are similarly equivalent within vector analysis.

Imagine a gigantic, beautiful graph whose vertices are the concepts of mathematics and whose edges are relationships between two concepts. A casual observer might take this as a finished object and consider all of the relationships to be already proved; the dedicated mathematician would want to establish all the relationships in a logical fashion. Upon arriving at a new concept, she gets one free relationship: the one embodied in the concept's definition. She must prove the others.

For example, we can define $\sin x$ with right triangles and geometry (one relationship). Then we can show it is a solution to the differential equation $y'' + y = 0$, that its Maclaurin series is $x - \frac{x^3}{3!} + \frac{x^5}{5!} \ldots$, and that its inverse $\sin^{-1} x$ is $\int \frac{1}{\sqrt{1-x^2}}\, dx$ (three more). Or we could define $\sin x$ using the Maclaurin series and show that the other three relationships hold.

We are in the same sort of situation here with $\theta(\mathbf{a}, \mathbf{b})$. Which relationship, the formula $\frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$ or the Law of Cosines, would you choose as the "freebie"?

What now?

This marks the end of the first part of this document. We have seen what a vector space is and how the list-vectors provide a canonical example. We have looked at the dot product and how it works on list-vectors; we put considerable effort into making it work sanely on complex numbers. This was accomplished with the greatest possible formality and the greatest possible generality.

In the coming sections we will examine several more operations and functions, including the generalized cross product and the hash product. We will also do some vector calculus, apply vectors to a physics problem, and investigate a beautiful generalization of matrices. However, the rigor of the first part will be replaced by a lighter, more intuitive discussion, and "vector" henceforth means "real vector" unless otherwise stated.

4 Volume and Related Issues

4.1 Making parallelepipeds

Given m vectors in n dimensions, one can make a parallelepiped out of them. Call the vectors $\mathbf{a}_i$ for i from 1 to m. Here the subscript i is a list index and does not indicate a single component of a vector. Start with a single point, and gradually move it through the translation represented by $\mathbf{a}_1$; it sweeps out a line segment. Now gradually move this segment through the translation represented by $\mathbf{a}_2$ to get a parallelogram. This process can be repeated for all m vectors, forming an m-dimensional parallelepiped if all is well. If some vector $\mathbf{a}_i$ is $\mathbf{0}$ or lies in the $(i-1)$-dimensional plane containing the parallelepiped formed from the previous vectors, then this procedure produces a degenerate parallelepiped.

If $m > n$, then the parallelepiped will definitely be degenerate. This is because, once the first n vectors have been used, either some vector $\mathbf{a}_i$ lies in the plane of the previous ones (in which case the parallelepiped is already degenerate), or a nondegenerate n-dimensional parallelepiped has been created. In the second case, the "plane" of the first n vectors is the entire coordinate space, and no matter what $\mathbf{a}_{n+1}$ is, it will lie in this "plane".

4.2 Linear dependence

The parallelepiped is nondegenerate exactly when the vectors are **linearly independent**. There are two similar ways of defining this, and it is nice to have both available. The first corresponds to the condition above that some vector lies in the plane of the previous vectors.

***4.1** *Definition.* A list of n vectors $\mathbf{a}_i$ is **linearly dependent** if and only if, for some k and some real numbers b_i for i from 1 to $k-1$, $\mathbf{a}_k = \sum_{i=1}^{k-1} b_i \mathbf{a}_i$. If no such k and b_i exist, the list is **linearly independent**. □

***4.2** *Definition.* A list of n vectors $\mathbf{a}_i$ is **linearly dependent** if and only if there exist real numbers c_i for i from 1 to n, not all zero, such that $\sum_i c_i \mathbf{a}_i = \mathbf{0}$. If setting $\sum_i c_i \mathbf{a}_i = \mathbf{0}$ forces one to conclude that all the c_i are zero, the list must be **linearly independent**. □

***4.3** *Theorem.* *4.1 and *4.2 are equivalent. □

Proof: We'll show that, if one definition proclaims a list of n vectors $\mathbf{a}_i$ linearly dependent, the other definition must do so.

Suppose the $\mathbf{a}_i$ are linearly dependent according to *4.1. Then we have k and b_i such that $\mathbf{a}_k = \sum_{i=1}^{k-1} b_i \mathbf{a}_i$. Consider the list of n real numbers c_i, where $c_i = b_i$ for i from 1 to $k-1$, $c_k = -1$, and $c_i = 0$ for $i > k$. Then:

$$\begin{aligned}
\textstyle\sum_{i=1}^{n} c_i \mathbf{a}_i &= \textstyle\sum_{i=1}^{k} c_i \mathbf{a} && \text{(because } c_i = 0 \text{ for } i > k\text{)} \\
&= \textstyle\sum_{i=1}^{k-1} b_i \mathbf{a}_i + (-1)\mathbf{a}_k && \text{(splitting the } \textstyle\sum\text{)} \\
&= \mathbf{a}_k + (-1)\mathbf{a}_k && \text{(because } \mathbf{a}_k = \textstyle\sum_{i=1}^{k-1} b_i \mathbf{a}_i\text{)} \\
&= 0 && \text{(canceling)}
\end{aligned}$$

Since $\sum_{i=1} n c_i \mathbf{a}_i = 0$ and c_k is nonzero (it is -1), the $\mathbf{a}_i$ are linearly dependent according to *4.2.

Now, suppose some list $\mathbf{a}_i$ is linearly dependent according to *4.2. Then there exist c_i such that $\sum_i c_i \mathbf{a}_i = \mathbf{0}$. Let k be the greatest index such that $c_k \neq 0$. Let $b_i = -c_i / c_k$ for i from 1 to $k-1$. Then:

$$
\begin{aligned}
\textstyle\sum_{i=1}^{k-1} b_i\mathbf{a}_i &= \textstyle\sum_{i=1}^{k-1}(-c_i/c_k)\mathbf{a}_i && \text{(definition of the } b_i\text{)}\\
&= (-1/c_k)\textstyle\sum_{i=1}^{k-1} c_i\mathbf{a}_i && \text{(pulled } -1/c_k \text{ out of the } \textstyle\sum\text{)}\\
&= \mathbf{a}_k + (-1/c_k)\textstyle\sum_{i=1}^{k-1} c_i\mathbf{a}_i - \mathbf{a}_k && \text{(added a convenient form of } \mathbf{0}\text{)}\\
&= \mathbf{a}_k - (\textstyle\sum_{i=1}^{k-1} c_i\mathbf{a}_i + c_k\mathbf{a}_k)/c_k\\
&= \mathbf{a}_k - (\textstyle\sum_{i=1}^{k} c_i\mathbf{a}_i)\\
&= \mathbf{a}_k - (\textstyle\sum_{i=1}^{k} c_i\mathbf{a}_i + \sum_{i=k+1} nc_i\mathbf{a}_i) && \text{(we added } \mathbf{0} \text{ because } c_i = 0 \text{ for } i > k\text{)}\\
&= \mathbf{a}_k - (\textstyle\sum_{i=1}^{n} c_i\mathbf{a}_i) && \text{(combining the } \textstyle\sum\text{s)}\\
&= \mathbf{a}_k && \text{(because } \textstyle\sum_{i=1}^{n} c_i\mathbf{a}_i = 0\text{)}
\end{aligned}
$$

Since $\sum_{i=1}^{k-1} b_i\mathbf{a}_i = \mathbf{a}_k$, the $\mathbf{a}_i$ are linearly dependent according to *4.1. △

With this theorem proved, one can say "linearly dependent" and use either definition at will.

The following is a theorem of linear algebra and will not be proved here:

***4.4** *Theorem.* A square matrix is singular (has no multiplicative inverse) if and only if its rows, when interpreted as vectors, are linearly dependent. An analogous result holds for columns. □

4.3 Determinants

The **determinant** is a function that takes a square matrix A and returns a real number denoted $\det A$. It has the properties stated in *4.6 through *4.14 below, and it happens to be the only function with these properties. There are two formulas that can be used to define the determinant. This one is my personal favorite because it is nonrecursive and makes all the properties easy to prove:

***4.5** *Definition.* Let $\operatorname{Inv} p$ denote the number of inversions of the permutation p, which is the number of pairs (i, j) for which $i < j$ and $p(i) > p(j)$. Then the **determinant** of an $n \times n$ matrix A is the sum below, taken over all permutations p of $1, \ldots, n$:

$$\sum_p \left((-1)^{\operatorname{Inv} p} \prod_i A_{i,p(i)} \right)$$

□

Theorems analogous to all the results below about rows of a matrix hold for columns. This is noted in the statement of each such result, but proofs are only presented for rows. (The proofs for columns are analogous, or one could apply *4.6 below.)

***4.6** *Theorem.* For any square matrix A, $\det A = \det A^T$. □

Proof: This follows from *4.5 because a permutation has the same number of inversions as its inverse. △

***4.7** *Theorem.* Multiplying one row (column) of a square matrix by a real number r multiplies its determinant by r. □

Proof: Each term of the $\sum$ in *4.5 gets multiplied by r; this new factor appears in the $\prod$ when i is the index of the row that got multiplied. △

***4.8** *Theorem.* Let A, B, and C be square matrices of the same size. Suppose A, B, and C are identical except for row (column) i, for some i. Suppose further that row (column) i of C is the sum of row (column) i of A and row (column) i of B. Then $\det C = \det A + \det B$. □

Proof: This is massive application of the distributive property on *4.5. The ith factor in the $\prod$ in $\det C$ can be written as a sum of a value from A and a value from B; pull this addition to the outside of the formula. △

***4.9** *Theorem.* Interchanging two rows (columns) of a square matrix negates its determinant. □

Proof: This follows from *4.5 because swapping two values in a permutation p changes $\operatorname{Inv} p$ by an odd number. △

***4.10** *Theorem.* The determinant of a multiplicative identity matrix is 1. □

Proof: All terms of the $\sum$ in *4.5 are zero except one: the term with $p(i) \equiv i$. This p has no inversions, and the $\prod$ evaluates to 1. △

***4.11** *Theorem.* If a square matrix has an entire row (column) of zeros, its determinant is 0. □

Proof: By *4.7, multiplying the row of zeros by 0 multiples the determinant by 0, making the determinant 0. However, this operation does not change the matrix, so the determinant of the original matrix must be 0. △

***4.12** *Theorem.* If a square matrix has two equal rows (columns), its determinant is 0. □

Proof: By *4.9, swapping the two equal rows negates the determinant. But since this swap doesn't change the matrix, the determinant can't change. The only way this can happen is if the determinant is 0. △

***4.13** *Theorem.* Let B be the matrix produced by adding r times row (column) i of a square matrix A to row j. Then $\det B = \det A$ for any row (column) indices i and j ($i \neq j$) and real number r. □

Proof: Consider the matrix C that matches A in all rows except row j. Row j of C is row i of A. By *4.12, $\det C = 0$.

Now consider the matrix D that matches C in all rows except row j. Row j of D is r times row j of C. By *4.7, $\det D = r \det C = r(0) = 0$.

Row j of B is the sum of row j of A and row j of D, while B, A, and D match in all other rows. By *4.8, $\det B = \det A + \det D = \det A + 0 = \det A$. △

***4.14** *Theorem.* $\det AB = (\det A)(\det B)$ for any square matrices A and B of the same size. □

Proof: Magical but messy. It appears in Section 4.7. △

***4.15** *Theorem.* Let A be any $n \times n$ square matrix. The following statements are equivalent:

(1). A is singular.

(2). The rows (columns) of A, when interpreted as vectors, are linearly dependent.

(3). $\det A = 0$.

□

Proof: *4.4 establishes that *4.15.(1) ⇒ *4.15.(2). We will show that *4.15.(2) ⇒ *4.15.(3) and that *4.15.(3) ⇒ *4.15.(1). This circle of implication shows that, if one of the three statements is true, all three must be true, so they are equivalent.

*4.15.(3) ⇒ *4.15.(1): Suppose $\det A = 0$ but A has a multiplicative inverse A^{-1}. Let I_n be the $n \times n$ multiplicative identity matrix. Then:

$$
\begin{aligned}
1 &= \det I_n && \text{(*4.10)} \\
&= \det AA^{-1} && \text{(because } A \text{ and } A^{-1} \text{ are multiplicative inverses)} \\
&= (\det A)(\det A^{-1}) && \text{(*4.14)} \\
&= 0(\det A^{-1}) && (\det A = 0) \\
&= 0
\end{aligned}
$$

This is a contradiction, so A can have no multiplicative inverse.

*4.15.(2) $\Rightarrow$ *4.15.(3): Let $\mathbf{a}_i$ be the ith row of A, interpreted as a vector. Using *4.1, there exists a k and numbers b_i such that $\mathbf{a}_k = \sum_{i=1} k - 1 b_i \mathbf{a}_i$.

Suppose, for each i from 1 to $k-1$, we add $-b_i$ times row i to row k of A. Since $\mathbf{a}_k = \sum_{i=1}^{k-1} b_i \mathbf{a}_i$, row k of the resulting matrix has all zeros. By *4.11, this matrix must have a determinant of 0. By *4.13, none of the row additions changed the determinant of the matrix, so A must also have a determinant of 0. △

4.4 Volume

The following is a formula for the volume of a parallelepiped based on the vectors for its sides:

***4.16** *Definition.* The m-dimensional **volume** of a list of m vectors $\mathbf{a}_i$ in n dimensions is the nonnegative real number $\operatorname{vol}_i\ \mathbf{a}_i = \sqrt{\det D}$, where $D = [_{1 \le i \le m, 1 \le j \le m}\ \mathbf{a}_i \cdot \mathbf{a}_j]$. □

It is interesting that, in order to compute the volume of a list of vectors, one need know only their mutual dot products, not the vectors themselves.

As defined here, "volume" is but an operation of vector analysis. If one wishes to formally connect vector analysis to geometry (see Section 1.3), one would want to prove that this formula gives the correct geometric volume. To do so, one would need a volume formula from geometry for parallelepipeds in any number of dimensions. One attractive plan is to make *this* the volume formula and assume it true without proof, which would put it on equal footing with formulas like $A = bh$. Another possibility is to define the volume of any set through integrals; the Hubbards manage this in [2].

Given a list of m vectors $\mathbf{a}_i$ in n dimensions, consider the matrix $A = [_{1 \le i \le m, 1 \le j \le n}\ (\mathbf{a}_i)_j]$. The vectors $\mathbf{a}_i$ appear as rows of A, while they appear as columns of A^T (the transpose of A: flip rows and columns). In forming the product AA^T, each row of A is put onto each column of A^T; this has the effect of taking the dot product of two vectors in the list. Thus AA^T is equal to the D of *4.16.

Here is a powerful theorem about volume:

***4.17** *Theorem.* The volume of a list of vectors is zero if and only if the list is linearly dependent. □

Proof: Let D be the matrix in *4.16. We wish to show that $\sqrt{\det D} = 0$ exactly when the $\mathbf{a}_i$ are linearly dependent. Let $\mathbf{u}_i$ be the ith row of M, interpreted as a vector; that is, $\mathbf{u}_i = \left\langle_j\ \mathbf{a}_i \cdot \mathbf{a}_j \right\rangle$. By *4.15, $\sqrt{\det D} = 0$ if and only if the $\mathbf{u}_i$ are linearly dependent, so we need only show that the $\mathbf{a}_i$ are linearly dependent if and only if the $\mathbf{u}_i$ are.

As a lemma, we can find a simpler form for a linear combination of the $\mathbf{u}_i$. This is presented here because it will be used in both directions of the proof to follow.

$$\begin{aligned}
\textstyle\sum_i b_i\mathbf{u}_i &= \textstyle\sum_i b_i \left\langle_j\, \mathbf{a}_i \cdot \mathbf{a}_j \right\rangle && \text{(definition of the } \mathbf{u}_i\text{)}\\
&= \textstyle\sum_i \left\langle_j\, b_i(\mathbf{a}_i \cdot \mathbf{a}_j) \right\rangle && \text{(*2.11)}\\
&= \left\langle_j\, \textstyle\sum_i b_i(\mathbf{a}_i \cdot \mathbf{a}_j) \right\rangle && \text{(repeated use of *2.2)}\\
&= \left\langle_j\, \textstyle\sum_i (b_i\mathbf{a}_i) \cdot \mathbf{a}_j \right\rangle && \text{(*3.29)}\\
&= \left\langle_j\, (\textstyle\sum_i b_i\mathbf{a}_i) \cdot \mathbf{a}_j \right\rangle && \text{(repeated use of *3.25.3)}
\end{aligned}$$

For the forward direction, suppose the $\mathbf{a}_i$ are linearly dependent. Then there are real numbers b_i, not all zero, such that $\sum_i b_i\mathbf{a}_i = \mathbf{0}$. The same real numbers b_i prove the $\mathbf{u}_i$ linearly dependent. This is why:

$$\begin{aligned}
\textstyle\sum_i b_i\mathbf{u}_i &= \left\langle_j\, (\textstyle\sum_i b_i\mathbf{a}_i) \cdot \mathbf{a}_j \right\rangle && \text{(by the lemma)}\\
&= \left\langle_j\, \mathbf{0} \cdot \mathbf{a}_j \right\rangle && \text{(because } \textstyle\sum_i b_i\mathbf{a}_i = \mathbf{0}\text{)}\\
&= \left\langle_j\, 0 \right\rangle && \text{(*3.35)}\\
&= \mathbf{0} && \text{(*2.3)}
\end{aligned}$$

For the backward direction, suppose the $\mathbf{u}_i$ are linearly dependent. Then there are real numbers b_i, not all zero, such that $\sum_i b_i\mathbf{u}_i = \mathbf{0}$. Again, the same numbers b_i prove the $\mathbf{a}_i$ linearly dependent. By the lemma, $\left\langle_j\, (\sum_i b_i\mathbf{a}_i) \cdot \mathbf{a}_j \right\rangle = \mathbf{0}$. This means that, for every j, $(\sum_i b_i\mathbf{a}_i) \cdot \mathbf{a}_j = 0$. Now:

$$\begin{aligned}
(\textstyle\sum_i b_i\mathbf{a}_i) \cdot (\textstyle\sum_j b_j\mathbf{a}_j) &= \textstyle\sum_j ((\textstyle\sum_i b_i\mathbf{a}_i) \cdot \mathbf{a}_j) && \text{(repeated use of *3.31)}\\
&= \textstyle\sum_j 0 && \text{(because } (\textstyle\sum_i b_i\mathbf{a}_i) \cdot \mathbf{a}_j = 0 \text{ for every } j\text{)}\\
&= 0 && (0 + 0 + \ldots = 0)
\end{aligned}$$

But $(\sum_i b_i\mathbf{a}_i)$ and $(\sum_j b_j\mathbf{a}_j)$ are really the same vector because i and j are dummy variables. So we have $(\sum_i b_i\mathbf{a}_i) \cdot (\sum_i b_i\mathbf{a}_i) = 0$. By *3.25.4, $\sum_i b_i\mathbf{a}_i = \mathbf{0}$, so the $\mathbf{a}_i$ are linearly dependent. △

The theorems below describe how volume is affected when a list of vectors is changed in certain ways. In general, they are proved by showing how the changes affect the matrix D in *4.16, and properties of determinants (see Section 4.3) are used to find how the changes to D affect its determinant.

***4.18** *Theorem.* Interchanging two vectors in a list does not change its volume. □

Proof: Interchanging $\mathbf{a}_i$ and $\mathbf{a}_j$ has the effect of swapping rows i and j in the matrix D of *4.16 and then swapping columns i and j. By *4.9, each such swap negates $\det D$, so the two swaps together leave $\det D$ and hence $\mathrm{vol}_i\, \mathbf{a}_i$ unchanged. △

***4.19** *Theorem.* Multiplying a vector in a list by a real number r multiplies the volume of the list by $|r|$. □

Proof: By *3.29, replacing $\mathbf{a}_i$ by $r\mathbf{a}_i$ has the effect of multiplying row i and column i of D each by r. (This means the entry $D_{i,i}$, which is $\mathbf{a}_i \cdot \mathbf{a}_i$, gets multiplied by r^2.) By *4.7, these multiplications cause $\det D$ to be multiplied by r^2, so $\mathrm{vol}_i\, \mathbf{a}_i = \sqrt{\det D}$ is multiplied by $|r|$. △

***4.20** *Theorem.* Adding a multiple of one vector in a list to another vector in the list does not change the volume of the list. □

Proof: By *3.25.3 and *3.31, replacing $\mathbf{a}_j$ by $\mathbf{a}_j + r\mathbf{a}_i$ has the effect of adding r times row i of D to row j and then adding r times column i to column j. By *4.13, both additions leave $\det D$ and hence $\mathrm{vol}_i\, \mathbf{a}_i$ unchanged. △

***4.21** *Theorem.* The volume of an orthonormal list of vectors is 1. (A list of vectors is **orthonormal** if and only if every vector has magnitude 1 and any two vectors from the list are orthogonal.) □

Proof: Let the vectors be $\mathbf{a}_i$ for i from 1 to m. Then $\mathbf{a}_i \cdot \mathbf{a}_j = \delta_{i,j}$ because the $\mathbf{a}_i$ are orthonormal. Consequently, D is the $m \times m$ multiplicative identity matrix, whose determinant is 1 by *4.10. Then, $\mathrm{vol}_i\, \mathbf{a}_i = \sqrt{\det D} = 1$. △

4.5 Orientation

When one sets up a 3D coordinate system in the real world, apart from matters of scale and rotation, one has two different choices for how to orient the axes, typically called "right-handed" and "left-handed". Once the x and y axes have been placed, the z axis can point in one of two opposite directions. If one makes an L with the thumb and index finger of one's right hand and then bends the middle finger so it is orthogonal to the other two, the thumb, index, and middle fingers will form the x, y, and z axes of a right-handed coordinate system. If one uses the left hand, a left-handed coordinate system will be formed.[9]Right-handed coordinate systems are more commonly used, especially in physics.

One can also speak of the "handedness" of a list of 3 vectors in the real world. In this case, the thumb, index, and middle fingers of one hand are pointed in the direction of the first, second, and third vectors, respectively, and the vectors are found to be either "left-handed" or "right-handed". (Order is extremely important!)

The terms "left-handed" and "right-handed" rely on the real world for their meaning; they are meaningless in vector analysis and geometry. It is possible, however, to introduce a worthwhile notion of "handedness" in vector analysis. We'll say a list of vectors is positive-handed if they have the same handedness as the vectors $\mathbf{1}_i$ (the positive directions of the coordinate axes) and negative-handed if they have the opposite handedness. This will be formalized in Section 4.6.

Handedness can be generalized to n vectors in n dimensions with a suitable "hyperhand", although this is beyond normal human intuition. If one has fewer than n vectors in n dimensions, or if the collection of n vectors is linearly dependent, either hand can fit the vectors, and handedness is meaningless.

4.6 Signed volume

*4.16 reduces to a rather simple form if $m = n$. Let A be the matrix in the remarks following *4.16. When $m = n$, A and A^T are square matrices, so $\det A$ and $\det A^T$ are defined. By *4.14 we can split up the determinant in *4.16: $\det AA^T = (\det A)(\det A^T)$. $\det A = \det A^T$, so the volume of the vectors is $|\det A|$.

This formula explicitly throws out the sign of $\det A$ in order to provide a nonnegative volume. In fact, the sign gives the handedness of the vectors. This makes sense because, if the vectors are linearly dependent, their volume is zero, and handedness is meaningless. If zero volume and undefined handedness did not occur in exactly the same cases, this "encoding" of volume and handedness in a single real number would not work.

***4.22** *Definition.* The signed volume of a list of n vectors $\mathbf{a}_i$ in n dimensions is $\mathrm{svol}_i\, \mathbf{a}_i = \det A$, where $A = [_{1 \le i \le n, 1 \le j \le n} (\mathbf{a}_i)_j]$. □

The derivation above proves this theorem:

***4.23** *Theorem.* For any list of n vectors $\mathbf{a}_i$ in n dimensions, $\mathrm{vol}_i\, \mathbf{a}_i = |\mathrm{svol}_i\, \mathbf{a}_i|$. □

***4.24** *Definition.* A list of n vectors $\mathbf{a}_i$ in n dimensions has **positive handedness** if and only if $\mathrm{svol}_i\, \mathbf{a}_i > 0$ and negative handedness if and only if $\mathrm{svol}_i\, \mathbf{a}_i < 0$. □

[9]This explanation of handedness with fingers is from [2].

This theorem, which is a special case of *4.17, can also be proved easily from *4.22 and *4.15:

***4.25** *Theorem.* The signed volume of a list of vectors is nonzero if and only if the list is linearly independent. □

The theorems below are similar to *4.18 through *4.21; they describe how *signed* volume responds to certain changes in a list of vectors. Since the matrix A in *4.22 contains the vectors $\mathbf{a}_i$ themselves, not their mutual dot products, the theorems below are somewhat easier to prove than those of Section 4.4; their proofs have been omitted.

***4.26** *Theorem.* Interchanging two vectors in a list negates its signed volume. □

***4.27** *Theorem.* Multiplying a vector in a list by a real number r multiplies the signed volume of the list by r. □

***4.28** *Theorem.* Adding a multiple of one vector in a list to another does not change the signed volume of the list. □

This is almost a special case of *4.21, but it makes an additional statement about handedness:

***4.29** *Theorem.* In any number of dimensions, the signed volume of the vectors $\mathbf{1}_i$ is 1. □

By *4.29, the vectors $\mathbf{1}_i$ have positive handedness, and the definition of positive and negative handedness in *4.24 agrees with the earlier remark about handedness.

4.7 Proof of *4.14

Here is the "magical but messy" proof of a result about determinants:

***4.14** *Theorem.* $\det AB = (\det A)(\det B)$ for any square matrices A and B of the same size. □

Proof:

$$\det AB = \sum_p \left((-1)^{\operatorname{Inv} p} \prod_i (AB)_{i,p(i)} \right) \qquad \text{(*4.5)}$$

$$= \sum_p \left((-1)^{\operatorname{Inv} p} \prod_i \left(\sum_k A_{i,k} B_{k,p(i)} \right) \right) \qquad \text{(definition of matrix multiplication)}$$

The $\prod_i$ represents the product of n sums, one for each i. Each sum has n terms, one for each value of k. By "repeated use" of the distributive property, turn this into a sum of n^n terms. Each of these n^n terms is a product of n factors, one for each i; each factor has a certain value of k. For each term, let's think of the sequence of ks in its factors as a function $k(i)$, where $k : \{1, \ldots, n\} \to \{1, \ldots, n\}$. There are clearly n^n possible functions k, which correspond to the n^n terms in the sum. The result:

$$\sum_p \left((-1)^{\operatorname{Inv} p} \prod_i \left(\sum_{k=1}^{n} A_{i,k} B_{k,p(i)} \right) \right) = \sum_p \left((-1)^{\operatorname{Inv} p} \sum_{k:\{1,\ldots,n\}\to\{1,\ldots,n\}} \left(\prod_i A_{i,k(i)} B_{k(i),p(i)} \right) \right)$$

$$= \sum_{p,k} \left((-1)^{\operatorname{Inv} p} \prod_i A_{i,k(i)} B_{k(i),p(i)} \right) \qquad \text{(move summing on } k \text{ to the outside)}$$

$$= \sum_k \left(\prod_i A_{i,k(i)} \right) \left(\sum_p (-1)^{\operatorname{Inv} p} \left(\prod_i B_{k(i),p(i)} \right) \right) \qquad \text{(move summing on } p \text{ to the inside)}$$

Now, the $\sum_p$ part on the right has the form of the determinant of a matrix D whose rows are all rows of B; precisely, row i of D is row $k(i)$ of B. Suppose the function k is such that $k(a) = k(b)$ for some $a \neq b$. Then rows a and b of D are equal, its determinant is zero by *4.12, and the entire term for that function k drops out of the main sum. Thus, we may change the main sum to restrict k to *permutations* of $\{1, \ldots, n\}$, not just any functions. We have:

$$\sum_k \left(\prod_i A_{i,k(i)}\right)\left(\sum_p (-1)^{\operatorname{Inv} p}\left(\prod_i B_{k(i),p(i)}\right)\right) = \sum_{\text{perms}\, k}\left(\prod_i A_{i,k(i)}\right)\left(\sum_{\text{perms}\, p} (-1)^{\operatorname{Inv} p}\left(\prod_i B_{k(i),p(i)}\right)\right)$$

Let q be the permutation with $q(x) = p(k^{-1}(x))$, so that $p(i) = q(k(i))$. Then:

$$\sum_k \left(\prod_i A_{i,k(i)}\right)\left(\sum_p (-1)^{\operatorname{Inv} p}\left(\prod_i B_{k(i),p(i)}\right)\right) = \sum_k \left(\prod_i A_{i,k(i)}\right)\left(\sum_p (-1)^{\operatorname{Inv} p}\left(\prod_i B_{k(i),q(k(i))}\right)\right)$$

In the $\prod_i$, we can throw out the uses of k, because we multiply over the same values of i either way. Plus, for a given k (determined by the outer $\sum$), permutations p and q match one-to-one, so the inner $\sum$ can be over all q instead of all p:

$$\sum_k \left(\prod_i A_{i,k(i)}\right)\left(\sum_p (-1)^{\operatorname{Inv} p}\left(\prod_i B_{k(i),q(k(i))}\right)\right) = \sum_k \left(\prod_i A_{i,k(i)}\right)\left(\sum_q (-1)^{\operatorname{Inv} p}\left(\prod_i B_{i,q(i)}\right)\right)$$

It can be shown that $\operatorname{Inv} p \equiv \operatorname{Inv} q + \operatorname{Inv} k \pmod 2$ since p is the composition of q and k. Brief proof: if q has an odd (even) number of inversions, it can be built from the identity permutation in an odd (even) number of swaps, so p can be built from k in an odd (even) number of swaps. Each swap changes $\operatorname{Inv} k$ by an odd amount, so at the end, $\operatorname{Inv} p \equiv (\#\,\text{of swaps}) + \operatorname{Inv} k \equiv \operatorname{Inv} q + \operatorname{Inv} k \pmod 2$. This result tells us that $(-1)^{\operatorname{Inv} p} = (-1)^{\operatorname{Inv} q}(-1)^{\operatorname{Inv} k}$, so:

$$\sum_k \left(\prod_i A_{i,k(i)}\right)\left(\sum_q (-1)^{\operatorname{Inv} p}\left(\prod_i B_{i,q(i)}\right)\right) = \sum_k \left(\prod_i A_{i,k(i)}\right)\left(\sum_q (-1)^{\operatorname{Inv} q}(-1)^{\operatorname{Inv} k}\left(\prod_i B_{i,q(i)}\right)\right)$$

Now, we just pull the $(-1)^{\operatorname{Inv} k}$ part to the outside and then pull apart the $\sum$s:

$$\sum_k \left(\prod_i A_{i,k(i)}\right)\left(\sum_q (-1)^{\operatorname{Inv} q}(-1)^{\operatorname{Inv} k}\left(\prod_i B_{i,q(i)}\right)\right) = \left(\sum_k (-1)^{\operatorname{Inv} k}\prod_i A_{i,k(i)}\right)\left(\sum_q (-1)^{\operatorname{Inv} q}\prod_i B_{i,q(i)}\right)$$

$$= \det A \det B \quad \text{(*4.5 twice)}$$

YES!! △

5 The Cross Product and Its Generalization

5.1 The ordinary cross product

The cross product is a well-known vector operation, but it's slightly more complex than the ones discussed in Sections 2 and 3. Its usual version is defined as follows:

***5.1** *Definition.* The **cross product** of two 3-dimensional vectors $\mathbf{a}$ and $\mathbf{b}$ is the 3-dimensional vector

$$\mathbf{a} \times \mathbf{b} = \left\langle a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1\right\rangle.$$

□

Here is a convenient memory aid for the definition of the cross product:

***5.2**

$$\mathbf{a} \times \mathbf{b} = \det \begin{bmatrix} \mathbf{1}_1 & \mathbf{1}_2 & \mathbf{1}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

□

This isn't really a determinant because not all of its entries are real numbers: it contains a row of vectors. However, if one expands it symbolically, ignoring this issue, a vector is obtained (rather than a real number), and it is the correct answer. Let's define the determinant for "matrices" that have a row or column of vectors:

***5.3** *Definition.* Let A be a "matrix" with a single row or column that contains k-dimensional vectors instead of real numbers. Then $\det A$ is a k-dimensional vector, and $(\det A)_j$ is the determinant of the matrix obtained by replacing each vector in A with its jth component. □

With this definition, *5.2 becomes

$$\mathbf{a} \times \mathbf{b} = \left\langle \det \begin{bmatrix} 1 & 0 & 0 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}, \det \begin{bmatrix} 0 & 1 & 0 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}, \det \begin{bmatrix} 0 & 0 & 1 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \right\rangle,$$

which is what we want. It turns out that most of the properties of determinants still hold for these "vector determinants". One notable exception: it makes little sense to multiply matrices containing vectors, so there is no corresponding property of the determinant.

Here is an important theorem about vector determinants that will be necessary for proofs of several of the properties of the cross product:

***5.4** *Theorem.* To compute the dot product of a n-dimensional vector $\mathbf{a}$ with an vector determinant one row of which is $\mathbf{1}_1, \ldots, \mathbf{1}_n$, simply replace that row by the components of $\mathbf{a}$ and evaluate the resulting real determinant. □

Proof: Let R be the row of the vector determinant containing $\mathbf{1}_1, \ldots, \mathbf{1}_n$. When we expand this determinant using *5.3, we get a vector of determinants, and the ith determinant has for its row R the vector $\mathbf{1}_i$. When the dot product of this vector with $\mathbf{a}$ is taken, the ith determinant is multiplied by a_i; we may bring the constant into the determinant by changing row R to $a_i(\mathbf{1}_i)$. The dot product then sums the n determinants. However, these determinants are identical except for row R, so repeated use of *4.8 allows us to just add their rows R. The resulting row R is $\sum_i a_i(\mathbf{1}_i) = \mathbf{a}$, completing the proof. △

The cross product has many properties; here are some of the more important ones.

***5.5** *Theorem.* $\mathbf{a} \times \mathbf{b}$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. That is, $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\mathbf{a} \times \mathbf{b}) = 0$. □

Proof: Using *5.4,

$$\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) = \det \begin{bmatrix} a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}.$$

This determinant has two equal rows, so it is 0. The proof that $\mathbf{b} \cdot (\mathbf{a} \times \mathbf{b}) = 0$ is similar. △

In fact, if $\mathbf{a}$ and $\mathbf{b}$ are linearly independent (which means in this case that they are both zero and they are not parallel), then the vectors orthogonal to both are exactly the scalar multiples of $\mathbf{a} \times \mathbf{b}$.

***5.6** *Theorem.* $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin \theta(\mathbf{a})\mathbf{b}$. □

If the reader wishes to connect vectors and geometry (see Section 1.3), then this theorem states that $|\mathbf{a} \times \mathbf{b}|$ is the area of the parallelogram with sides formed by $\mathbf{a}$ and $\mathbf{b}$. (If $\mathbf{a}$ forms the base of the parallelogram, then $|\mathbf{a}|$ is the length of the base and $|\mathbf{b}| \sin \theta(\mathbf{a})\mathbf{b}$ is the height.)

The next two properties of the cross product follow directly from properties of determinants.

***5.7** *Theorem.* $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$. That is, interchanging the operands negates the cross product. □

***5.8** *Theorem.* $\mathbf{a} \times (s\mathbf{b} + t\mathbf{c}) = s(\mathbf{a} \times \mathbf{b}) + t(\mathbf{a} \times \mathbf{c})$ and $(s\mathbf{a} + t\mathbf{b}) \times \mathbf{c} = s(\mathbf{a} \times \mathbf{c}) + t(\mathbf{b} \times \mathbf{c})$. That is, the cross product is linear in either input vector. □

This one is a consequence of *5.3:

***5.9** *Theorem.* $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \operatorname{svol}(\mathbf{a}, \mathbf{b}, \mathbf{c})$. □

***5.10** *Theorem.* If the list $\mathbf{a} \times \mathbf{b}, \mathbf{a}, \mathbf{b}$ is linearly independent, it has positive handedness. □

Proof: $\operatorname{svol}(\mathbf{a} \times \mathbf{b}, \mathbf{a}, \mathbf{b}) = (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{b})$ by *5.9. By *3.25.4, $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{b}) \geq 0$, so if the list $\mathbf{a} \times \mathbf{b}, \mathbf{a}, \mathbf{b}$ has handedness, the handedness must be positive. △

Since coordinate systems in the real world are usually right-handed (especially in physics), there is an unfortunate tendency to state that the cross product obeys the "right-hand rule": one can point the thumb, index finger, and middle finger of the right hand in the directions of $\mathbf{a} \times \mathbf{b}$, $\mathbf{a}$, and $\mathbf{b}$ respectively. The correct statement is that the cross product obeys the right-hand rule in a right-handed coordinate system and the corresponding "left-hand rule" in a left-handed coordinate system.

5.2 Generalizing the cross product

Wouldn't it be nice if the cross product could be generalized to any number of dimensions? If possible, its generalization should have properties analogous to *5.5 through *5.10.

It's not immediately obvious how to generalize the formula in *5.1. The formula in *5.2 offers more possibilities because the determinant of any square matrix containing a row of vectors is defined by *5.3. If the vectors participating in the cross product are n-dimensional, $n-1$ of them must be used in order that the matrix be square. The generalized cross product is most logically defined as follows:

***5.11** *Definition.* The cross product of $n-1$ vectors $\mathbf{a}_i$ in n dimensions is

$$\mathop{\times}_{i=1}^{n-1}(\mathbf{a}_i) = \begin{vmatrix} \mathbf{1}_1 & \mathbf{1}_2 & \cdots & \mathbf{1}_n \\ (\mathbf{a}_1)_1 & (\mathbf{a}_1)_2 & \cdots & (\mathbf{a}_1)_n \\ \vdots & \vdots & & \vdots \\ (\mathbf{a}_{n-1})_1 & (\mathbf{a}_{n-1})_2 & \cdots & (\mathbf{a}_{n-1})_n \end{vmatrix}.$$

□

6 Components and Projections

7 Some Vector Calculus [empty]

8 Torques, Rotation, and Some Physics [empty]

9 Generalized Matrices [empty]

References

[1] Swokowski, Earl W. *Calculus with Analytic Geometry.*

[2] Hubbard, John H. and Barbara Burke. *Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach.*

[3] Baez, John. *The Octonions.* <http://math.ucr.edu/home/baez/Octonions/index.html>

Index of mathematical statements

Below are listed all mathematical statements in this document by number and page. Each statement has been given a short name or summary. A statement containing a *slanted term* is (or contains) a definition of that term.

The End