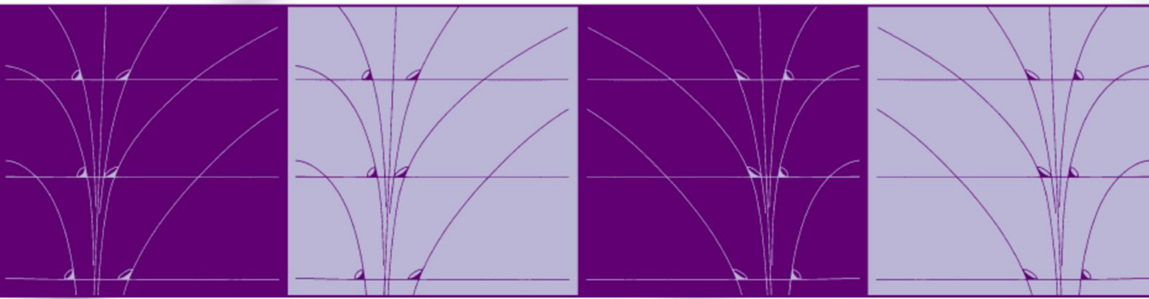
ROBERT H. WASSERMAN



Tensors and Manifolds | SECOND EDITION

With Applications to Physics

OXFORD

Tensors and Manifolds

Tensors and Manifolds

with Applications to Physics

Second edition

ROBERT H. WASSERMAN

Department of Mathematics
Michigan State University

OXFORD
UNIVERSITY PRESS

This book has been printed digitally and produced in a standard specification in order to ensure its continuing availability

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto
With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan South Korea Poland Portugal
Singapore Switzerland Thailand Turkey Ukraine Vietnam

Published in the United States
by Oxford University Press Inc., New York

Reprinted 2007

ISBN 978-0-19-851059-8

For Liz

PREFACE TO SECOND EDITION

In this edition I have corrected some errors, and expanded and clarified some of the exposition of the first edition. I have added a few problems and I have expanded the section on Notation.

The main change is the addition of 4 chapters extending the application of tensors and manifolds made in Chapters 15-17 on geometry to connections on fiber bundles and culminating with a brief description of gauge theory and the very elegant model of elementary particle physics based on this mathematics. In the new chapters I have tried to keep the mathematical level and style the same as in the first edition.

I would like to thank Prof. Wayne Repko for helpful discussions on elementary particle physics. Again I am very grateful to Cathy Friess for doing a remarkable job with the typing.

East Lansing, Michigan R. H. W.
June, 2003

PREFACE

This book is based on courses taken by advanced undergraduate and beginning graduate students in mathematics and physics at Michigan State University.

The courses were intended to present an introduction to the expanse of modern mathematics and its application in modern physics. The book gives an introductory perspective to young students intending to go into a field of pure mathematics, and who, with the usual "pigeon-holed" graduate curriculum, will not get an overall perspective for several years, much less any idea of application. At the same time, it gives a glimpse of a variety of pure mathematics for applied mathematics and physics students who will have to be carefully selective of the pure mathematics courses they can fit into their curriculum.

Thus, in brief, I have attempted to fill the gap between the basic courses and the highly technical and specialized courses that both mathematics and physics students require in their advanced training, while simultaneously trying to promote, at this early stage, a better appreciation and understanding of each other's discipline.

A third objective is to try to harmonize the two aspects which appear at this level, variously described on the one hand as the "classical," "index," or "local" approach, and on the other hand as the "modern," "intrinsic," or "global" approach.

An underlying theme is an emphasis on mathematical structures. To model a physical phenomenon in general we use some kind of mathematical "space" on which various "physical properties" are defined. For example, in fluid mechanics we make a model in which the fluid consists of points or regions of a space with a certain structure ("ordinary Euclidean space"), and pressure, rate of strain, etc. are mappings into other spaces with certain structures. In general, to model a physical phenomenon, $\mathbb{R}^n$ won't suffice for the domains and images of our mappings - we have to start with manifolds. Moreover, these manifolds have to have additional structures the basic ingredients of which are tensor algebras.

We begin with the algebraic structures we will need and go briefly to $\mathbb{R}^n$ with these in Chapter 7. Manifolds are introduced in Chapter 9 and these structures come together as tensor fields on manifolds in Chapter 11. Chapters 12-14 cover the rudiments of analysis on manifolds, and Chapters 15-17 are devoted to geometry. Finally, a modern treatment of the major ideas of classical analytical mechanics is given in Chapters 18-19, and the remaining chapters are dedicated to an exposition of special and general relativity.

A few words about terminology and notation are needed. I have tried to stick as closely as possible to the most popular current usages, sometimes inserting parenthetically strongly competing alternatives. However, since our text borrows from many different branches of mathematics and physics, we require terminology and notation for a very large variety of concepts and their interrelations. Consequently, the usual problem of how to tread between high precision and readability occurs in aggravated form. Sometimes dropping some notation which is really needed for precision can make it easier to read a given discussion and get the main ideas. On the other hand, sometimes keeping extra terminology and/or notation makes it easier by reminding us of certain important distinctions which might otherwise be temporarily forgotten. A separate section on "Terminology and Notation" is included for convenient reference to some of the conventions used in this book.

Finally, with respect to general style, I have endeavored to steer a safe course between the Scylla of rigor, and the Charybdis of informality. By my not being too heavy-handed with some of the details (including taking some notational liberties as indicated above), hopefully the young student will be able to sail through this passage to enjoy a panorama of interesting mathematics and physics.

This book owes a great deal to the efforts of many classes of students who struggled with earlier classnote versions, and helped hone it into this final form. I was also fortunate to have the services of excellent typists, in particular, Cathy Friess who did the entire final version.

East Lansing, Michigan R. H. W.
November, 1991

CONTENTS

1	**VECTOR SPACES**	**1**
	1.1 Definitions, properties, and examples	1
	1.2 Representation of vector spaces	5
	1.3 Linear mappings	6
	1.4 Representation of linear mappings	8
2	**MULTILINEAR MAPPINGS AND DUAL SPACES**	**11**
	2.1 Vector spaces of linear mappings	11
	2.2 Vector spaces of multilinear mappings	15
	2.3 Nondegenerate bilinear functions	19
	2.4 Orthogonal complements and the transpose of a linear mapping	20
3	**TENSOR PRODUCT SPACES**	**25**
	3.1 The tensor product of two finite-dimensional vector spaces	25
	3.2 Generalizations, isomorphisms, and a characterization	28
	3.3 Tensor products of infinite-dimensional vector spaces	31
4	**TENSORS**	**34**
	4.1 Definitions and alternative interpretations	34
	4.2 The components of tensors	36
	4.3 Mappings of the spaces V_s^r	39
5	**SYMMETRIC AND SKEW-SYMMETRIC TENSORS**	**46**
	5.1 Symmetry and skew-symmetry	46
	5.2 The symmetric subspace of V_s^0	48
	5.3 The skew-symmetric (alternating) subspace of V_s^0	54
	5.4 Some special properties of $\mathbf{S}^2(V^*)$ and $\bigwedge^2(V^*)$	59
6	**EXTERIOR (GRASSMANN) ALGEBRA**	**71**
	6.1 Tensor algebras	71
	6.2 Definition and properties of the exterior product	72
	6.3 Some more properties of the exterior product	77
7	**THE TANGENT MAP OF REAL CARTESIAN SPACES**	**84**
	7.1 Maps of real cartesian spaces	84
	7.2 The tangent and cotangent spaces at a point of $\mathbb{R}^n$	88
	7.3 The tangent map	96
8	**TOPOLOGICAL SPACES**	**102**
	8.1 Definitions, properties, and examples	102
	8.2 Continuous mappings	106

9 DIFFERENTIABLE MANIFOLDS **108**
9.1 Definitions and examples 108
9.2 Mappings of differentiable manifolds 116
9.3 The tangent and cotangent spaces at a point of M 119
9.4 Some properties of mappings 125

10 SUBMANIFOLDS **131**
10.1 Parametrized submanifolds 131
10.2 Differentiable varieties as submanifolds 133

11 VECTOR FIELDS, 1-FORMS, AND OTHER TENSOR FIELDS **136**
11.1 Vector fields 136
11.2 1-Form fields 143
11.3 Tensor fields and differential forms 146
11.4 Mappings of tensor fields and differential forms 149

12 DIFFERENTIATION AND INTEGRATION OF DIFFERENTIAL FORMS **153**
12.1 Exterior differentiation of differential forms 153
12.2 Integration of differential forms 158

13 THE FLOW AND THE LIE DERIVATIVE OF A VECTOR FIELD **168**
13.1 Integral curves and the flow of a vector field 168
13.2 Flow boxes (local flows) and complete vector fields 172
13.3 Coordinate vector fields 176
13.4 The Lie derivative 178

14 INTEGRABILITY CONDITIONS FOR DISTRIBUTIONS AND FOR PFAFFIAN SYSTEMS **186**
14.1 Completely integrable distributions 186
14.2 Completely integrable Pfaffian systems 191
14.3 The characteristic distribution of a differential system 192

15 PSEUDO-RIEMANNIAN GEOMETRY **198**
15.1 Pseudo-Riemannian manifolds 198
15.2 Length and distance 204
15.3 Flat spaces 208

16 CONNECTION 1-FORMS **212**
16.1 The Levi-Civita connection and its covariant derivative 212
16.2 Geodesics of the Levi-Civita connection 216
16.3 The torsion and curvature of a linear, or affine connection 219
16.4 The exponential map and normal coordinates 225
16.5 Connections on pseudo-Riemannian manifolds 226

17 CONNECTIONS ON MANIFOLDS **230**
17.1 Connections between tangent spaces 230
17.2 Coordinate-free description of a connection 231
17.3 The torsion and curvature of a connection 235
17.4 Some geometry of submanifolds 242

18 MECHANICS **248**
18.1 Symplectic forms, symplectic mappings, Hamiltonian vector fields, and Poisson brackets 248
18.2 The Darboux theorem, and the natural symplectic structure of T^*M 253
18.3 Hamilton's equations. Examples of mechanical systems 258
18.4 The Legendre transformation and Lagrangian vector fields 263

19 ADDITIONAL TOPICS IN MECHANICS **268**
19.1 The configuration space as a pseudo-Riemannian manifold 268
19.2 The momentum mapping and Noether's theorem 271
19.3 Hamilton-Jacobi theory 275

20 A SPACETIME **282**
20.1 Newton's mechanics and Maxwell's electromagnetic theory 282
20.2 Frames of reference generalized 287
20.3 The Lorentz transformations 289
20.4 Some properties and forms of the Lorentz transformations 294
20.5 Minkowski spacetime 298

21 SOME PHYSICS ON MINKOWSKI SPACETIME **306**
21.1 Time dilation and the Lorentz-Fitzgerald contraction 306
21.2 Particle dynamics on Minkowski spacetime 313
21.3 Electromagnetism on Minkowski spacetime 317
21.4 Perfect fluids on Minkowski spacetime 322

22 EINSTEIN SPACETIMES **326**
22.1 Gravity, acceleration, and geodesics 326
22.2 Gravity is a manifestation of curvature 328
22.3 The field equation in empty space 331
22.4 Einstein's field equation (Sitz, der Preuss Acad. Wissen., 1917) 334

23 SPACETIMES NEAR AN ISOLATED STAR **339**
23.1 Schwarzschild's exterior solution 339
23.2 Two applications of Schwarzschild's solution 344
23.3 The Kruskal extension of Schwarzschild spacetime 348
23.4 The field of a rotating star 352

24 NONEMPTY SPACETIMES **356**
24.1 Schwarzschild's interior solution 356
24.2 The form of the Friedmann-Robertson-Walker metric tensor and its properties 361
24.3 Friedmann-Robertson-Walker spacetimes 365

25 LIE GROUPS **369**
25.1 Definition and examples 369
25.2 Vector fields on a Lie group 371
25.3 Differential forms on a Lie group 377
25.4 The action of a Lie group on a manifold 380

26 FIBER BUNDLES **384**
26.1 Principal fiber bundles 384
26.2 Examples 388
26.3 Associated bundles 390
26.4 Examples of associated bundles 392

27 CONNECTIONS ON FIBER BUNDLES **394**
27.1 Connections on principal fiber bundles 394
27.2 Curvature 398
27.3 Linear Connections 401
27.4 Connections on vector bundles 404

28 GAUGE THEORY **409**
28.1 Gauge transformation of a principal bundle 409
28.2 Gauge transformations of a vector bundle 413
28.3 How fiber bundles with connections form the basic framework of the Standard Model of elementary particle physics 417

References **423**

Notation **427**

Index **435**

1

VECTOR SPACES

Since tensor algebras are built from vector spaces we will recall some of the theory of the latter. We will review the basic properties of vector spaces, their representations and mappings, and mention a generalization.

1.1 Definitions, properties, and examples

Definition If V is an Abelian group with elements $v, w, \ldots$, if $a, b, \ldots$ are elements of a field, $\mathbb{K}$, and if a mapping $\mathbb{K} \times V \longrightarrow V$ called "scalar multiplication" and denoted by

$$(a, v) \mapsto av \in V$$

is defined for all elements $a \in \mathbb{K}$ and all $v \in V$ such that

$$\begin{aligned} 1v &= v \\ a(v+w) &= av + aw \\ (a+b)v &= av + bv \\ (ab)v &= a(bv) \end{aligned}$$

then V is a *vector space.*

From these properties one can prove the additional properties

$$\begin{aligned} 0v = 0 \quad &\text{for all} \quad v \in V \\ a0 = 0 \quad &\text{for all} \quad a \in \mathbb{K} \end{aligned}$$

and, conversely, if $av = 0$ then either $a = 0$ or $v = 0$.

Definitions Let S be any set (not necessarily finite) of elements in a vector space, V. Then the intersection of all subspaces of V which contain S and the set of all (finite) linear combinations of elements of S are the same. This set is a subspace, $< S >$, of V called the *linear closure of* S. If a subspace, W, of V is the linear closure of some set $S \subset V$ then we say S *spans*, or *generates* W, and S is *a set of generators* of W.

Definitions Let $\{W_i\}$ be any set (not necessarily finite) of subspaces of a vector space, V. Then the intersection of all subspaces of V containing $\bigcup W_i$, and the set of all finite sums of the form $w_j + w_k + \cdots + w_p$ where $w_j \in W_j$, $w_k \in W_k, \dots, w_p \in W_p$ are the same. This set is a subspace, $\sum W_i$, of V called the *sum of the subspaces, W_i of V*. $\sum W_i$ is *direct* if $W_j \cap \sum_{k \neq j} W_k = 0$ for all j. (Note, there is no restriction on the cardinality of the index set we are using.)

Definitions If for all (finite) sums $\sum_i a_i v_i$ with $v_i \in S$ we have that $\sum_i a_i v_i = 0$ implies that $a_i = 0$ for all i, then S is *a linearly independent set.* This property is equivalent to $0 \notin S$ and every finite sum of subspaces of the form $\ell_v = \{av : a \in \mathbb{K}\}$ is direct. Finally, if W is a subspace of V, and if S spans W, and is a linearly independent set, then S is *a basis* of W. This property is equivalent to the property that for each $w \in W$ corresponding to each $v_i \in S$ there exists a unique element $a_i \in \mathbb{K}$ (with $a_i = 0$ for all but a finite number of i's) such that

$$w = \sum_i a_i v_i \tag{1.1}$$

Clearly, these definitions allow for the possibility that a vector space may contain an infinite linearly independent set, and an infinite basis.

EXAMPLES. (i) $\{(a^1, a^2, \dots, a^n, \dots) : a^i \in \mathbb{R}\}$ with "component wise" addition and scalar multiplication with $\mathbb{R}$ is a vector space over $\mathbb{R}$. The set $S = \{(0, 0, \dots, 1, 0, \dots) : 1 \text{ is in the } i\text{th place}, i = 1, 2, 3, \dots\}$ is a linearly independent set but is not a basis.

(ii) $\{(a^1, a^2, \dots, a^n, \dots) : \text{only a finite number of the } a\text{'s are nonzero}\}$ with operations as in (i) is a vector space and now S is a basis.

(iii) The set of continuous real-valued functions on $[0, \pi]$ with the usual operations: $\{\sin nx : n = 1, 2, \dots\}$ is a linearly independent set but not a basis.

(iv) The set of all ordered n-tuples of real numbers, $\mathbb{R}^n$, is an important example of a vector space having a finite basis. $\{(0, 0, \dots, 1, \dots, 0) : 1 \text{ in the } i\text{th place}, i = 1, \dots, n\}$ is called the *natural basis* of $\mathbb{R}^n$.

(v) As generalizations of (i) – (iv) we can consider the set $\mathbb{R}^S$, the set of all (real-valued) functions on an arbitrary set, S, and $\mathbb{R}^S_0$ the set of all (real-valued) functions on an arbitrary set which vanish at all but a finite number of points. Again with pointwise addition and multiplication with $\mathbb{R}$ these are both vector spaces over $\mathbb{R}$. $\mathbb{R}^S_0$ is called *the free vector space over S*. For each $x \in S$ we can define the function f_x by

$$f_x(y) = \begin{cases} 1 & \text{when } y = x \\ 0 & \text{when } y \neq x \end{cases}.$$

Then we have an injection $i : S \to \mathbb{R}^S$ defined by $x \mapsto f_x$. The image set of i, $i(S) \subset \mathbb{R}^S$ is a linearly independent set. It is a basis of $\mathbb{R}^S_0 \subset \mathbb{R}^S$ but not of $\mathbb{R}^S$. Every $f \in \mathbb{R}^S_0$ has the unique representation

$$f = \sum_{x \in S} f(x) f_x$$

(Only a finite number of terms of this sum are nonzero). $\mathbb{R}_0^S$ will be important in the construction of tensor product spaces in Chapter 3.

Theorem 1.1 *If $\boldsymbol{I}$ is a linearly independent set, and $\boldsymbol{G}$ is a set of generators of a vector space, V, and $\boldsymbol{I} \subset \boldsymbol{G}$, then there exists a basis $\boldsymbol{B}$ such that $\boldsymbol{I} \subset \boldsymbol{B} \subset \boldsymbol{G}$. In other words, every linearly independent set can be extended to a basis, and every set of generators contains a basis.*

Proof (1) If $\boldsymbol{G}$ is finite, we can let $\boldsymbol{G} = \{v_1, \ldots, v_n\}$ and proceed by induction on n. That is, let v_k be any element of $\boldsymbol{G}$ and consider the subspace, W, generated by $\boldsymbol{G} - v_k$. By the induction hypothesis, W has a basis $\boldsymbol{B}$ with $\boldsymbol{B} \subset \boldsymbol{G} - v_k$. If $v_k \in W$, then $W = V$ and we are done. If $v_k \notin W$, then $\boldsymbol{B} \cup v_k$ is a linearly independent set and spans V and is thus a basis of V.

(2) If $\boldsymbol{G}$ is infinite, we can use Zorn's Lemma, Problem 1.5. □

In particular, since every (nontrivial) vector space has a linearly independent set and a set of generators with $\boldsymbol{I} \subset \boldsymbol{G}$, every vector space has a basis.

Note, however, that for many important spaces the bases are uncountable. In particular, every basis of the vector space of example (i) is uncountable (for method of proof, see Problem 1.12), every basis of the vector space of example (iii) is uncountable (it is an infinite-dimensional Banach space with norm $\max\limits_{[0,\pi]} f$), and if X is an infinite set, then the bases of example (v) are uncountable.

In all the examples above, the field, $\mathbb{K}$, of scalars was $\mathbb{R}$, the field of real numbers. Many of the subsequent results are valid for an arbitrary field (of characteristic 0), in particular for the field of complex numbers, $\mathbb{C}$. It is often useful to bring in complex vector spaces in physical applications. Relations between real and complex vector spaces are indicated in Problems 1.7, 1.8, and 1.14. However, with rare exceptions all our vector spaces in the following will be real.

There is a generalization of the concept of a vector space which we will need when we get to vector fields in Chapter 11. In the definition of vector space we simply replace the field, $\mathbb{K}$, by a ring, $\mathbb{A}$ (with a unit). Such a structure is called an $\mathbb{A}$-*module.* All the definitions above in their various equivalent forms are valid for $\mathbb{A}$-modules. However, there are $\mathbb{A}$-modules that do not have a basis. For example, the module consisting of the elements $0, v_1, v_2, v_3$ over the integers, and with addition given by $v_i + v_i = 0, i = 1, 2, 3$ and $v_i + v_j = v_k$ for i, j, k all different, contains no linearly independent set, and hence has no basis. This example also shows that the property stated right after the definition of a vector space is not valid for modules.

An $\mathbb{A}$-module with a basis is called *a free module.* The $\mathbb{A}$-module we will be studying has a basis. There, $\mathbb{A}$ will be the algebra of (real-valued) functions on a certain set M; i.e., $\mathbb{R}^S$ with the structure of an algebra and with $S = M$.

Definition If a vector space V has a linearly independent set of n elements and no linearly independent set of $n+1$ elements, then V *has dimension* n, or $\dim V = n$.

Theorem 1.2 *If V has dimension n, then every linearly independent set of n elements is a basis, and every basis has exactly n elements.*

Proof Problem 1.6. □

(Note: Theorem 1.2 can be extended to the "infinite-dimensional" case. In particular, every two bases of V have the same cardinality (cf. Lang, 1965, p. 86).)

PROBLEM 1.1. In the definition of the linear closure of S we actually gave two different characterizations of this concept. Prove they are the same.

PROBLEM 1.2. The same as Problem 1.1 for the definition of the sum of subspaces.

PROBLEM 1.3. The same as Problem 1.1 for the definition of linearly independent set.

PROBLEM 1.4. The same as Problem 1.1 for the definition of basis.

PROBLEM 1.5. Prove Theorem 1.1 for the case $\boldsymbol{G}$ is an infinite set (cf., Greub, 1981, pp. 12 - 13).

PROBLEM 1.6. Prove Theorem 1.2.

PROBLEM 1.7. The set of all ordered n-tuples of complex numbers, $\mathbb{C}^n$, with "component-wise" addition and scalar multiplication with $\mathbb{C}$ is an n-dimensional vector space with the natural basis $S = \{(0,\ldots,1,0,\ldots,0) : 1$ is in the kth place, $k = 1,\ldots,n\}$ (cf., Example (iv)). Show that the *same set* of n-tuples is a real vector space with basis $S \cup \{(0,\ldots,i,0,\ldots,0) : i = \sqrt{-1}$ is in the kth place, $k = 1,\ldots,n\}$.

PROBLEM 1.8. Starting with a real vector space, V, we can construct a complex vector space, V_c, *the complexification of* V, whose elements are the elements of $V \times V$ with "component-wise" addition and with scalar multiplication defined by $z(v,w) = (av - bw, aw + bv)$ where $z = a + ib$. Show that $\dim V = \dim V_c$ (cf., Problem 1.14).

1.2 Representation of vector spaces

For an n-dimensional vector space, V^n (over $\mathbb{R}$), Theorem 1.2 implies that in the representation (1.1) we can choose the same finite set of basis elements v_i for all $v \in V^n$ so that (1.1) defines a $1-1$ correspondence between elements $v \in V^n$ and ordered n-tuples $(a^1, \ldots, a^n) \in \mathbb{R}^n$. In particular, $v_i \mapsto (0, \ldots, 1, \ldots, 0)$ with 1 in the ith place. With different bases we have different correspondences. That is, once a basis is chosen we can represent any n-dimensional vector space by the particular space, $\mathbb{R}^n$, of example (iv) above.

If $\{e_1, \ldots, e_n\}$ and $\{\bar{e}_1, \ldots, \bar{e}_n\}$ are two bases for V^n, the relation between them can be written in the form

$$\bar{e}_j = \sum_i a^i_j e_i \tag{1.2}$$

For any $v \in V^n$ we have $v = \sum v^i e_i$ and $v = \sum \bar{v}^j \bar{e}_j$. Substituting (1.2) into the second representation for v, and then comparing the two expressions for v, we see that

$$v^i = \sum_j \bar{v}^j a^i_j \tag{1.3}$$

Let

$$(a^i_j) = \begin{pmatrix} a^1_1 & a^1_2 & \cdots & a^1_n \\ & & & \\ a^2_1 & & & \vdots \\ \vdots & & & \\ a^n_1 & \cdots & & a^n_n \end{pmatrix}$$

Then (1.3) can be written in matrix notation as either

$$\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} = (a^i_j) \begin{pmatrix} \bar{v}^1 \\ \vdots \\ \bar{v}^n \end{pmatrix}$$

or

$$\begin{pmatrix} \bar{v}^1 \\ \vdots \\ \bar{v}^n \end{pmatrix} = (a^i_j)^{-1} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \tag{1.4}$$

Writing (1.2) in "matrix notation" as

$$\begin{pmatrix} \bar{e}_1 \\ \vdots \\ \bar{e}_n \end{pmatrix} = (a^i_j)^{tr} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \tag{1.5}$$

we see that the systems of basis vectors transform with the inverse transpose of the matrix of the transformation of the *components* of a vector. We call (a^i_j) *the change-of-basis* matrix.

Observe that we are using superscripts as well as subscripts in our notation. This enables us to use the "summation convention", which we will do from now on. That is, if in products, such as on the right side of (1.2), the same index occurs as a superscript and a subscript we will sum on that index omitting the Σ notation. Also, note that though in the matrix notation introduced above the superscript is the row index and the subscript is the column index this will not always be the case.

PROBLEM 1.9. Show that the matrix (a^i_j) defined above is nonsingular.

PROBLEM 1.10. If V^n is the direct sum of two subspaces and a change of basis is made in each of the subspaces, what is the form of the change of basis matrix for V^n?

1.3 Linear mappings

Definitions If V and W are two vector spaces, a mapping $\phi : V \to W$ is a *linear mapping* if $\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2)$ and $\phi(av) = a\phi(v)$ for all v_1, v_2, and v in V and a in the common field of V and W. For such a mapping we write $\phi(v) = \phi \cdot v$. $\phi(V)$ is the *image space* of V under ϕ, $\dim \phi(V)$ is the *rank* of ϕ, and $\phi^{-1}(0)$ is the inverse image of $0 \in W$, or the *null space*, N_ϕ, of ϕ, or the *kernel*, $\ker \phi$, of ϕ. $\phi(V)$ is a subspace of W and $\phi^{-1}(0)$ is a subspace of V. If $W = \mathbb{R}$, then ϕ is called *a linear function*, or *a linear functional*, or *a linear form*, or *a covector*. If $W = V$ then ϕ is called *a linear transformation* or *a linear operator*.

Again, we can generalize the preceding and much of the following to $\mathbb{A}$-modules.

We can construct many different linear mappings between two vector spaces (or, $\mathbb{A}$-modules) as the following theorem shows.

Theorem 1.3 *Given vector spaces V and W and a basis S of V, then there exists a unique linear mapping, $\phi : V \to W$ such that $\phi \cdot v_i = w_i$ where $v_i \in S$ and w_i are arbitrarily chosen elements of W.*

Proof (i) Define ϕ by "extending the given conditions linearly", that is, let $\phi(v) = a^i w_i$ for $v = a^i v_i$. Then it is easy to verify the properties required by the definition. (ii) If ψ is another linear mapping such that $\psi \cdot v_i = w_i$, then by linearity $\psi \cdot v = a^i w_i$ for $v = a^i v_i$. Hence $\psi \cdot v = \phi \cdot v$ for all v so $\psi = \phi$. □

Corollary *If $\dim V = \dim W$, then V and W are isomorphic.*

The converse of this corollary is also true; if ϕ is an isomorphism of V and W then they have the same dimensions. More generally, we have the following result for the dimension of V.

Theorem 1.4 *If $\phi : V \to W$ is linear, then*

$$\dim V = \text{rank } \phi + \dim \ker \phi$$

Proof Let $\{e_1, \dots, e_p\}$ be a basis of $\phi^{-1}(0)$, and let $\{e_1, \dots, e_p, \bar{e}_1, \dots, \bar{e}_q\}$ be a basis of V. (See Problem 1.5.) Then $\dim \phi^{-1}(0) = p$ and $\dim V = p + q$. Now we need only to show that $\{\phi \cdot \bar{e}_1, \dots, \phi \cdot \bar{e}_q\}$ is a basis of $\phi(V)$, that is that $\dim \phi(V) = q$. A direct calculation shows that $\phi \cdot \bar{e}_1, \dots, \phi \cdot \bar{e}_q$ is a linearly independent set and spans $\phi(V)$. □

If W is a subspace of V, and $v \in V$, we can form the subset $\{v + w : w \in W\}$. The set of all such subsets (as v ranges over V) is a *factor*, or *quotient space*, V/W. v_1 and v_2 are both in the same subset if and only if $v_1 - v_2 \in W$. The set (i.e., equivalence class) containing v is denoted by $[v]$. The dimension of V/W is called *the codimension of* W. The linear mapping $\pi : V \to V/W$ given by $v \mapsto [v]$ is onto, has the property $\pi^{-1}(0) = W$, and is called *the natural projection* of V onto V/W. From this and Theorem 1.4 we have codim $W = \dim V - \dim W$.

Conversely, starting with a linear mapping, $\phi : V \to W$, we have $\phi^{-1}(0)$ and we can construct a factor space $V/\phi^{-1}(0)$ and an induced isomorphism $V/\phi^{-1}(0) \cong \phi(V)$. There is also the factor space $W/\phi(V)$ called *the cokernel of* ϕ, and rank $\phi + \dim \text{coker } \phi = \dim W$.

PROBLEM 1.11. Let V and W be vector spaces. (i) Define operations in $V \times W$ so that it becomes a vector space (called *the exterior direct sum of* V *and* W and denoted by $V \oplus W$). (ii) Show that the *projections* given by $p_1 : (v, w) \mapsto (v, 0)$ and $p_2 : (v, w) \mapsto (0, w)$ are linear. (iii) Show that $V \oplus W = \ker p_1 \oplus (p_1(V \oplus W))$, where the first $\oplus$ on the right is the direct sum defined in Section 1.1.

PROBLEM 1.12. The set of linear functions, $\mathcal{L}(V, \mathbb{R})$, on V is a vector space (this is a special case of Theorem 2.1). Use Theorem 1.3 to show that if V has a countably infinite basis, then the basis of $\mathcal{L}(V, \mathbb{R})$ is uncountably infinite.

(Hint: By Theorem 1.3, if V has a countably infinite basis, then every sequence of real numbers determines a unique element $f \in \mathcal{L}(V, \mathbb{R})$. For each positive number, r, let f_r be the element determined by $(1, r, r^2, r^3, \dots)$. Consider the set of all such functions, and show that for any finite subset $\{f_1, \dots, f_n\}$ $a^1 f_1 + a^2 f_2 + \cdots + a^n f_n = 0$ implies $a^1 = a^2 = \cdots = a^n = 0$.)

PROBLEM 1.13. If $(a, b) \in \mathbb{R}^2$, let $W = \{t(a, b) : t \in \mathbb{R}\}$, a subset of the vector space $\mathbb{R}^2$. Give a geometrical interpretation of $\mathbb{R}^2/W$.

PROBLEM 1.14. Corresponding to each $a \in \mathbb{K}$ there is a linear operator ϕ on V given by $\phi : V \to aV$. In particular, for the space $\mathbb{C}^n$, corresponding to $i = \sqrt{-1}$ we have a linear operator, J with the property that $J^2 v = -v$. On the real vector space $\mathbb{R}^{2n}$, the linear operator J defined by

$$J : \begin{cases} e_k \mapsto e_{k+n} & k = 1, \ldots, n \\ e_k \mapsto e_{k-n} & k = n+1, \ldots, 2n \end{cases}$$

where $\{e_1, \ldots, e_{2n}\}$ is the natural basis of $\mathbb{R}^{2n}$, has this property. Show that for any real vector space which admits such an operator, called *a complex structure*, the same set of elements can be made into a complex vector space by defining multiplication by a complex scalar by $(a + bi)v = av + bJv$ (cf., Problems 1.7 and 1.8).

PROBLEM 1.15. A set, S, on which there is defined a vector space, V, of transformations is called *an affine space* of V, if (i) $0 \in V$ is the identity mapping, if (ii) for $v \neq 0$ and $p \in S, v(p) \neq p$, if (iii) $u + v$ is the composition, and if (iv) for each ordered pair (p, q) of elements of S, there is an element $v \in V$ such that $v(p) = q$. Show that $(-v)(q) = p$ if $v(p) = q$. Show that (p, q) determines a unique vector, so we can write it as $\overrightarrow{pq}$ and call it a *free vector*, or *translation of* S. If $S = \{(a, b, c) \in \mathbb{R}^3 : c = a + b + 1\}$ and V is the vector space $\mathbb{R}^2$, show that S is an affine space of $\mathbb{R}^2$ (cf., Problem 1.13).

1.4 Representation of linear mappings

Just as we have a $1-1$ correspondence between vectors in a real n-dimensional vector space, V^n, and elements of $\mathbb{R}^n$ once a basis is chosen in V^n, we can now, after choosing bases $\{e_1, \ldots, e_n\}$ in V^n and $\{E_1, \ldots, E_r\}$ in V^r, set up a $1-1$ correspondence between the set of linear mappings from V^n to V^r and the set of r by n matrices. Thus, given a linear mapping, ϕ, then for each $i = 1, \ldots, n$, there is a unique set of components, $\phi_i^j, j = 1, \ldots, r$, of the image of e_i under ϕ. That is,

$$\phi \cdot e_i = \phi_i^j E_j \tag{1.6}$$

If $w = \phi \cdot v$, then putting $v = v^i e_i$ and $w = w^j E_j$ we have $w^j E_j = w = \phi \cdot v^i e_i = v^i \phi \cdot e_i = v^i \phi_i^j E_j$, so

$$w^j = v^i \phi_i^j \qquad \begin{matrix} i = 1, \ldots, n \\ j = 1, \ldots, r \end{matrix} \tag{1.7}$$

or, in matrix notation, writing

$$(\phi_i^j) = \begin{pmatrix} \phi_1^1 & \phi_2^1 & \cdots & \phi_n^1 \\ \phi_1^2 & & & \\ \vdots & & & \\ \phi_1^r & & \cdots & \phi_n^r \end{pmatrix}$$

we have

$$\begin{pmatrix} w^1 \\ \vdots \\ w^r \end{pmatrix} = (\phi_i^j) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \tag{1.8}$$

(Notice that the matrix of coefficients of the system (1.6), as it stands, is the transpose of (ϕ_i^j). Also, compare eqs. (1.6) - (1.8) respectively with eqs. (1.2) - (1.4).)

If $n = r$, then the mapping can be represented by a square matrix. For example, if $V^n = V^r = \mathbb{R}^{2n}$, and ϕ is the linear operator J in Problem 1.14, then in the natural basis of $\mathbb{R}^{2n}$ $\phi = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ where I is the $m \times m$ identity matrix.

If $r = 1$, then the mapping can be represented by a $1 \times n$ matrix, or a row vector. Thus, linear functions, or covectors, can be represented by row vectors.

Conversely, starting with an arbitrary r by r matrix, (ϕ_i^j), we can let the columns be the components, in the $\{E_i\}$ basis of V^r, of the images of the basis elements, e_i, of V^n, i.e., $e_i \mapsto \phi_i^1 E_1 + \phi_i^2 E_2 + \cdots + \phi_i^r E_r$. An arbitrary choice of these images determines a unique linear mapping by Theorem 1.3.

With the representation of vectors by elements of cartesian spaces and the representation of linear mappings by matrices, we see that the specific "concrete" examples of linear mappings of the form $\phi : \mathbb{R}^n \to \mathbb{R}^r$ given by

$$(v^1, \ldots, v^n) \mapsto (\phi_i^1 v^i, \ldots, \phi_i^r v^i) \tag{1.9}$$

or, in matrix notation

$$\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \mapsto \begin{pmatrix} \phi_1^1 & \cdots & \phi_n^1 \\ \vdots & & \vdots \\ \phi_1^r & \cdots & \phi_n^r \end{pmatrix} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

are really all there are.

Just as a vector will have two different sets of components in two different bases, a linear mapping will have different matrices if different bases are chosen in V^n and V^r.

Definition Two r by n matrices are *equivalent* if they represent the same linear mapping ϕ, relative to different bases in V^n and V^r.

From matrix theory we know that two matrices are equivalent if and only if they have the same rank.

An important special case of the ideas above occurs when $V^r = V^n$, i.e., ϕ is a linear transformation. In this case, two n by n matrices representing ϕ are called *similar*, and matrices (ϕ^i_j) and (ψ^i_j) are similar if and only if there exists a nonsingular matrix, M such that $(\psi^i_j) = M^{-1}(\phi^i_j)M$. M is the change-of-basis matrix. For a linear transformation we also have the important concepts of eigenvectors, eigenvalues, invariant subspaces, etc. We will come back to these ideas later when we need them in Chapter 5.

PROBLEM 1.16. (i) Show that the correspondence $e_i \mapsto (0, \ldots 1, \ldots, 0)$ between V^n and $\mathbb{R}^n$ described in Section 1.2 is an isomorphism. (ii) What is the form of (1.8) for this linear mapping if we choose in $\mathbb{R}^n$ (a) the natural basis, (b) an arbitrary basis $\{(a_{i1}, \ldots, a_{in}) : i = 1, \ldots, n\}$?

2

MULTILINEAR MAPPINGS AND DUAL SPACES

We will discuss the important space, V^*, of linear functions on a vector space, V. We will describe isomorphisms between spaces of multilinear mappings, and, finally, we will focus on special properties of bilinear functions.

2.1 Vector spaces of linear mappings

In Section 1.3 we discussed briefly the idea of a linear mapping between vector spaces V and W. Now we consider the set of all linear mappings from V to W.

Theorem 2.1 *The set, $\mathcal{L}(V, W)$, of all linear mappings from V to W with operations $\phi + \psi$ and $a\phi$ defined by*

$$(\phi + \psi) \cdot v = \phi \cdot v + \psi \cdot v$$
$$(a\phi) \cdot v = a(\phi \cdot v)$$

is a vector space.

Proof $\phi + \psi$ is a linear mapping since

$$\begin{aligned}(\phi + \psi) \cdot (av + bw) &= \phi \cdot (av + bw) + \psi \cdot (av + bw) \\ &= a\phi \cdot v + b\phi \cdot w + a\psi \cdot v + b\psi \cdot w \\ &= a(\phi + \psi) \cdot v + b(\phi + \psi) \cdot w.\end{aligned}$$

Similarly for $a\phi$. Each of the vector space properties of $\mathfrak{L}(V, W)$ comes from the corresponding property of W. $\square$

Theorem 2.2 *With the standard definitions of addition and scalar multiplication, the set of r by n matrices is a vector space, and the $1-1$ correspondence described in Section 1.4 between $\mathfrak{L}(V^n, V^r)$ and the vector space of r by n matrices is an isomorphism.*

Proof Problem 2.2. $\square$

Thus, in particular, the set of $n \times n$ matrices, $\mathcal{M}_n$, forms a vector space isomorphic to the vector space of linear transformations of V^n.

Theorem 2.3 *Suppose $\{e_i\}$ and $\{E_j\}$ are bases of finite-dimensional spaces V and W respectively. Then the elements e^i_j of $\mathcal{L}(V, W)$ defined by $e^i_j : v \mapsto v^i E_j$ form a basis for $\mathcal{L}(V, W)$. Moreover, for $\phi \in \mathfrak{L}(V, W)$,*

$$\phi = \phi^j_i e^i_j$$

where ϕ^i_j are given by $\phi \cdot e_k = \phi^j_k E_j$.

Proof First of all, note that the definition of e^i_j is equivalent to $e^i_j : e_k \mapsto \delta^i_k E_j$ with e^i_j extended to V by linearity.

(i) $a^j_i e^i_j(e_k) = a^j_i \delta^i_k E_j = a^j_k E_j$ so $a^j_i e^i_j = 0$ implies that $a^j_k E_j = 0$ for all k, and since $\{E_i\}$ is linearly independent, $a^j_i = 0$.

(ii) Given ϕ in $\mathcal{L}(V, W)$, let $\phi \cdot e_k = \phi^j_i E_j$. Then $\phi^j_i e^i_j(e_k) = \phi^j_i \delta^i_k E_j = \phi^j_k E_j = \phi(e_k)$ so $\phi = \phi^j_i e^i_j$. □

There are two important special cases of $\mathcal{L}(V, W)$; namely, $\mathcal{L}(\mathbb{R}, W)$ and $\mathcal{L}(V, \mathbb{R})$. We will devote our attention exclusively to the second case after giving one result for the first.

Theorem 2.4 *$\mathfrak{L}(\mathbb{R}, W)$ is isomorphic with W.*

Proof Note that every nonzero element of $\mathbb{R}$ is a basis of $\mathbb{R}$. In particular, 1 is the natural basis of $\mathbb{R}$. By Theorem 1.3 for each $w \in W$, there is an element $\bar{w}$ of $\mathcal{L}(\mathbb{R}, W)$ given by $\bar{w} : 1 \mapsto w$. Two different w's must clearly lead to two different mappings, so the correspondence $W \to \mathcal{L}(\mathbb{R}, W)$ is one-to-one. On the other hand, given any $\phi \in \mathcal{L}(\mathbb{R}, W), \phi \cdot 1 \in W$ determines a mapping ψ given by $1 \mapsto \phi \cdot 1$. By the "uniqueness" part of Theorem 1.3 $\psi = \phi$ so the correspondence is onto. The linearity follows from the definition of the operations in $\mathcal{L}(\mathbb{R}, W)$. □

This isomorphism will be invoked to make a certain "identification" when we study tangent maps in Section 7.3.

The space, $\mathfrak{L}(V, \mathbb{R})$, of linear functions on V is also denoted by V^*. As noted in Section 1.3, its elements, f, are also called linear forms (or linear functionals, or covectors).

As a concrete *example* of a space of linear forms we have $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$, whose elements, f, are given by $f : (a^1, \ldots, a^n) \mapsto f_i a^i$, or, in matrix form,

$$\begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix} \mapsto (f_1 \cdots f_n) \begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix}$$

Clearly the elements of $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ are just special cases of the linear mappings, (1.9). As there, this example can be thought of as being quite general, in the sense that once we choose a basis, $\{e_i\}$, in any n-dimensional space, V^n, and a basis, $E \in \mathbb{R}$, then $v \in V^n$ can be represented by $\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$, $f \in V^{n*}$ can be represented by a $1 \times n$ matrix (or row vector), $(f_1, \cdots f_n)$, where the f_i are given by $f \cdot e_i = f_i E$ (cf., eq. (1.6)), and

$$(f_1 \cdots f_n) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

represents the image of v under f.

From the fact that V^{n*} is isomorphic to the set of $1 \times n$ matrices, it is clear the $\dim V^{n*} = n$; i.e., $\dim V^{n*} = \dim V^n$. It follows as a special case of Theorem 2.3 that the elements $\varepsilon^i \in V^{n*}$ defined by $\varepsilon^i \cdot e_k = \delta^i_k E \quad i, k = 1, \ldots, n, \quad E \in \mathbb{R}$, form a basis for V^{n*} and

$$f = f_i \varepsilon^i$$

for any f in V^{n*}.

Definition If we choose $E = 1$, then the basis $\{\varepsilon_i\}$ of V^{n*} given by

$$\varepsilon^i \cdot e_k = \delta^i_k, \qquad i, k = 1, \ldots, n$$

is called *the dual basis* of $\{e_i\}$.

Since, when V is finite-dimensional, $\dim V = \dim V^*$, it follows that we have a result for $\mathfrak{L}(V, \mathbb{R})$ analogous to Theorem 2.4 for $\mathfrak{L}(\mathbb{R}, W)$; namely, $\mathfrak{L}(V, \mathbb{R})$ is isomorphic to V.

In general, however, if V is not finite-dimensional, V^* is not necessarily isomorphic with V. It is generally larger than V. See Problem 1.12. However, there is a subspace of V^* which is isomorphic with V.

Definition Let $\{e_i\}$ be a basis of V. We denote the set $\{f \in V^* : f \cdot e_i = 0$ for all but a finite number of the e_i's$\}$ by V_0^*.

Theorem 2.5 *The set, $\{\varepsilon^i\}$, of elements of V^* defined by $\varepsilon_i \cdot e_k = \delta^i_k$ is a basis for V_0^*, and $V_0^* \cong V$.*

Proof (i) Consider any linear combination $a_i \varepsilon^i$. Then

$$(a_i \varepsilon^i) \cdot e_k = a_i(\varepsilon^i \cdot e_k) = a_k$$

so $a_i \varepsilon^i = 0$ implies that $a_k = 0$. (ii) Given f in V_0^*, let $f \cdot e_i = f_i$, and look at $f_i \varepsilon^i$. $(f_i \varepsilon^i) \cdot e_j = f_i \delta^i_j = f_j = f \cdot e_j$ for all e_j, so $f = f_i \varepsilon^i$. (i) and (ii) show that $\{\varepsilon^i\}$ is a basis for V_0^*. (iii) Notice that the condition $\varepsilon^i \cdot e_k = \delta^i_k$ defines a $1 - 1$ correspondence between the basis sets $\{e_i\}$ and $\{\varepsilon^i\}$; namely, $e_k \mapsto \varepsilon^i$, the basis element of V_0^* which takes e_i to 1 and all the others to zero. If we extend this correspondence by linearity, then the resulting mapping is $1 - 1$ and onto V_0^*. □

Having constructed a new vector space V^* from V we can now ask about linear transformations on V^*; in particular, we can study the set $\mathfrak{L}(V^*, \mathbb{R}) = V^{**}$.

Theorem 2.6 *Let v be any arbitrarily chosen (fixed) element of V. Then the map $\bar{v} : V^* \to \mathbb{R}$ defined by $\bar{v} : f \mapsto f \cdot v$ is in V^{**}.*

Proof Problem 2.4. □

Since for each $v \in V$ we have a $\bar{v} \in V^{**}$ by Theorem 2.6, that theorem gives us a mapping $\mathcal{I} : V \to V^{**}$. With the notation of Theorem 2.6, $\mathcal{I}(v) = \bar{v}$ and the map defined in that theorem can be written

$$\mathcal{I}(v) \cdot f = f \cdot v \tag{2.1}$$

Theorem 2.7 *The mapping $\mathcal{I} : V \to \mathcal{I}(V) \subset V^{**}$ is an isomorphism.*

Proof

$$\text{(i)} \qquad \begin{aligned} \mathcal{I}(av_1 + bv_2) \cdot f &= f \cdot (av_1 + bv_2) && \text{by (2.1)} \\ &= af \cdot v_1 + bf \cdot v_2 && \text{by linearity of } f \\ &= a\mathcal{I}(v_1) \cdot f + b\mathcal{I}(v_2) \cdot f && \text{by (2.1)} \\ &= (a\mathcal{I}(v_1) + b\mathcal{I}(v_2)) \cdot f \end{aligned}$$

by definition of operations in V^{**}. Hence $\mathcal{I}(av_1 + bv_2) = a\mathcal{I}(v_1) + b\mathcal{I}(v_2)$, so $\mathcal{I}$ is linear.

(ii) $\mathcal{I}(v_1) = \mathcal{I}(v_2)$ implies $f \cdot v_1 = f \cdot v_2$ for all $f \in V^*$, so $v_1 = v_2$ and $\mathcal{I}$ is $1 - 1$. (See Problem 2.7.) □

Corollary *If V is finite-dimensional, then V^{**} is isomorphic with V. Thus, if we identify V and V^{**} we can think of V as the space of linear functions on V^*. (See Section 2.2.)*

Definition $\mathcal{I}$ is called *the natural injection of V into V^{**}*.

PROBLEM 2.1. (i) Prove Theorem 2.1 with $\mathcal{L}(V, W)$ replaced by W^V, the set of all maps from V to W. (ii) Prove Theorem 2.1 with $\mathcal{L}(V, W)$ replaced by W^S where S is an arbitrary set. (Compare with $\mathbb{R}^S$ in Section 1.1.)

PROBLEM 2.2. Prove Theorem 2.2.

PROBLEM 2.3. Suppose $\{e_i\}$ and $\{\bar{e}_i\}$ are bases of $V, i = 1, \ldots, n$, and suppose $\{\varepsilon^i\}$ and $\{\bar{\varepsilon}^i\}$ are corresponding dual bases of V^*. Show that if

$$(\bar{e}_1 \cdots \bar{e}_n) = (e_1 \cdots e_n) \begin{pmatrix} a_1^1 & \cdots & a_n^1 \\ \vdots & & \vdots \\ a_1^n & \cdots & a_n^n \end{pmatrix}$$

$$\text{then} \quad \begin{pmatrix} \bar{\varepsilon}^1 \\ \vdots \\ \bar{\varepsilon}^n \end{pmatrix} = \begin{pmatrix} b_1^1 & \cdots & b_n^1 \\ \vdots & & \vdots \\ b_1^n & \cdots & b_b^n \end{pmatrix} \begin{pmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^n \end{pmatrix}$$

where $a_j^i b_k^j = \delta_k^i$. That is, the change-of-basis matrix (b_j^i) is the inverse of the change-of-basis matrix (a_j^i). Write the relation between the components, f_i and $\bar{f}_i$, of an element of V^* in the two bases in terms of (a_j^i).

PROBLEM 2.4. Write out the matrix of the linear transformation e_j^i in Theorem 2.3.

PROBLEM 2.5. In Theorem 2.5 we got a basis for a space of linear forms on a vector space which was permitted to be either finite- or infinite-dimensional. In Theorem 2.3 we got bases for spaces of mappings, and the proof given for Theorem 2.3 resembles that of Theorem 2.5. If in Theorem 2.3 V and/or W are infinite-dimensional we can define e_j^i as in the finite case. Precisely where will the proof fail if we try to prove Theorem 2.3 if V and/or W are infinite-dimensional?

PROBLEM 2.6. Prove Theorem 2.6.

PROBLEM 2.7. In the proof of part (ii) of Theorem 2.7 we stated that $f \cdot v_1 = f \cdot v_2$ for all $f \in V^*$ implies that $v_1 = v_2$. This is equivalent to the statement that if $v_1 \neq v_2$, then there is an f in V^* such that $f \cdot v_1 \neq f \cdot v_2$, or if $v \neq 0$, then there is an f in V^* such that $f \cdot v \neq 0$. Prove this.

PROBLEM 2.8. Prove that two matrices (ϕ_j^i) and (ψ_j^i) represent (with respect to two bases in V and their duals in V^*) the same linear mapping of V to V^* if and only if there is a matrix M such that $(\psi_j^i) = M^{tr}(\phi_j^i)M$, and M is the change of basis matrix, (a_j^i). Two such matrices are called *congruent* (cf., definition of similar matrices in Section 1.4).

2.2 Vector spaces of multilinear mappings

Definition Given vector spaces $V_1, V_2, \ldots, V_p, W$. A mapping $\phi : V_1 \times \cdots \times V_p \to W$ from the cartesian product $V_1 \times \cdots \times V_p$ to W is *multilinear* if it is linear in each argument.

Examples (i) We get an important example of a bilinear function ($p = 2$, $W = \mathbb{R}$) if we choose $V_1 = \mathbb{R}^n, V_2 = \mathbb{R}^m$. Then, given any a_{ij} $i = 1, \ldots, n$, $j = 1, \ldots, m$, the map $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ defined by

$$((v^1, \ldots, v^n), (w^1, \ldots, w^m)) \mapsto a_{ij} v^i w^j \tag{2.2}$$

is bilinear. Moreover, every bilinear function $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ can be written in this form. For, given ϕ, then

$$\phi : ((0, \ldots, \underset{i\text{th place}}{1}, \ldots, 0)(0, \ldots, \underset{j\text{th place}}{1}, \ldots, 0)) \mapsto a_{ij}$$

for some a_{ij} and then by bilinearity $\phi((v^1, \dots, v^n), (w^1, \dots, w^m)) = a_{ij}v^i w^j$. Clearly, these bilinear functions are generalizations of the dot product of vector analysis.

(ii) Another example of a bilinear function (in a certain sense both more and less general than the previous one) which we will see again in Chapter 3 is obtained if V_1 and V_2 are two vector spaces and we are given $f \in \mathcal{L}(V_1, \mathbb{R})$ and $g \in \mathcal{L}(V_2, \mathbb{R})$. Then $\phi : V_1 \times V_2 \to \mathbb{R}$ given by $(v_1, v_2) \mapsto (f \cdot v_1)(g \cdot v_2)$ is bilinear.

(iii) Finally, the "oriented volume" of a parallelepiped in "ordinary" space (a normed determinant function) is an example of a trilinear function.

The bilinear functions of example (i) are analogous to the "concrete" examples given by eq. (1.9) in the sense that every linear map is represented by the latter when bases are chosen in V and W, and every bilinear function is represented by the former when bases are chosen in V, W, and $\mathbb{R}$.

Theorem 2.8 *Given vector spaces $V_1, \dots, V_p, W$ and bases S_i of V_i, $i = 1, \dots, p$, then there exists a unique multi linear mapping $\phi : V_1 \times \cdots \times V_p \to W$ such that $\phi(v_{k_1}, v_{k_2}, \dots, v_{k_p}) = w_{k_1 \cdots k_p}$ where $v_{k_i} \in S_i$, and $w_{k_1 \cdots k_p}$ are arbitrary elements of W.*

Proof See Theorem 1.3. □

Theorem 2.9 *The set, $\mathfrak{L}(V_1, V_2, \dots, V_p; W)$ of all p-linear maps from $V_1 \times \cdots \times V_p$ to W (with the obvious definitions for the operations) is a vector space.*

Proof Problem 2.9. □

If the vector spaces $V_1, \dots, V_p, W$ are finite-dimensional, then $\mathfrak{L}(V_1, \dots, V_p; W)$ has a basis in terms of those of $V_1, \dots, V_p, W$ and the dimension of $\mathfrak{L}(V_1, \dots, V_p; W)$ is the product of the dimensions of $V_1, \dots, V_p, W$. More explicitly, we have the following special case.

Theorem 2.10 *Suppose $\{e_i\}$ and $\{E_i\}$ are bases of V and W respectively. Then the functions $f^{ij} : V \times W \to \mathbb{R}$ defined by $f^{ij} : (v, w) \mapsto v^i w^j$ where v^i and w^i are the components of v and w in the chosen bases, form a basis for $\mathfrak{L}(V, W; \mathbb{R})$. Moreover, for $b \in \mathfrak{L}(V, W; \mathbb{R})$*

$$b = b_{ij} f^{ij}$$

where $b_{ij} = b(e_i, E_j)$ (cf. Theorem 2.3).

Proof First of all, note that the definition of f^{ij} is equivalent to $f^{ij} : (e_k, E_l) \mapsto \delta^i_k \delta^j_l$ with f^{ij} extended to $V \times W$ by bilinearity.

(i) $b_{ij} f^{ij}(e_k, E_l) = b_{ij}\delta^i_k \delta^j_l = b_{kl}$ so $b_{ij} f^{ij} = 0$ implies that $b_{kl} = 0$ for all k, l. Hence the f^{ij} form a linearly independent set.

(ii) If $b \in \mathcal{L}(V, W; \mathbb{R})$ then $b(v, w) = v^i w^j b(e_i, E_j) = b(e_i, E_j) f^{ij}(v, w)$ for all $v \in V$ and $w \in W$, so $b = b(e_k, E_j) f^{ij}$. □

In the sequel we will be dealing with a variety of special cases of these spaces of multilinear mappings. It will be important to observe that these special cases are not all really different from one another – that certain spaces of mapping can be identified with certain others. What we mean by saying that two spaces can be "identified" with one another is that an isomorphism can be constructed between the two which does not require any choice of bases. Note that the isomorphisms established in Theorem 2.2 and Theorem 2.4 *do* require a choice of a basis. In the finite-dimensional case, isomorphisms between V and V^* require choosing bases, but the isomorphism $\mathcal{I}$ between V and V^{**} in the corollary of Theorem 2.7 does not. Hence we do not identify V and V^*, but we do identify V and V^{**}.

Definition A linear mapping of vector spaces which is independent of the choice of bases is called *natural* (or *canonical*).

Theorem 2.11 *There is a natural isomorphism between* $\mathfrak{L}(V_1, V_2; W)$ *and* $\mathfrak{L}(V_1, \mathfrak{L}(V_2, W))$.

Proof Let b be an element of $\mathfrak{L}(V_1, V_2; W)$. We define a map $\phi : \mathfrak{L}(V_1, V_2; W) \to \mathfrak{L}(V_1, \mathfrak{L}(V_2, W))$ by $b \mapsto \phi(b) \in \mathfrak{L}(V_1, \mathfrak{L}(V_2, W))$ where $\phi(b)$ is the linear map whose values, $\phi(b) \cdot v_1 \in \mathfrak{L}(V_2, W)$ are linear mappings given by

$$\phi(b) \cdot v_1 : v_2 \mapsto b(v_1, v_2) \in W. \tag{2.3}$$

That is, $\phi(b)$ is defined by its values, $\phi(b) \cdot v_1$, on V_1, each of which (i.e., for each fixed v_1) is a linear function from V_2 to W defined by (2.3).

We also write, for given b, and v_1, the partial map, $b(v_1, -) : v_2 \mapsto b(v_1, v_2)$, so with this notation $\phi(b) \cdot v_1$ is defined to be $b(v_1, -)$, that is, $\phi(b) \cdot v_1 = b(v_1, -)$.

We will prove that ϕ is $1-1$ and onto. (The linearity of ϕ comes by following through the definitions.) Let $\mathcal{A} \in \mathfrak{L}(V_1, \mathfrak{L}(V_2, W))$. We define a map $\psi : \mathfrak{L}(V_1, \mathfrak{L}(V_2, W)) \to \mathfrak{L}(V_1, V_2; W)$ by $\psi(\mathcal{A}) : (v_1, v_2) \mapsto (\mathcal{A} \cdot v_1) \cdot v_2$ (Figure 2.1). Now we form the compositions $\psi \circ \phi$ and $\phi \circ \psi$, and evaluate them on their respective domains.

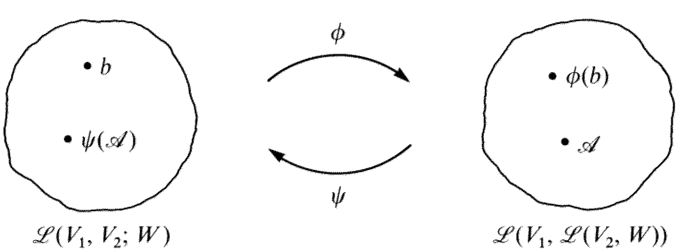


Figure 2.1

(i) For all b, v_1, v_2

$$\begin{aligned}\psi(\phi(b))(v_1, v_2) &= (\phi(b) \cdot v_1) \cdot v_2 \quad &&\text{(by the definition of } \psi)\\ &= b(v_1, -) \cdot v_2 &&\text{(by the definition of } \phi(b))\\ &= b(v_1, v_2)\end{aligned}$$

Hence $\psi(\phi(b))(v_1, v_2) = b(v_1, v_2)$ for all v_1, v_2, and b, so $\psi(\phi(b)) = b$ for all b and $\psi \circ \phi =$ identity on $\mathfrak{L}(V_1, V_2; W)$ which implies that ϕ is $1 - 1$.

(ii) For all $\mathcal{A}, v_1, v_2$

$$\begin{aligned}\phi(\psi(\mathcal{A})) \cdot v_1) \cdot v_2 &= \psi(\mathcal{A})(v_1, -) \cdot v_2 \quad &&\text{(by the definition of } \phi(b))\\ &= \psi(\mathcal{A})(v_1, v_2)\\ &= (\mathcal{A} \cdot v_1) \cdot v_2 &&\text{(by the definition of } \psi)\end{aligned}$$

Again, since this is valid for all v_1, v_2 and $\mathcal{A}$, we have $\phi \circ \psi =$ identity on $\mathfrak{L}(V_1, \mathfrak{L}(V_2, W))$ which implies that ϕ is onto. □

By Theorem 2.11 to each bilinear function $b \in \mathfrak{L}(V, W; \mathbb{R})$ there corresponds a unique linear map $\mathcal{A}_1 = \phi_1 \cdot b \in \mathfrak{L}(V, \mathfrak{L}(W, \mathbb{R}) \cong \mathfrak{L}(V, W^*)$ where $\phi_1 : \mathfrak{L}(V, W; \mathbb{R}) \to \mathfrak{L}(V, \mathfrak{L}(W, \mathbb{R}))$. Since, $\mathfrak{L}(V, \mathfrak{L}(W, \mathbb{R})) \cong \mathfrak{L}(W, \mathfrak{L}(V, \mathbb{R}))$, (see, Problem 2.14) corresponding to b there is also a unique linear map $\mathcal{A}_2 = \phi_2 \cdot b \in \mathfrak{L}(W, \mathfrak{L}(V, \mathbb{R})) \cong \mathfrak{L}(W, V^*)$ where $\phi_2 : \mathfrak{L}(V, W; \mathbb{R}) \to \mathfrak{L}(W, \mathfrak{L}(V, \mathbb{R}))$. $\phi_1 \cdot b : V \to W^*$ is given by $v \mapsto b(v, -)$ and $\phi_2 \cdot b : W \to V^*$ is given by $w \mapsto b(-, w)$.

If V and W are finite-dimensional and $\{e_i\}$ is a basis of V and $\{E_i\}$ is a basis of W, then $\phi_1 \cdot b : e_i \mapsto a_{ij}\Xi^j$ where $\{\Xi^j\}$ is the basis dual to $\{E_i\}$. But $\phi_1 \cdot b : e_i \mapsto b(e_i, -)$, so $a_{ij}\Xi^j \cdot E_k = b(e_i, E_k)$ and $a_{ik} = b(e_i, E_k) = b_{ik}$. That is the ikth component of b in the $\{f^{ik}\}$ basis is the ikth element of the matrix of the linear map $\mathcal{A}_1$. Similarly, if $\phi_2 \cdot b : E_i \mapsto a_{ij}\varepsilon^j$, then $a_{ik} = b_{ki}$, which says that the ikth component of b in the $\{f^{ik}\}$ basis is the ikth element of the transpose of the matrix of the linear map $\mathcal{A}_2$.

More generally, we have the following.

Theorem 2.12 *There is a natural isomorphism between* $\mathfrak{L}(V_1, V_2, \ldots, V_p; W)$ *and* $\mathfrak{L}(V_i, \mathfrak{L}(V_1, \ldots, \hat{V}_i, \ldots, V_p; W))$. *($\hat{V}_i$ means that argument is to be omitted.)*

Proof Problem 2.13. □

On the basis of these and other similar types of isomorphisms (see, e.g., Problem 2.14), we will be able to describe tensor spaces in a variety of equivalent ways - which can be convenient and also confusing. Moreover, such isomorphisms are used for the construction of "vector-valued forms" which we will encounter in geometry (cf., Section 4.1).

PROBLEM 2.9. Prove Theorem 2.9.

PROBLEM 2.10. Prove the general result stated below Theorem 2.9.

PROBLEM 2.11. We mentioned above Theorem 2.8 that every bilinear function is represented by the mapping (2.2) when bases are chosen. What are the coefficients of (2.2) for the f^{ij} of Theorem 2.10?

PROBLEM 2.12. Verify the statements in Theorem 2.11 that $\psi \circ \phi =$ identity on $\mathfrak{L}(V_1, V_2; W)$ implies that ϕ is $1-1$, and $\phi \circ \psi =$ identity on $\mathfrak{L}(V_1 : \mathfrak{L}(V_2, W))$ implies that ϕ is onto.

PROBLEM 2.13. Prove Theorem 2.12.

PROBLEM 2.14. Prove that there is a natural isomorphism between $\mathfrak{L}(V_1, \ldots, V_p; W)$ and $\mathfrak{L}(V_{i_1}, \ldots, V_{i_s}; \mathfrak{L}(V_{i_{s+1}}, \ldots, V_{i_p}; W))$ where $i_1, \ldots, i_p$ is a permutation of $1, \ldots, p$.

PROBLEM 2.15. The cartesian product $W_1 \times \cdots \times W_n$ of vector spaces with the operations

$$\begin{aligned}(w_1, \ldots, w_n) + (x_1, \ldots, x_n) &= (w_1 + x_1, \ldots, w_n + x_n)\\ a(w_1, \ldots, w_n) &= (aw_1, \ldots, aw_n)\end{aligned}$$

is a vector space denoted by $W_1 \oplus W_2 \oplus \cdots \oplus W_n$ and called *the exterior direct sum* (see Problem 1.11). Notice that the sets $\tilde{W}_i = \{(0, \ldots, w_i, \ldots, 0)\}$ are subspaces of $W_1 \oplus \cdots \oplus W_n$ and $W_1 \oplus \cdots \oplus W_n$ is the (interior) direct sum of $\tilde{W}_i$ as defined in Section 1.1. Show that if $\dim V = n$, then $\mathfrak{L}(V, W) \cong W \oplus \cdots \oplus W$, the direct sum of n copies of W (see Theorem 2.4).

2.3 Nondegenerate bilinear functions

Starting with a vector space V, we constructed a second vector space V^* each of whose elements is a mapping from V to $\mathbb{R}$. Instead of fixing an element of V^* and letting v vary in V to get a mapping with values in $\mathbb{R}$, we can, as we did in Theorem 2.6, fix an element $v \in V$ and let f vary in V^* and get a mapping $f \mapsto f \cdot v$ with values in $\mathbb{R}$. These two mappings appear in more symmetrical roles as partial mappings of a mapping $V^* \times V \to \mathbb{R}$ of the cartesian product of V^* and V to $\mathbb{R}$.

Definition The function $\delta : V^* \times V \to \mathbb{R}$ defined by $(f, v) \mapsto f \cdot v$ is called *the Kronecker delta*, or *the natural pairing of* V^* *and* V *into* $\mathbb{R}$. We write $\delta(f, v) = \langle f, v\rangle$, and $\delta = \langle -, -\rangle$; i.e., $\langle f, v\rangle = f \cdot v$.

The partial function $\langle f, -\rangle : V \to \mathbb{R}$ given by $v \mapsto f \cdot v$ is the same as f itself. The partial function $\langle -, v\rangle : V^* \to \mathbb{R}$ given by $f \mapsto f \cdot v$ is the function $\bar{v}$ in Theorem 2.6. It is easy to see that δ is a bilinear function, i.e., $\delta \in \mathfrak{L}(V^*, V; \mathbb{R})$.

By Theorem 2.10, if V is finite-dimensional and $\{e_i\}$ and $\{\varepsilon^i\}$ are dual bases, then $f_i^j : (\varepsilon^k, e_l) \mapsto \delta_i^k \delta_l^j$ define a basis for $\mathfrak{L}(V^*, V; R)$ and

$$\delta = \delta_j^i f_i^j = f_1^1 + f_2^2 + \cdots + f_n^n$$

δ has a property which defines an important class of bilinear functions.

Definitions A bilinear function $b : V \times W \to \mathbb{R}$ is (weakly) *nondegenerate* if $b(v, w) = 0$ for all $v \in V$ implies $w = 0$, and $b(v, w) = 0$ for all $w \in W$ implies $v = 0$. If b is a nondegenerate bilinear function on $V \times W$, then V *and* W *are dual spaces with respect to* b.

The Kronecker delta is nondegenerate, for the fact that $\langle f, v\rangle = 0$ for all v implies $f = 0$ is just the definition of $f = 0$. The fact that $\langle f, v\rangle = 0$ for all f implies $v = 0$ was already used in Theorem 2.6. See Problem 2.7. Thus, the natural pairing $\langle -, -\rangle$ of V^* and V is just one particular nondegenerate bilinear function, and V^* and V are dual with respect to $\langle -, -\rangle$.

Recall that in Section 2.2. we gave two examples of bilinear functions. In the first example, the bilinear function is nondegenerate and $\mathbb{R}^n$ and $\mathbb{R}^m$ are dual with respect to this function if and only if $m = n$, and the matrix (a_{ij}) is nonsingular. In this case b is a nondegenerate bilinear function of the form $b : V \times V \to \mathbb{R}$ and V is *self-dual* (with respect to that bilinear function). We will study important examples in Section 5.4.

PROBLEM 2.16. Prove the necessary and sufficient conditions stated for the nondegeneracy of the bilinear function $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ of the example given above.

PROBLEM 2.17. If V and W are dual with respect to b, then the maps $\mathcal{A}_1$ and $\mathcal{A}_2$ in Section 2.2 are injective.

PROBLEM 2.18. Choose a basis in V. (i) Express $\langle f, v\rangle$ in terms of this basis and its dual in V^*. (ii) What are the matrices, relative to these bases, of the linear maps $\phi_1 \cdot b$ and $\phi_2 \cdot b$ in Section 2.2 where b is the natural pairing of V and V^*?

PROBLEM 2.19. Give an example of a vector space with two nonisomorphic dual spaces.

2.4 Orthogonal complements and the transpose of a linear mapping

Definitions Suppose b is a bilinear function on $V \times W$, and $S \subset V$ and $T \subset W$. $S^\perp = \{w \in W : b(v, w) = 0 \text{ for all } v \in S\}$ is called *the orthogonal complement*, or *annihilator of* S *with respect to* b. ${}^\perp T = \{v \in V : b(v, w) = 0 \text{ for all } w \in T\}$ is *the orthogonal complement*, or *annihilator of* T *with respect to* b.

Theorem 2.13 *For any set* $S \subset V, S^\perp$ *is a subspace of* W, *and for any set* $T \subset W, {}^\perp T$ *is a subspace of* V.

Proof Problem 2.20. □

Theorem 2.14 *$N_1 = {}^{\perp}W$ and $N_2 = V^{\perp}$ where $N_1 \subset V$ is the null space of $\phi_1 \cdot b$ and $N_2 \subset W$ is the null space of $\phi_2 \cdot b$.*

Proof Problem 2.21. □

Corollary *b is nondegenerate if and only if $N_1 = 0$ and $N_2 = 0$.*

Theorem 2.15 *If finite-dimensional vector spaces V and W are dual with respect to b, then $W \cong V^*$ and $V \cong W^*$.*

Proof Since $N_1 = N_2 = 0$ the linear maps $\phi_1 \cdot b$ and $\phi_2 \cdot b$ are $1-1$. Hence $\dim V \leq \dim W^* = \dim W \leq \dim V^* = \dim V$. So $\dim V = \dim W^*$ and $\dim W = \dim V^*$ and the maps are onto. □

Corollary *If V and W are dual with respect to b, then* $\dim V = \dim W$.

On the basis of Theorem 2.15 we frequently say that V^* is *the* dual space of V, and $V^{**} = V$ is *the* dual space of V^*.

(Note that in the infinite-dimensional case the conclusion of Theorem 2.15 is not necessarily valid. If it is, we say that b is strongly nondegenerate, or *nonsingular*.)

Theorem 2.16 *If $S \subset V$ and $T \subset W$, $\langle S \rangle \subset {}^{\perp}(S^{\perp})$ and $\langle T \rangle \subset ({}^{\perp}T)^{\perp}$.*

Proof If $v \in \langle S \rangle$, then for all $w \in S^{\perp}, b(v, w) = a_i b(v^i, w) = 0$ so $v \in {}^{\perp}(S^{\perp})$. Similarly for the second part. □

We can strengthen this result, and get other interesting relations in the case $W = V^*$ and, $b = \delta$. Hence, *the orthogonal complement will be with respect to δ from now on.*

Theorem 2.17 *If $S \subset V$, then $\langle S \rangle = {}^{\perp}(S^{\perp})$.*

Proof From Theorem 2.16 we know $\langle S \rangle \subset {}^{\perp}(S^{\perp})$. Now suppose $\tilde{v} \in {}^{\perp}(S^{\perp})$ and $\tilde{v} \notin \langle S \rangle$. We can choose the $\tilde{v}$ to be one of the basis elements of V by Theorem 1.1. Let $\tilde{f}$ be the element of V^* which maps $\tilde{v}$ to 1, and the other basis elements to zero. $\tilde{f}$ exists and is unique by Theorem 1.3. In particular, for all $v \in \langle S \rangle, \langle \tilde{f}, v \rangle = 0$. That is, $\tilde{f} \in \langle S \rangle^{\perp}$. But $S \subset \langle S \rangle$, so $\tilde{f} \in S^{\perp}$. Hence, for all $v \in {}^{\perp}(S^{\perp}), \langle \tilde{f}, v \rangle = 0$. In particular, since by assumption $\tilde{v} \in {}^{\perp}(S^{\perp})$, we must have $\langle \tilde{f}, \tilde{v} \rangle = 0$, but this contradicts $\tilde{f} \cdot \tilde{v} = 1$, so ${}^{\perp}(S)^{\perp} \subset \langle S \rangle$. □

Theorem 2.18 *If V is finite-dimensional and S is a subspace of V then* $\dim S^{\perp} = \dim V - \dim S$ *($\perp$ is with respect to δ).*

Proof Problem 2.23. □

Now suppose we have two bilinear functions $b_1 : V_1 \times W_1 \to \mathbb{R}$ and $b_2 : V_2 \times W_2 \to \mathbb{R}$ and a linear mapping $\Phi : V_1 \to V_2$. Φ determines a *dual* linear mapping, Φ^*, from W_2 to W_1 by

$$b_1(v_1, \Phi^* \cdot w_2) = b_2(\Phi \cdot v_1, w_2) \tag{2.4}$$

We will confine ourselves to the special case $b_1 = b_2 = \delta$.

Definition If $\Phi : V \to W$, then the *transpose*, or *dual of* Φ is the map $\Phi^* : W^* \to V^*$ defined by $\Phi^* : w^* \mapsto w^* \circ \Phi$ where $w^* \in W^*$.

We can think of Φ^* as taking functions defined on W and "pulling them back" to functions defined on V.

$$\begin{tikzcd}
& \mathbb{R} & \\
V \arrow[ur, "\Phi^*(w^*) = w^* \circ \Phi"] \arrow[rr, "\Phi"'] & & W \arrow[ul, "w^*"']
\end{tikzcd}$$

Theorem 2.19 *Φ^* is linear.*

Proof We evaluate $\Phi^*(aw_1^* + bw_2^*)$ at an arbitrary $v \in V$. Thus,

$$\begin{aligned}
\Phi^*(aw_1^* + bw_2^*) \cdot v &= ((aw_1^* + bw_2^*) \circ \Phi) \cdot v \qquad \text{by definition of } \Phi^* \\
&= (aw_1^* + bw_2^*) \cdot (\Phi \cdot v) = a(w_1^* \cdot (\Phi \cdot v)) + b(w_2^* \cdot (\Phi \cdot v)) \\
&= a(\Phi^*(w_1^*) \cdot v) + b(\Phi^*(w_2^*) \cdot v) \quad \text{by definition of } \Phi^* \\
&= (a\Phi^*(w_1^*) + b\Phi^*(w_2^*)) \cdot v \\
&\qquad \text{by definition of operations in } V^*
\end{aligned}$$

So $\Phi^*(aw_1^* + bw_2^*) = a\Phi^*(w_1^*) + b\Phi^*(w_2^*)$. □

We discussed an example of a mapping and its transpose near the end of Section 2.2. More precisely, if $\mathcal{A}_1 = \phi_1 \cdot b : V \to W^*$ and $\mathcal{A}_2 = \phi_2 \cdot b : W \to V^*$, then $\mathcal{A}_2 = \mathcal{A}_1^*\big|_W$. That is, $(\mathcal{A}_2 \cdot w) \cdot v = (\mathcal{A}_1^* \cdot w^{**}) \cdot v$ when $w^{**} = w$ (by the isomorphism $\mathcal{I}$). This is seen by rewriting the right side: $(\mathcal{A}_1^* \cdot w^{**}) \cdot v = [w^{**} \circ (\phi_1 \cdot b)](v) = w^{**} \cdot w^* = w^* \cdot w = (\phi_1 \cdot b) \cdot v \cdot w = (\phi_2 \cdot b) \cdot w \cdot v$. Moreover, that discussion (in Section 2.2) shows that the transpose of the matrix of a linear map is the matrix of the transpose of that map.

Theorem 2.20 *If $\Phi_1 : V \to W$ and if $\Phi_2 = \Phi_1^*$, then $\Phi_1 = \Phi_2^*\big|_V$.*

Proof Problem 2.24. □

Theorem 2.21 *Φ^* is the transpose of Φ if and only if*

$$\langle \Phi^* \cdot w^*, v \rangle = \langle w^*, \Phi \cdot v \rangle \tag{2.5}$$

for all $v \in V$ and $w^* \in W^*$.

Proof If we evaluate $\Phi^* \cdot w^* = w^* \circ \Phi$ at v we get (2.5) immediately. □

Theorem 2.22 *If $\Phi : V \to W$ and Φ^* is its transpose, then*

(i) $\ker \Phi^* = (\Phi(V))^\perp$

(ii) $\Phi(V) = {}^\perp(\ker \Phi^*)$

(iii) $\ker \Phi = {}^{\perp}(\Phi^*(W^*))$
(iv) $\Phi^*(W^*) = \ker(\Phi)^{\perp}$
See Figure 2.2.

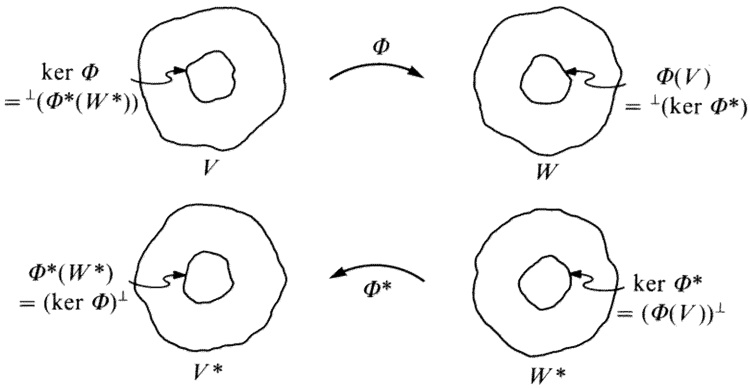


Figure 2.2

Proof (i) If $w^* \in \ker \Phi^*$, then $\Phi^* \cdot w^* = 0$ and by eq. (2.5) $\langle w^*, \Phi \cdot v\rangle = 0$ for all $v \in V$, so $w^* \in \Phi(V)^{\perp}$. On the other hand, if $w^* \in \Phi(V)^{\perp}$, then $\langle w^*, \Phi{\cdot}v\rangle = 0$ for all $v \in V$. By eq. (2.5) $\langle \Phi^*{\cdot}w^*, v\rangle = 0$ for all $v \in V$. Since $\langle -, -\rangle$ is nondegenerate $\Phi^* \cdot w^* = 0$ so $W^* \in \ker \Phi^*$.

(ii) By Theorem 2.17 $\Phi(V) = {}^{\perp}(\Phi(V)^{\perp})$ which is ${}^{\perp}(\ker \Phi^*)$ by part (i).

(iii) and (iv): Problem 2.26. □

Corollary *Φ is onto if and only if Φ^* is $1-1$. Φ is $1-1$ if and only if Φ^* is onto.*

Some of the results above are the algebraic abstractions of important results in other areas of mathematics. Thus, e.g., Theorem 2.22(ii) has the interpretation that a nonhomogeneous system of linear equations has a solution if and only if the nonhomogeneous part is orthogonal to every solution of the adjoint homogeneous system. This is known as the "Fredholm Alternative" in the theory of integral equations, or, more generally, in functional analysis (A. Friedman, p. 189 ff).

PROBLEM 2.20. Prove Theorem 2.13.

PROBLEM 2.21. Prove Theorem 2.14.

PROBLEM 2.22. Prove that $\mathcal{I}({}^\perp S^*) = \mathcal{I}(V) \cap {}^\perp(S^*)$ for $S^* \subset V^*$.

PROBLEM 2.23. Prove Theorem 2.18.

PROBLEM 2.24. Prove Theorem 2.20.

PROBLEM 2.25. In the finite-dimensional case with given bases, write (2.4) in matrix notation and find the relation between the matrices of Φ and Φ^*.

PROBLEM 2.26. Prove parts (iii) and (iv) of Theorem 2.22.

PROBLEM 2.27. Using parts (iii) and (iv) of Theorem 2.22 prove that if $T \subset W$, then $\langle T \rangle = ({}^\perp T)^\perp$.

PROBLEM 2.28. In the general case of the dual linear mapping, Φ^*, prove that Φ^* satisfies parts (i) and (iii) of Theorem 2.22.

PROBLEM 2.29. (Abstract homology, cf., Whitney, p. 346 ff or Greub, p. 178 ff). Suppose $W = V$ and Φ is a linear transformation on V with the property $\Phi^2 = 0$. (Φ is called *a differential operator*). Show (i) $\Phi(V) \subset \ker \Phi$, (ii) $\Phi^{*2} = 0$, (iii) $\Phi^*(V^*) \subset \ker \Phi^*$. Finally, the factor spaces $\ker \Phi/\Phi(V)$ and $\ker \Phi^*/\Phi^*(V^*)$, called the *homology* and *cohomology* spaces of V, respectively, are isomorphic. These ideas will arise in a more specific context in Chapter 12.

3

TENSOR PRODUCT SPACES

We will describe the space of bilinear functions on a pair of finite-dimensional vector spaces as the tensor product space of the duals of those two vector spaces. We then consider tensor product spaces of more than two vector spaces. Finally we define tensor product spaces in the general case so that they reduce to spaces of multilinear functions in finite dimensions. From now on we will denote vectors by $v, w, \ldots$, which may have subscripts, and linear functions by Greek letters $\sigma, \tau, \ldots$, which may have superscripts.

3.1 The tensor product of two finite-dimensional vector spaces

Consider the set of bilinear functions $\mathcal{L}(V, W; \mathbb{R})$, with V and W finite-dimensional. Notice that this set has certain elements of the following type. Let $\sigma \in V^*$ and $\tau \in W^*$. Then with each fixed pair $(\sigma, \tau) \in V^* \times W^*$ we have the function $V \times W \to \mathbb{R}$ given by

$$(v, w) \mapsto (\sigma \cdot v)(\tau \cdot w) \tag{3.1}$$

By the linearity of σ and τ this function is bilinear, so it is in $\mathcal{L}(V, W; \mathbb{R})$. (See example (ii) in Section 2.2.) Since the values of this function are products of the values of σ and τ, we use the product notation $\sigma \otimes \tau$ for this function. That is, $\sigma \otimes \tau : V \times W \to \mathbb{R}$ is the element of $\mathcal{L}(V, W; \mathbb{R})$ given by (3.1).

For an example, consider the case where $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$. Then $\sigma \in \mathbb{R}^{n*}$ will have the form $\sigma : (v^1, \ldots, v^n) \mapsto \sigma_i v^i$, and $\tau \in \mathbb{R}^{m*}$ will have the form $\tau : (w^1, \ldots, w^m) \mapsto \tau_i w^i$. (See below Theorem 2.4.) According to (3.1) $\sigma \otimes \tau$ will be given by $\sigma \otimes \tau : ((v^1, \ldots, v^n), (w^1, \ldots, w^m)) \mapsto \sigma_i \tau_j v^i w^j$. Note that this is a special case of example (i) in Section 2.2. Note also, that with bases in V and W every element in $\mathfrak{L}(V, W; \mathbb{R})$ of the form $\sigma \otimes \tau$ can be represented by such examples.

Now, for each pair $(\sigma, \tau) \in V^* \times W^*$ we have an element $\sigma \otimes \tau$ of $\mathfrak{L}(V, W; \mathbb{R})$. Thus, $\mathfrak{L}(V, W; \mathbb{R})$ contains a set of "products", the image set of a mapping

$$\otimes : V^* \times W^* \to \mathfrak{L}(V, W; \mathbb{R})$$

given by

$$(\sigma, \tau) \mapsto \sigma \otimes \tau \tag{3.2}$$

One easily verifies that $\otimes$ is bilinear. Thus, $\sigma \otimes \tau$ plays two roles: it is a bilinear function given by (3.1), and it is in the image of a bilinear map, $\otimes$, given by (3.2).

The image set of $\otimes$ is not a subspace of $\mathfrak{L}(V, W; \mathbb{R})$ and, in particular, not $\mathfrak{L}(V, W; \mathbb{R})$, since $\sigma^1 \otimes \tau^1 + \sigma^2 \otimes \tau^2$ is not always of the form $\sigma \otimes \tau$. However, when V and W are finite-dimensional spaces, if $\{\varepsilon^i\}$ is a basis of V^* and $\{\Xi^i\}$ is a basis of W^*, then $\{\varepsilon^i \otimes \Xi^j\} = \{f^{ij}\}$ is a basis of $\mathfrak{L}(V, W; \mathbb{R})$ (cf., Theorem 2.10), so the set of products does span $\mathfrak{L}(V, W; \mathbb{R})$. This accounts for the following terminology and notation.

Definition If V and W are finite-dimensional spaces, then the vector space $\mathfrak{L}(V, W; \mathbb{R})$ is called *the tensor* (or *Kronecker*) *product* of V^* and W^*. We write $\mathfrak{L}(V, W; \mathbb{R}) = V^* \otimes W^*$. (Note that we are using the notation $\otimes$ in three different ways.)

Similarly, since in $\mathfrak{L}(V^*, W^*; \mathbb{R})$ we have the "product" elements

$$v \otimes w : (\sigma, \tau) \mapsto (\sigma \cdot v)(\tau \cdot w),$$

(cf., corollary of Theorem 2.7), the vector space $\mathfrak{L}(V^*, W^*; \mathbb{R})$ is the *tensor* (or *Kronecker*) *product* of V and W, and we write $\mathfrak{L}(V^*, W^*; \mathbb{R}) = V \otimes W$.

$\mathfrak{L}(V, V^*; \mathbb{R}) = V^* \otimes V$ is an example of a tensor product space. So is $\mathfrak{L}(V^*, V; \mathbb{R}) = V \otimes V^*$. The natural pairing, $\langle -, - \rangle$ (Section 2.3) is an element of $V \otimes V^*$.

Theorem 3.1 *Given two finite-dimensional vector spaces V and W. Form $V \otimes W = \mathfrak{L}(V^*, W^*; \mathbb{R})$ and the map $\otimes : V \times W \to V \otimes W$ with $\otimes(v, w)$ defined as above. The pair $(V \otimes W, \otimes)$ has the following properties:*

(i) $\otimes$ is bilinear, and $\otimes(V \times W)$ spans $V \otimes W$

(ii) If Z is any vector space, and b is any bilinear map, $b : V \times W \to Z$, then there exists a unique linear map, $\phi : V \otimes W \to Z$, onto $\langle b(V \times W) \rangle$, the linear closure of the range of b, and such that $b = \phi \circ \otimes$; i.e., any bilinear map, b, has a unique factorization into the product of $\otimes$ and a linear map. (The fact that every bilinear map, b, can be factored into a linear map and a particular unique map is referred to as the universal property of bilinear maps.)

Proof (i) These two properties of $\otimes : V \times W \to V \otimes W$ are precisely analogous to those of the mapping given by eq. (3.2). (ii) Let $\{v_i \otimes w_j\}$ be a basis of $V \otimes W$. Define ϕ by mapping $v_i \otimes w_j \mapsto b(v_i, w_j)$ and then extending this correspondence to all of $V \otimes W$ by linearity. Every $z \in \langle b(V \times W) \rangle$ can be written $z = a^{ij} b(v_i, w_j)$. But this is the image of $a^{ij} v_i \otimes w_j$ under ϕ, so ϕ is onto $\langle b(V \times W) \rangle$. If $\phi_1 \circ \otimes = \phi_2 \circ \otimes$, then $(\phi_1 - \phi_2) \cdot \otimes(v, w) = 0$ for all v, w. But, since $\{\otimes(v, w) = v \otimes w : v \in V, w \in W\}$ spans $V \otimes W, (\phi_1 - \phi_2) \cdot A = 0$ for all $A \in V \otimes W$, which implies that $\phi_1 - \phi_2$. $\square$

$$\begin{tikzcd} V \times W \arrow[r, "\otimes"] \arrow[dr, "b"'] & V \otimes W \arrow[d, "\phi"] \\ & Z \end{tikzcd}$$

Corollary *For a bilinear map, b, in Theorem 3.1 (ii) with the additional property that the dimension of the linear closure $\langle b(V \times W)\rangle$ of its range is* $\dim V \times W$, *the corresponding linear map, ϕ, is an isomorphism of $V \times W$ with $\langle b(V \times W)\rangle$.*

Proof By the theorem, ϕ is onto $\langle b(V \times W)\rangle$. This plus the fact that $\dim\langle b(V \times W)\rangle = \dim V \times W$ makes ϕ an isomorphism. □

Corollary *Given spaces V, W, and Z, there is a natural isomorphism between the space $\mathfrak{L}(V, W; Z)$ of bilinear mappings and the space $\mathfrak{L}(V \otimes W, Z)$ of linear mappings.*

Proof Problem 3.4. □

At the beginning of this section we saw that if $\{\varepsilon^i\}$ is a basis of V^* and $\{\Xi^i\}$ is a basis of W^*, then $\{\varepsilon^i \otimes \Xi^j\}$ form a basis of $V^* \otimes W^*$, so for any $A \in V^* \otimes W^*$ we can write

$$A = A_{ij}\varepsilon^i \otimes \Xi^j \tag{3.3}$$

and since, by eq. (3.1), $\varepsilon^i \otimes \Xi^j(v, w) = v^i w^j$, A has values

$$A(v, w) = A_{ij} v^i w^j. \tag{3.4}$$

Similarly $\{e_i \otimes E_j\}$ form a basis of $V \otimes W$, so for any $A \in V \otimes W$,

$$A = A^{ij} e_i \otimes E_j \tag{3.5}$$

and, since $e_i \otimes E_j(\sigma, \tau) = \sigma_i \tau_j$, A has values

$$A(\sigma, \tau) = A^{ij}\sigma_i \tau_j. \tag{3.6}$$

PROBLEM 3.1. Write the explicit form of the product $v \otimes w \in V \otimes W$ as a bilinear function if $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$.

PROBLEM 3.2. Give a specific example in which $\sigma^1 \otimes \tau^1 + \sigma^2 \otimes \tau^2$ in $V^* \otimes W^*$ is not a product.

PROBLEM 3.3. Prove part (i) of Theorem 3.1.

PROBLEM 3.4. Prove the second corollary of Theorem 3.1.

3.2 Generalizations, isomorphisms, and a characterization

We can generalize the development of the last section and construct tensor products of any finite number of vector spaces. For example, we note that with every triple $(\sigma, \tau, \omega) \in V^* \times W^* \times X^*$ we have an element of $\mathfrak{L}(V, W, X; \mathbb{R})$

$$(v, w, x) \mapsto (\sigma \cdot v)(\tau \cdot w)(\omega \cdot x)$$

and we call this function $\sigma \otimes \tau \otimes \omega$. That is, we have a map $\otimes : V^* \times W^* \times X^* \to \mathfrak{L}(V, W, X; \mathbb{R})$ whose images are of the form $\sigma \otimes \tau \otimes \omega$. We have a theorem analogous to Theorem 3.1. In particular, $\otimes$ is trilinear and $\otimes(V^* \times W^* \times X^*)$ spans $\mathfrak{L}(V, W, X; \mathbb{R})$, which we now denote by $V^* \otimes W^* \otimes X^*$. Also, any trilinear map has a unique factorization into the product of $\otimes$ and a linear map.

We can generalize eqs. (3.3) - (3.6). In particular, if $\{\varepsilon_1^i\}$, $\{\varepsilon_2^i\}$, and $\{\varepsilon_3^i\}$ are bases respectively of V^*, W^*, and X^*, then $\{\varepsilon_1^i \otimes \varepsilon_2^j \otimes \varepsilon_3^k\}$ is a basis of $V^* \otimes W^* \otimes X^*$ and for any $A \in V^* \otimes W^* \otimes X^*$,

$$A = A_{ijk}\varepsilon_1^i \otimes \varepsilon_2^j \otimes \varepsilon_3^k \tag{3.7}$$

and A has values

$$A(v, w, x) = A_{ijk}v^i w^j x^k \tag{3.8}$$

Generally, given m vector spaces $V_1, \ldots, V_m$, we can construct $V_1 \otimes \cdots \otimes V_m$. If $\{e_{i_1}^1\}, \ldots, \{e_{i_m}^m\}$ are bases of $V_1, \ldots, V_m (i_k = 1, \ldots, n_k$ where $n_k = \dim V_k)$ then $\{e_{i_1}^1 \otimes \cdots \otimes e_{i_m}^m\}$ is a basis of $V_1 \otimes \cdots \otimes V_m$ and for any $A \in V_1 \otimes \cdots \otimes V_m$,

$$A = A^{i_1 \cdots i_m} e_{i_1}^1 \otimes \cdots \otimes e_{i_m}^m \tag{3.9}$$

and A has values

$$A(\sigma^1, \ldots, \sigma^m) = A^{i_1 \cdots i_m} \sigma_{i_1}^1 \cdots \sigma_{i_m}^m \tag{3.10}$$

Note that $\dim V_1 \otimes \cdots \otimes V_m = n_1 \cdots n_m$. This is a special case of the general result stated above Theorem 2.10. (See Problem 2.10.)

Since we can construct the tensor product of any two spaces, we can form the tensor product of tensor products. In general, the operations of taking tensor products and/or duals can be iterated, resulting in an apparently bewildering proliferation of vector spaces. The following theorems help to keep things under control.

Theorem 3.2 *The vector spaces $V \otimes W \otimes X$, $(V \otimes W) \otimes X$, and $V \otimes (W \otimes X)$ are naturally isomorphic.*

Proof A bilinear map $b : (V \otimes W) \times X \to V \otimes W \otimes X$ is determined by $(v \otimes w, x) \mapsto v \otimes w \otimes x$. (For each fixed $v \otimes w$ this prescription determines a linear map from X to $V \otimes W \otimes X$, and for each fixed x, by Theorem 3.1, since $(v, w) \mapsto v \otimes w \otimes x$ is bilinear, it determines a linear map from $V \otimes W$ to

$V \otimes W \otimes X$.) Then by Theorem 3.1 we have the linear map $\phi : v(\otimes w) \otimes x \mapsto v \otimes w \otimes x$.

$$\begin{tikzcd}
(V \otimes W) \times X \arrow[r, "\otimes"] \arrow[dr, "b"'] & (V \otimes W) \otimes X \arrow[d, "\phi"] \\
& V \otimes W \otimes X
\end{tikzcd}$$

From the trilinear map $t : V \times W \times X \to (V \otimes W) \otimes X$ given by $(v, w, x) \mapsto (v \otimes w) \otimes x$ we have the linear map $\psi : v \otimes w \otimes x \mapsto (v \otimes w) \otimes x$

$$\begin{tikzcd}
V \times W \times X \arrow[r, "\otimes"] \arrow[dr, "t"'] & V \otimes W \otimes X \arrow[d, "\psi"] \\
& (V \otimes W) \otimes X
\end{tikzcd}$$

But $\phi \circ \psi$ and $\psi \circ \phi$ are both identities so ϕ is an isomorphism. A similar argument shows that $V \otimes (W \otimes X)$ and $V \otimes X \otimes X$ are isomorphic. □

Theorem 3.3 *$V \otimes W$ and $W \otimes V$ are naturally isomorphic.*

Proof Problem 3.7. □

We saw that $\otimes$ is bilinear, and, in particular, distributive with respect to $+$. We might be tempted to infer from Theorems 3.2 and 3.3 that $\otimes$ is also associative and commutative. However, in general, these properties are not even defined for $\otimes$ and, when they are, they are not generally true (see Problem 3.8). When we define a closely related mapping in Section 4.3 this situation will be partially rectified (see Theorem 4.6).

Theorem 3.4 *$(V \otimes W)^*$ is naturally isomorphic to $V^* \otimes W^*$.*

Proof The result follows immediately from the second corollary of Theorem 3.1 for the case $Z = \mathbb{R}$. □

Corollary *$(V^* \otimes W^*)^*$ is naturally isomorphic to $V \otimes W$.*

Theorem 3.4 implies that $V \otimes W$ and $V^* \otimes W^*$ are dual spaces with respect to the natural pairing. It is an example of "duality" in tensor product spaces (see Section 4.1). It is important to note that Theorem 3.4 is not necessarily valid if V and W are not both finite-dimensional.

We have seen that when V and W are finite-dimensional, the pair $(V \otimes W, \otimes)$ has the properties listed in Theorem 3.1, which with its corollaries, in turn, lead to the isomorphisms described above. Now suppose we abstract this situation a bit; suppose V, W, and X are any vector spaces, and $\mathfrak{F}$ is any map $V \times W \to X$, and suppose the pair $(X, \mathfrak{F})$ satisfies the properties listed in Theorem 3.1. It turns out that $(X, \mathfrak{F})$ is really no more general than $(V \otimes W, \otimes)$; that is, the properties of Theorem 3.1 essentially uniquely determine the pair $(X, \mathfrak{F})$.

Theorem 3.5 *If $(X, \mathfrak{F})$ has the properties of Theorem 3.1, then $X \cong V \otimes W$ and $\mathfrak{F} = \phi \circ \otimes$ where ϕ is the isomorphism $V \otimes W \to X$.*

Proof Letting $\mathfrak{F}$ have the role of b in Theorem 3.1 we have $\mathfrak{F} = \phi \circ \otimes$ where $\phi : V \otimes W \to X$. Then letting $\otimes$ have the role of b in Theorem 3.1 we have $\otimes = \psi \circ \mathfrak{F}$ where $\psi : X \to V \otimes W$. Combining these two equalities we get $\mathfrak{F} = \phi \circ \psi \circ \mathfrak{F}$ and $\otimes = \psi \circ \phi \circ \otimes$. Since $\mathfrak{F}(V \times W)$ spans X, $\phi \circ \psi$ is the identity on X, so ϕ is onto. Since $\otimes(V \times W)$ spans $V \otimes W$, $\psi \circ \phi$ is the identity on $V \otimes W$ so ϕ is $1-1$. Hence ϕ is an isomorphism. □

On the basis of Theorem 3.5 we can now say that given two finite-dimensional vector spaces, there is one and only one tensor product space with the properties of Theorem 3.1, and that tensor product space is the space $V \otimes W$ defined in Section 3.1. The major significance, however, of the characterization of Theorem 3.1 is that it can be extended to any two vector spaces, not necessarily finite-dimensional. The uniqueness requirement is shown exactly as in the proof of Theorem 3.5, as in that proof the nature of $V \otimes W$ is irrelevant. The existence requirement will be satisfied by construction, in the next section. Clearly, the constructed space will have to reduce to $V \otimes W = \mathfrak{L}(V^*, W^*; \mathbb{R})$ in the finite-dimensional case.

PROBLEM 3.5. Prove that $\mathfrak{L}(V_1, \ldots, V_n, W^*; \mathbb{R}) \cong \mathfrak{L}(V_1 \otimes \cdots \otimes V_n, W)$.

PROBLEM 3.6. The proof of Theorem 3.2 requires only that the tensor products satisfy Theorem 3.1 and its generalizations. Hence it is valid for arbitrary vector spaces (not necessarily finite-dimensional). Give a simpler proof of Theorem 3.2 by using Theorem 3.4. (Note, however, that Theorem 3.4 is only valid in the finite-dimensional case.)

PROBLEM 3.7. Prove Theorem 3.3 for arbitrary vector spaces (not necessarily finite-dimensional).

PROBLEM 3.8. If V and W are the same, then $V \otimes W$ and $W \otimes V$ are the same space, but in general $v \otimes w \neq w \otimes v$.

PROBLEM 3.9. Generalize Theorem 3.4 to $(V_1 \otimes V_2 \otimes \cdots \otimes V_p)^* \cong V_1^* \otimes V_2^* \otimes \cdots \otimes V_p^*$.

3.3 Tensor products of infinite-dimensional vector spaces

In Section 3.1, we defined tensor products of finite-dimensional vector spaces. In the general case where V and/or W may not be finite-dimensional we still have the maps $\otimes : V^* \times W^* \to \mathfrak{L}(V, W; \mathbb{R})$ given by eq. (3.2) and $\otimes : V \times W \to \mathfrak{L}(V^*, W^*; \mathbb{R})$ in Theorem 3.1. However, now the sets $\otimes(V^* \times W^*)$ and $\otimes(V \times W)$ do not necessarily span the spaces $\mathfrak{L}(V, W; \mathbb{R})$ and $\mathfrak{L}(V^*, W^*; \mathbb{R})$. These latter spaces may be too large.

It turns out that there are a space and a mapping which satisfy Theorem 3.1. As we have already pointed out, this can be the only such pair. We will denote it by $(V \otimes W, \otimes)$. That is, the tensor product of V and W, $(V \otimes W, \otimes)$, will be defined by the properties of Theorem 3.1.

Given V and W to get a space $V \otimes W$, and a mapping, $\otimes$ which satisfy Theorem 3.1, we start out by considering the vector space, $\mathbb{R}_0^{V\times W}$, of all real-valued functions with domain $V \times W$ which vanish at all but a finite number of points. As we saw in Section 1.1, there is a $1-1$ correspondence between elements $(v, w) \in V \times W$ and functions $f_{(v,w)} \in \mathbb{R}_0^{V\times W}$ which are 1 on (v, w) and 0 otherwise, and $\{f_{(v,w)} : (v, w) \in V \times W\}$ is the basis of $\mathbb{R}_0^{V\times W}$. We can think of $V \times W \subset \mathbb{R}_0^{V\times W}$, and we will use the notation (v, w) for $f_{(v,w)}$.

Now consider the set of all elements of $\mathbb{R}_0^{V\times W}$ of the form

$$\begin{aligned}(a^1v_1 + a^2v_2, b^1w_1 + b^2w_2) &- a^1b^1(v_1, w_1)\\ &- a^1b^2(v_1, w_2) - a^2b^1(v_2, w_1) - a^2b^2(v_2, w_2)\end{aligned}$$

with $v_1, v_2 \in V$, $w_1, w_2 \in W$ and $a^1, a^2, b^1, b^2 \in \mathbb{R}$. They generate a subspace, N. Then we put

$$V \otimes W = \mathbb{R}_0^{V\times W}/N.$$

That is, $V \otimes W$ is a set of equivalence classes of functions from $V \times W$ to $\mathbb{R}$. The map

$$\otimes = \Pi\big|_{V\times W}$$

is the restriction to $V \times W$ of the projection $\Pi : \mathbb{R}_0^{V\times W} \to \mathbb{R}_0^{V\times W}/N$.

Theorem 3.6 *The pair $(V \otimes W, \otimes)$ defined above has the properties of Theorem 3.1.*

Proof (i) Since Π is linear and maps all of N to 0,

$$\begin{aligned}\Pi(a^1v_1 + a^2v_2, b^1w_1 + b^2w_2) = a^1b^1\Pi(v_1, w_1) &+ a^1b^2\Pi(v_1, w_2)\\ &+a^2b^1\Pi(v_2, w_1) + a^2b_2\Pi(v_2, w_2)\end{aligned}$$

so $\otimes(a^1v_1 + a^2v_2, b^1w_1 + b^2w_2) = a^1b^1 \otimes (v_1, w_2) + \cdots$ for all $v_1, v_2 \in V$, $w_1, w_2 \in W$, and $a^1, a^2, b^1, b^2 \in \mathbb{R}$. That is, $\otimes$ is bilinear. Also, since Π is linear and $V \times W$ spans $\mathbb{R}_0^{V\times W}$, $\otimes(V \times W) = \Pi(V \times W)$ spans $V \otimes W$.

(ii) Now suppose we have $b : V \times W \to Z$. We define a linear map $\Phi : \mathbb{R}_0^{V\times W} \to Z$ by $a^{ij}(v_i, w_j) \mapsto a^{ij}b(v_i, w_j)$. Then we define a map $\phi : V \otimes W \to Z$