William B. Heard

Rigid Body Mechanics

Mathematics, Physics and Applications

WILEY-VCH

WILEY-VCH Verlag GmbH & Co. KGaA

This Page Intentionally Left Blank

William B. Heard
Rigid Body Mechanics

This Page Intentionally Left Blank

William B. Heard

Rigid Body Mechanics

Mathematics, Physics and Applications

WILEY-VCH
WILEY-VCH Verlag GmbH & Co. KGaA

The Author

William B. Heard
Alexandria, VA
USA

For a Solutions Manual, lecturers should contact the editorial department at physics@wiley-vch.de, stating their affiliation and the course in which they wish to use the book

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

Printed in the Federal Republic of Germany
Printed on acid-free paper

Printing betz-druck GmbH, Darmstadt
Binding Litges & Dopf Buchbinderei GmbH, Heppenheim

ISBN-13: 978-3-527-40620-3
ISBN-10: 3-527-40620-4

This book is dedicated to
Peggy
and the memory of my
Mother and Father

This Page Intentionally Left Blank

Contents

Rigid Body Mechanics: Mathematics, Physics and Applications. William B. Heard

ISBN: 3-527-40620-4

Preface

This is a textbook on rigid body mechanics written for graduate and advanced undergraduate students of science and engineering. The primary reason for writing the book was to give an account of the subject which was firmly grounded in both the classical and geometrical foundations of the subject. The book is intended to be accessible to a student who is well prepared in linear algebra and advanced calculus, who has had an introductory course in mechanics and who has a certain degree of mathematical maturity. Any mathematics needed beyond this is included in the text.

Chapter 1 deals with the rotations, the basic operation in rigid body theory. Rotations are presented in several parameterizations including axis angle, Euler angle, quaternion, and Cayley–Klein parameters. The rotations form a Lie group which underlies all of rigid body mechanics.

Chapter 2 studies rigid body motions, angular velocity, and the physical concepts of angular momentum and kinetic energy. The fundamental idea of angular velocity is straight from the Lie algebra theory. These concepts are illustrated with several examples from physics and engineering.

Chapter 3 studies rigid body dynamics in vector, Lagrangian, and Hamiltonian formulations. This chapter introduces many geometric concepts as dynamics occurs on differential manifolds and for rigid body mechanics the manifold is often a Lie group. The idea of the adjoint action is seen to be basic to the rigid body equations of motion. This chapter contains many examples from physics and engineering.

Chapter 4 considers the dynamics of constrained systems. Here Lagrange multipliers are introduced and several ways of determining or eliminating them are considered. This topic is rich in geometrical interactions and there are several examples, some standard and some not.

Chapter 5 considers the integrable problems of free rotation, Lagrange's top, and the gyrostat. The Kowalevsky top and Lax equations are also considered. Geometrical topics include the Poinsot construction, the geometric phase, and

Rigid Body Mechanics: Mathematics, Physics and Applications. William B. Heard

ISBN: 3-527-40620-4

Liouville tori. Verification and validation are of the utmost importance in the world of scientific and engineering computing and analytical solutions are treasured. In addition to the validation service they also play an important role in developing our intuitive understanding of the subject, not to mention their intrinsic and historical worth.

The importance of numerical methods in today's applications of mechanics cannot be overstated. Chapter 6 discusses classical numerical methods and more recent methods tailored specifically for Lie groups. This chapter includes a case study of a complicated rigid body motion, the wobblestone, which is naturally studied with numerical methods.

The final chapter, Chapter 7, applies the previous material to phenomena ranging in scale from the astronomical to the molecular. The largest scale concerns precession and nutation of the Earth. The Earth is nonspherical – an oblate spheroid to first approximation – and the axis of the Earth's rotation is observed to move on the celestial sphere. Most of this motion is attributed to the torque exerted on the nonspherical Earth by the Moon and the Sun and rigid body dynamics explains the effect. Next we study satellite gravity gradient stabilization. If Earth's satellites are not stabilized by some mechanism, they will tumble – as do the asteroids – and will not be useful platforms. One stabilization mechanism uses the gradient in the Earth's gravitational field and the basics of this mechanism are explained by rigid body theory. On the same scale, but down to the Earth, we consider the motion of a multibody, a mechanical system consisting of interconnected rigid bodies. Rigid body dynamics is used to study the motion of a robot arm. At the smallest scale we examine the techniques of molecular dynamics. Physicists and chemists study properties of matter by simulating the motions of a large number of interacting molecules. At the most basic level this is a problem in quantum mechanics. However, there are properties which can be calculated by idealizing the material to be a collection of interacting rigid bodies.

Three appendices are provided on spherical trigonometry, elliptic functions and Lie groups and Lie algebras. Lie groups and Lie algebras unify the subject. Appendix C provides background on all the Lie group concepts used in the text.

Some choices made in the course of writing the book should be mentioned. The application of mechanics often boils down to making a calculation and getting the useful number. I have tried to keep calculations foremost in mind. There are no theorem–proof structures in the book. Rather observations are made and substantiated by methods which are decidedly computational. Rarely are equations put in dimensionless form because one can rely on dimensional checks to avoid algebra mistakes and misconceived physics. Of course, when it comes time to compute one is well advised to pay attention to the scalings. Various notations are used in the text ranging from "old tensor" to intrinsic operators independent of coordinates and the concise notations of [1,23] which facilitate the computations. There is a glossary in Appendix D of the notations used. Examples are set aside in italic type and end with the symbol ◊.

William B. Heard

Alexandria, Virginia, U.S.A., 23 August 2005

This Page Intentionally Left Blank

1
Rotations

This chapter is devoted to rotations in three-dimensional space. Rotations are fundamental to rigid body dynamics because there is a one-one correspondence between orientations of a rigid body and rotations in three-dimensional space. The theory of rotations is a classical subject with a rich history and a variety of modes of expression. We shall begin by expressing the elements of the theory in terms of vectors and linear operators. Next, quaternions will be introduced and additional elements of the theory developed with them. This will lead to some elegant connections between rotations and spherical geometry. This leads, via the stereographic projection and Möbius transformations, to a description of rotations in terms of complex variables.

1.1
Rotations as Linear Operators

One way to approach rotations is to study their effect on spatial objects. The language of vectors and matrices provides a natural calculus. This section reviews some basic algebra of vector spaces and establishes our notation. Then the angle of rotation and axis of rotation as well as the Euler angles are studied as ways to parameterize rotation matrices.

1.1.1
Vector Algebra

Let us first establish some notation. Let $\mathbf{V}$ be a finite-dimensional vector space over the real numbers, $\mathbb{R}$. The elements of $\mathbf{V}$, the vectors, will be denoted by bold, lower case Latin letters, $\boldsymbol{u}$. A basis of $\mathbf{V}$ is a set of vectors $\{\boldsymbol{e}_i\}$ having the property that every vector has a unique representation as a linear combination of basis elements. The basis vectors will be indexed with subscripts. Let $\mathbf{V}^*$ be the dual vector space of $\mathbf{V}$ – the space of all linear, real valued functions on $\mathbf{V}$. The elements of $\mathbf{V}^*$, the covectors, will be denoted by bold, lower case Greek

Rigid Body Mechanics: Mathematics, Physics and Applications. William B. Heard

ISBN: 3-527-40620-4

letters, $\boldsymbol{v}$. For each basis $\{\boldsymbol{e}_i\}$ of $\mathbf{V}$ there is a basis $\{\boldsymbol{\epsilon}^i\}$ of $\mathbf{V}^*$ defined by

$$\boldsymbol{\epsilon}^i(\boldsymbol{e}_j) = \delta_{ij}$$

where δ_{ij} is the Kronecker delta. If $\boldsymbol{u} = u^i \boldsymbol{e}_i$, [1] then $\mathbf{e}$ and u will denote

$$\mathbf{e} = \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{pmatrix} \qquad u = \begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix}$$

and the representation of a vector $\boldsymbol{u}$ in the basis $\mathbf{e}$ is denoted by

$$\boldsymbol{u} = \mathbf{e}u$$

The basis $\mathbf{e}$ will also be referred to as a *frame*. Similarly, if $\boldsymbol{v} = v_i \boldsymbol{\epsilon}^i$, then $\boldsymbol{\epsilon}$ and v will denote

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon^1 \\ \vdots \\ \epsilon^n \end{pmatrix} \qquad v = \begin{pmatrix} v_1 \cdots v_n \end{pmatrix}$$

and the representation of covector $\boldsymbol{v}$ in the basis $\boldsymbol{\epsilon}$ is denoted by

$$\boldsymbol{v} = v\boldsymbol{\epsilon}$$

This notation is also found in [1] which is a good source of material on geometrical aspects of mechanics.

The distinction between a vector $\boldsymbol{u} = \mathbf{e}u$ and its components u relative to a basis $\mathbf{e}$ is of prime importance. The former is invariant under a change of basis and the latter, of course, is not. We shall always reserve the bold-face font for the invariant form and the regular typeface components relative to a basis. A *linear operator* $\mathcal{A}$ on $\boldsymbol{V}$ takes vector $\boldsymbol{x} \in \boldsymbol{V}$ to vector $\mathcal{A}(\boldsymbol{x}) \in \boldsymbol{V}$ and preserves the operations of the vector space

$$\mathcal{A}(a\boldsymbol{x} + b\boldsymbol{y}) = a\mathcal{A}(\boldsymbol{x}) + b\mathcal{A}(\boldsymbol{y}), \text{ for all } a, b \in \mathbb{R} \text{ and } \boldsymbol{x}, \boldsymbol{y} \in \mathbf{V}$$

Given a linear operator $\mathcal{A}$ and a basis $\{\mathbf{e}_i\}$ define the matrix representation $A = [A^i_j]$ (row superscript i and column subscript j) by

$$\mathcal{A}(\mathbf{e}_j) = A^i_j \boldsymbol{e}_i$$

The action of $\mathcal{A}$ on any $\mathbf{u} = u^j \mathbf{e}_j$ is then represented as

$$\mathcal{A}\mathbf{u} = A^i_j u^j \boldsymbol{e}_i$$

1) The summation convention that repeated indices indicate a sum over their range is used here and throughout the text.

The matrix can be regarded as operating on a row of basis vectors from the right according to

$$\mathcal{A}(\boldsymbol{u}) = u^j \boldsymbol{f}_j \qquad \text{with} \quad \boldsymbol{f}_j = A^i_j \boldsymbol{e}_i$$

This can be expressed in a matrix form as

$$\begin{pmatrix} \boldsymbol{f}_1 & \boldsymbol{f}_2 & \boldsymbol{f}_3 \end{pmatrix} = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{pmatrix} \begin{pmatrix} A^1_1 & A^1_2 & A^1_2 \\ A^2_1 & A^2_2 & A^2_3 \\ A^3_1 & A^3_2 & A^3_3 \end{pmatrix}$$

or in the shorthand notation

$$\boldsymbol{f} = \mathbf{e}A$$

Alternatively it can be regarded as operating on a column of components from the left according to

$$\mathcal{A}(\boldsymbol{u}) = v^i \boldsymbol{e}_i \qquad \text{with} \quad v^i = A^i_j u_j$$

This can be expressed as

$$\begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \begin{pmatrix} A^1_1 & A^1_2 & A^1_2 \\ A^2_1 & A^2_2 & A^2_3 \\ A^3_1 & A^3_2 & A^3_3 \end{pmatrix} \begin{pmatrix} u^1 \\ u^2 \\ u^3 \end{pmatrix}$$

or

$$v = Au$$

Given a pair of vectors $\boldsymbol{u}, \boldsymbol{v}$, a scalar or inner product on $\boldsymbol{V}$ assigns a non-negative, real number $\langle \mathbf{u}, \mathbf{v} \rangle$ which has the following properties:

$$\text{if } \mathbf{u} \neq 0 \text{ then } \langle \mathbf{u}, \mathbf{u} \rangle > 0$$

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$$

$$\langle a^i \mathbf{e}_i, b^j \mathbf{e}_j \rangle = a^i b^j \langle \mathbf{e}_i, \mathbf{e}_j \rangle$$

An inner product may be expressed in the following equivalent ways:

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u} \cdot \boldsymbol{v} = (u^i \mathbf{e}_i) \cdot (v^j \mathbf{e}_j) = u^i v^i G_{ij}$$

where the real numbers

$$G_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j$$

are components of a symmetric, positive definite matrix G called the metric tensor. The *Euclidean* inner product is distinguished by $G_{ij} = \delta_{ij}$. In the Euclidean case

$$\boldsymbol{u} \cdot \boldsymbol{v} = u^t v = v^t u$$

Given an inner product we can define the length or norm of a vector and the angle between two vectors. The norm of $\mathbf{u}$ is

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$$

The angle between $\mathbf{u}$ and $\mathbf{v}$ is

$$\arccos\left(\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right)$$

Tensors are important objects in rigid body mechanics and we now set down the basics of tensor algebra. We start with the algebraic definition of tensors of rank 2.[2] A tensor of rank 2 assigns a real number to a pair of vectors or covectors and is linear in each argument. A covariant tensor $\boldsymbol{T}$ of rank 2 assigns a real number to pairs of vectors $\boldsymbol{T}(\mathbf{u}, \mathbf{v}) \in \mathbb{R}$. The metric tensor is an example of a covariant tensor. A contravariant tensor $\boldsymbol{T}$ of rank 2 assigns a real number to pairs of covectors $\boldsymbol{T}(\boldsymbol{v}, \boldsymbol{\nu}) \in \mathbb{R}$. A mixed tensor $\boldsymbol{T}$ of rank 2 assigns a real number to a vector–covector pair $\boldsymbol{T}(\mathbf{u}, \boldsymbol{\nu}) \in \mathbb{R}$. The components of a tensor are its values on basis vectors. Thus, a covariant tensor $\boldsymbol{A}$ has components $a_{ij} = \boldsymbol{A}(\mathbf{e}_i, \mathbf{e}_j)$ and a mixed tensor $\boldsymbol{B}$ has components $b^i_j = \boldsymbol{B}(\mathbf{e}_i, \epsilon^j)$.

Tensors can be formed from tensor products of vectors and covectors. The tensor products are denoted with the symbol $\otimes$ and are defined by their action on their arguments. Thus we define a covariant tensor $\boldsymbol{v} \otimes \boldsymbol{\nu}$ by $\boldsymbol{v} \otimes \boldsymbol{\nu}(\mathbf{u}, \mathbf{v}) = \boldsymbol{v}(\mathbf{u})\boldsymbol{\nu}(\mathbf{v})$ and define the mixed tensor $\mathbf{u} \otimes \boldsymbol{\nu}$ by $\mathbf{u} \otimes \boldsymbol{\nu}(\boldsymbol{v}, \mathbf{v}) = \boldsymbol{v}(\mathbf{u})\boldsymbol{\nu}(\mathbf{v})$. It is not the case that every tensor is a tensor product but every tensor is a linear combination of tensor products of basis vectors. For example,

$$\boldsymbol{T}(\mu_i \epsilon^i, v^j \mathbf{e}_j) = \mu_i v^j \boldsymbol{T}(\epsilon^i, \mathbf{e}_j)$$

Addition and scalar multiplication of the tensors can be defined by their action on vectors

$$(a\boldsymbol{T} + b\boldsymbol{S})(\mathbf{u}, \mathbf{v}) = a\boldsymbol{T}(\mathbf{u}, \mathbf{v}) + b\boldsymbol{S}(\mathbf{u}, \mathbf{v})$$

In the case of mixed tensors, the result of this construction is a new vector space $\boldsymbol{V} \otimes \boldsymbol{V}^*$. When $\boldsymbol{V}$ has dimension n, $\boldsymbol{V} \otimes \boldsymbol{V}^*$ has dimension n^2 . The same construction can be carried through for pairs of covectors and the resulting vector space $\boldsymbol{V}^* \otimes \boldsymbol{V}^*$ again has dimension n^2.

The subspace $\bigwedge^2 \boldsymbol{V}^*$ of $\boldsymbol{V}^* \otimes \boldsymbol{V}^*$ consists of skew-symmetric covariant tensors, that is, of covariant tensors $\boldsymbol{\Omega}$ which satisfy

$$\boldsymbol{\Omega}(\mathbf{u}, \mathbf{v}) = -\boldsymbol{\Omega}(\mathbf{v}, \mathbf{u})$$

2) The development extends to any rank but we will need only rank 2.

Example 1.1 *The tensor* $\boldsymbol{T} = \boldsymbol{\nu} \otimes \boldsymbol{v} - \boldsymbol{v} \otimes \boldsymbol{\nu}$ *is skew-symmetric because*

$$\boldsymbol{T}(\mathbf{u}, \mathbf{v}) = \boldsymbol{\nu}(\mathbf{u})\boldsymbol{v}(\mathbf{v}) - \boldsymbol{v}(\mathbf{v})\boldsymbol{\nu}(\mathbf{u})$$

and

$$\boldsymbol{T}(\mathbf{v}, \mathbf{u}) = \boldsymbol{\nu}(\mathbf{v})\boldsymbol{v}(\mathbf{u}) - \boldsymbol{v}(\mathbf{u})\boldsymbol{\nu}(\mathbf{v}) = -\boldsymbol{T}(\mathbf{u}, \mathbf{v}). \quad \diamondsuit$$

The members of $\bigwedge^2 \boldsymbol{V}^*$ are called 2-forms over $\boldsymbol{V}$. Covectors, members of $\boldsymbol{V}^*$, are also called 1-forms. There is a product, the *wedge product*, which produces a 2-form $\boldsymbol{\omega} \wedge \boldsymbol{\mu}$ from two 1-forms $\boldsymbol{\omega}$ and $\boldsymbol{\mu}$. The wedge product is defined by its action on its arguments

$$\boldsymbol{\omega} \wedge \boldsymbol{\mu}(\mathbf{u}, \mathbf{v}) = \boldsymbol{\omega}(\mathbf{u})\boldsymbol{\mu}(\mathbf{v}) - \boldsymbol{\omega}(\mathbf{v})\boldsymbol{\mu}(\mathbf{u})$$

The wedge product is basic to the geometric treatment of Hamiltonian mechanics.

Now we follow [1] to establish the connection between rank 2 mixed tensors and linear operators. If $\mathcal{A}$ is a linear operator, let $\boldsymbol{T}_{\mathcal{A}}$ be the tensor defined by the action $\boldsymbol{T}_{\mathcal{A}}(\mathbf{v}, \boldsymbol{\nu}) = \boldsymbol{\nu}(\mathcal{A}(\mathbf{v}))$. The components of $\boldsymbol{T}_{\mathcal{A}}$ are given by

$$\boldsymbol{T}_{\mathcal{A}}{}^i_j = \boldsymbol{T}_{\mathcal{A}}(\mathbf{e}_j, \boldsymbol{\epsilon}^i) = \boldsymbol{\epsilon}^i(\mathcal{A}(\mathbf{e}_j)) = A^i_j$$

Thus the components of $\boldsymbol{T}_{\mathcal{A}}$ are the same as those of $\boldsymbol{A}$. We will exploit this correspondence, borrowing from the theory of dyadics, to represent linear transformations by

$$\mathcal{A}(u^k \mathbf{e}_k) = \mathbf{e}_i A^i_j \epsilon^j (u^k \mathbf{e}_k) = \mathbf{e}_i A^i_j u^k \delta_{jk} = \mathbf{e}_i A^i_j u^j \tag{1.1}$$

Thus the action of the second-rank tensor depends on its object. If applied to a vector–covector pair it produces a scalar and if applied to a vector it returns another vector. This is reminiscent of the dot and double-dot products of dyadics [2] which are closely related to second-rank tensors.

The combinations $\boldsymbol{\epsilon}\mathbf{e}$ and $\mathbf{e} \otimes \boldsymbol{\epsilon}$ arise frequently in the manipulation of tensors. The product $\boldsymbol{\epsilon}\mathbf{e}$ is the identity matrix because

$$\begin{bmatrix} \boldsymbol{\epsilon}^1 \\ \boldsymbol{\epsilon}^2 \\ \boldsymbol{\epsilon} \end{bmatrix} \left(\begin{bmatrix} \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3 \end{bmatrix} \right) = [\boldsymbol{\epsilon}^i(\mathbf{e}_j)] = I$$

The product $\mathbf{e} \otimes \boldsymbol{\epsilon}$ is the identity operator because

$$\begin{bmatrix} \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3 \end{bmatrix} \otimes \begin{bmatrix} \boldsymbol{\epsilon}^1 \\ \boldsymbol{\epsilon}^2 \\ \boldsymbol{\epsilon} \end{bmatrix} = \mathbf{e}_i \otimes \boldsymbol{\epsilon}^i$$

and for any vector $\boldsymbol{v} = v^i \mathbf{e}_i$

$$(\mathbf{e}_i \otimes \boldsymbol{\epsilon}^i)(v^k \mathbf{e}_k) = v^k \mathbf{e}_i \delta_{ik} = \boldsymbol{v}$$

The above representation leads to a succinct representation of linear operators

$$\mathcal{A} = \mathbf{e}A\boldsymbol{\epsilon}$$

which is shorthand for $\mathcal{A} = A^i_j \mathbf{e}_i \otimes \boldsymbol{\epsilon}^i$ so that

$$\mathcal{A}(\mathbf{v}) = \mathbf{e}A\boldsymbol{\epsilon}(\mathbf{e}v) = \mathbf{e}Av$$

Now consider the effect of a change of basis. Suppose $\mathbf{e}$ and $\bar{\mathbf{e}}$ are bases of $\boldsymbol{V}$ linearly related by

$$\mathbf{e} = \bar{\mathbf{e}}B$$

Then a vector $\mathbf{v}$ has the representations

$$\mathbf{v} = \mathbf{e}v = \bar{\mathbf{e}}Bv = \bar{\mathbf{e}}\bar{v}$$

so that the components in the two bases are related by

$$\bar{v} = Bv$$

The covector relations

$$\boldsymbol{I} = \mathbf{e}\boldsymbol{\epsilon} = \bar{\mathbf{e}}B\boldsymbol{\epsilon} = \bar{\mathbf{e}}\bar{\boldsymbol{\epsilon}}$$

give

$$\bar{\boldsymbol{\epsilon}} = B\boldsymbol{\epsilon}$$

Thus a covector $\boldsymbol{\mu}$ has the representations

$$\boldsymbol{\mu} = \mu\boldsymbol{\epsilon} = \mu B^{-1}\bar{\boldsymbol{\epsilon}} = \bar{\mu}\bar{\boldsymbol{\epsilon}}$$

so that the components in the two bases are related by

$$\mu = \bar{\mu}B$$

The effect of change of basis on a linear operator follows immediately

$$\mathcal{A} = \mathbf{e}A\boldsymbol{\epsilon} = \bar{\mathbf{e}}BAB^{-1}\bar{\boldsymbol{\epsilon}} = \bar{\mathbf{e}}\bar{A}\bar{\boldsymbol{\epsilon}}$$

or

$$\bar{A} = BAB^{-1}$$

Given the linear operators $\mathcal{A}$, $\mathcal{B}$ with the matrix representations A, B, the matrix representation of the composite map $\mathcal{B} \circ \mathcal{A}$, $\mathcal{A}$ followed by $\mathcal{B}$, is the matrix product BA. The basis vectors transform as

$$\mathbf{e}' = \mathbf{e}BA$$

and the components transform as

$$v' = BAv$$

For any n the vector space $\mathbb{R}^n$ has the standard basis $\{e_i\}$ where e_i is the column of length n having a single nonzero entry, 1 in row i. The inner product in the standard basis is

$$\boldsymbol{u} \cdot \boldsymbol{v} = u^t = v^u = u_i v_i$$

where no distinction is made between u_i and u^i. In $\mathbb{R}^3$ there is also a *vector product* or cross product

$$(a_i \mathbf{e}_i) \times (b_j \mathbf{e}_j) = \epsilon_{ijk} a_i b_j \mathbf{e}_k = (a_2 b_3 - a_3 b_2)\mathbf{e}_1 + (a_3 b_1 - a_1 b_3)\mathbf{e}_2 + (a_1 b_2 - a_2 b_1)\mathbf{e}_3$$

where ϵ_{kij} is the permutation symbol

$$\epsilon_{kij} = \begin{cases} 1 & \text{if } i\,j\,k \text{ is an even permutation of } 123 \\ -1 & \text{if } i\,j\,k \text{ is an odd permutation of } 123 \\ 0 & \text{otherwise} \end{cases}$$

We will have no further need in this chapter to distinguish between subscripts and superscripts, so subscripts will be used. In this case matrix entries will be denoted by A_{ij} with row index i and column index j. Superscripts will return with a vengeance when generalized coordinates are considered.

1.1.2
Rotation Operators on $\mathbb{R}^3$

A rotation is a linear transformation, $\mathcal{R}$, that fixes the origin, preserves the lengths of vectors, and preserves the orientation of bases. That is,

$$\mathcal{R} : \mathbf{V} \to \mathbf{V} : \boldsymbol{x} \mapsto \mathcal{R}(\boldsymbol{x})$$

$$\mathcal{R}(\mathbf{0}) = \mathbf{0}$$

$$\mathcal{R}(\boldsymbol{x}) \cdot \mathcal{R}(\boldsymbol{x}) = \boldsymbol{x} \cdot \boldsymbol{x}$$

$$\mathcal{R}(\mathbf{e}_1) \cdot (\mathcal{R}(\mathbf{e}_2) \times \mathcal{R}(\mathbf{e}_3)) = \mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3)$$

The length-preserving property becomes, in a matrix form,

$$\mathcal{R}(\boldsymbol{x}) \cdot \mathcal{R}(\boldsymbol{x}) = (x_j R_{ij} \mathbf{e}_i) \cdot (x_l R_{kl} \mathbf{e}_k) = x_j x_l R_{ij} R_{il} = x_i x_i \text{ for all } \{x_i\} \in \mathbb{R}^3$$

which implies that

$$R_{ij} R_{il} = \delta_{jl} \quad \text{or} \quad R^t R = I$$

This defines an *orthogonal matrix*. The orientation preserving condition becomes

$$R_{k1} \epsilon_{kij} R_{i2} R_{j3} = 1 \quad \text{or} \quad \det R = +1$$

When R and S are orthogonal, then $(RS)^t RS = S^t R^t RS = I$. When R is orthogonal $R^{-1} = R^t$ and $(R^{-1})^t R^{-1} = RR^t = I$. Therefore RS and R^{-1} are also orthogonal. Clearly I is orthogonal. The product of orthogonal matrices preserves the determinant, that is, if $\det R = \det S = 1$ then $\det RS = \det R \det S = 1$. If $\det R = 1$ then $\det R^{-1} = 1$ because $\det RR^{-1} = \det I = 1$.

It follows from these facts that the orthogonal matrices of determinant 1 form a group.[3] The group is called $SO(3)$ – the *special orthogonal group* of order 3. This group has the additional structure of a three-dimensional manifold and is therefore a Lie group. In the next section we begin to study parameterizations, or coordinates, of $SO(3)$. The basics of the Lie group theory are outlined in Appendix C.

In addition to preserving lengths, orthogonal matrices preserve angles because they preserve inner products

$$(Ru)^t Rv = u^t R^t Rv = u^t v$$

Members of $SO(3)$ also preserve $\mathbb{R}^3$ vector products in the sense that $A(x \times y) = Ax \times Ay$. To prove this, it is enough to show it for e_1, e_2, e_3, the standard basis of $\mathbb{R}^3$. Let $A \in SO(3)$ be presented in terms of its orthonormal columns, $A = [c_1 c_2 c_3]$ so that $Ae_i = c_i$. Then

$$A(e_i \times e_j) = A(\epsilon_{ijk} e_k) = \epsilon_{ijk} Ae_k = \epsilon_{ijk} c_k = c_i \times c_j = Ae_i \times Ae_j$$

Every member of $SO(3)$ fixes not only the origin but actually an entire line. This follows from the structure of eigensystems of rotation matrices. The length preserving property of a rotation requires that eigenvalues have magnitude 1. To show this we must allow for complex eigenvalues and eigenvectors and use the norm $\|x\|^2 = (x^*)^t x$ where x^* is the complex conjugate of x. Then $Rx = \lambda x$ implies that $\|Rx\|^2 = \lambda^* \lambda (x^*)^t x = |\lambda|^2 \|x\|^2$. In other words $\|Rx\| = \|x\|$ implies $|\lambda| = \pm 1$. The characteristic polynomial of a rotation matrix is a cubic and one of its roots must be $+1$ because $\det R = 1$. Thus, any eigenvector corresponding to $\lambda = 1$ is fixed and real. The set of all such eigenvectors forms the *axis of rotation*.

1.1.3
Rotations Specified by Axis and Angle

First consider plane rotations. Represent vectors as complex numbers $(x, y) \leftrightarrow x + \imath y = \rho \exp(\imath\theta)$. Then a counterclockwise rotation by angle ϕ is simply multiplication by $\exp(\imath\phi)$: $z \rightarrow \exp(\imath\phi) z = \rho \exp[\imath(\theta + \phi)]$. In rectangular components

$$x + \imath y = z \rightarrow z' = e^{\imath\theta}(x + \imath y) = (x \cos\theta - y \sin\theta) + \imath(x \sin\theta + y \cos\theta)$$

3) A group G is a set equipped with a binary operation such that if $a, b \in G$ then $ab \in G$. There is an identity element e such that $ea = a$ for every $a \in G$ and for every $a \in G$ there is an inverse a^{-1} such that $aa^{-1} = e$.

This shows that rotations by angle θ about the x, y, z axes are represented by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{pmatrix} \quad \begin{pmatrix} c & 0 & s \\ 0 & 1 & 0 \\ -s & 0 & c \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

respectively, where $c = \cos\theta,\ s = \sin\theta$.

These matrices have alternative expressions as operators which emphasize the role of the axis of rotation and the rotation angle. To state the alternative forms we first introduce the idea of duality between $\mathbb{R}^3$ and $\mathfrak{so}(3)$, which is the set of all skew-symmetric 3×3 matrices, that is matrices which have the property that $A^t = -A$. In fact $\mathfrak{so}(3)$ is the Lie algebra of the Lie group $SO(3)$ (see Appendix C). To each vector v there corresponds a skew-symmetric matrix $\widehat{v}$,

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix} = \widehat{v}$$

and to each skew-symmetric matrix M there is a vector $\vec{M}$,

$$M = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} \longleftrightarrow \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \vec{M}$$

In the language of Lie algebra $\widehat{v} = ad_v$ (see Section C.3.2).

Using the canonical basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$, this is expressed in terms of tensor products as

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \longleftrightarrow u_1\widehat{\mathbf{i}} + u_2\widehat{\mathbf{j}} + u_3\widehat{\mathbf{k}}$$

where

$$\widehat{\mathbf{i}} = \mathbf{k} \otimes \mathbf{j} - \mathbf{j} \otimes \mathbf{k} \qquad \widehat{\mathbf{j}} = \mathbf{i} \otimes \mathbf{k} - \mathbf{k} \otimes \mathbf{i} \qquad \widehat{\mathbf{k}} = \mathbf{j} \otimes \mathbf{i} - \mathbf{i} \otimes \mathbf{j}$$

are the invariant forms of the $\widehat{\ }$ operator. Here we have identified $\mathbb{R}^3$ and $(\mathbb{R}^3)^*$ via

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \longleftrightarrow \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix}$$

or

$$\mathbf{u} \otimes \boldsymbol{v} \longleftrightarrow uv^t$$

With these preliminaries we can write that the rotations about the coordinate axes are

$$\mathbf{i}\otimes\mathbf{i}+c(I-\mathbf{i}\otimes\mathbf{i})+s\widehat{\mathbf{i}}$$
$$\mathbf{j}\otimes\mathbf{j}+c(I-\mathbf{j}\otimes\mathbf{j})+s\widehat{\mathbf{j}}$$
$$\mathbf{k}\otimes\mathbf{k}+c(I-\mathbf{k}\otimes\mathbf{k})+s\widehat{\mathbf{k}}$$

These expressions can be generalized to an arbitrary axis of rotation determined by the unit vector $\mathbf{n}$. The form of the expressions suggests that one might construct a basis containing $\mathbf{n}$ and write immediately that

$$\mathcal{R}_{\mathbf{n}}(\theta)=\mathbf{n}\otimes\mathbf{n}+\cos\theta(I-\mathbf{n}\otimes\mathbf{n})+\sin\theta\widehat{\mathbf{n}}$$

and we now proceed to justify this. Let P represent the operator relative to the standard basis which projects a vector onto the axis of rotation (Fig. 1.1)

$$P=nn^t$$

P has the property that, for any r, Pr is a multiple of n

$$Pr=nn^tr=\lambda n \qquad \lambda=n^tr$$

and P is idempotent

$$P^2=(nn^t)(nn^t)=n(n^tn)n^t=nn^t=P$$

$I-P$ then represents the operator which projects a vector onto the plane normal to $\mathbf{n}$ because for any r

$$n^t(I-P)r=n^tr-n^tnn^tr=0$$

and

$$(I-P)^2=I-P-P+P^2=I-P$$

Let $N=\widehat{n}$. Any rotation of r through an angle θ leaves Pr invariant and rotates $(I-P)r$ by an angle θ in the plane spanned by $(I-P)r$, $n\times r$. Thus the rotation is given by $r'=R_n(\theta)r$ with

$$R_n(\theta)=nn^t+\cos\theta\left(I-nn^t\right)+\sin\theta\, N$$

It is easy to show that the operators $P, I-P, N$ form a closed system defined by Table 1.1. Here we sketch the proof of one entry as an illustration

$$N^2r=n\times(n\times r)=(r\cdot n)n-r=(P-I)r$$

The relations in Table 1.1 can be used to recast the rotation operator entirely in terms of N,

$$R=I+\sin\theta N+(1-\cos\theta)N^2 \qquad (1.2)$$