

Springer Series in Computational Physics

Editors: R. Glowinski M. Holt P. Hut H.B. Keller J. Killeen
S. A. Orszag V.V. Rusanov

A Computational Method in Plasma Physics

F. Bauer, O. Betancourt, P. Garabedian

Implementation of Finite Element Methods for Navier-Stokes Equations

F. Thomasset

Finite-Difference Techniques for Vectorized Fluid Dynamics Calculations

Edited by D. Book

Unsteady Viscous Flows. D. P. Telionis

Computational Methods for Fluid Flow. R. Peyret, T. D. Taylor

Computational Methods in Bifurcation Theory and Dissipative Structures

M. Kubicek, M. Marek

Optimal Shape Design for Elliptic Systems. O. Pironneau

The Method of Differential Approximation. Yu. I. Shokin

Computational Galerkin Methods. C. A. J. Fletcher

Numerical Methods for Nonlinear Variational Problems

R. Glowinski

Numerical Methods in Fluid Dynamics. Second Edition

M. Holt

Computer Studies of Phase Transitions and Critical Phenomena

O. G. Mouritsen

Finite Element Methods in Linear Ideal Magnetohydrodynamics

R. Gruber, J. Rappaz

Numerical Simulation of Plasmas. Y.N. Dnestrovskii, D.P. Kostomarov

Computational Methods for Kinetic Models of Magnetically Confined Plasmas

J. Killeen, G.D. Kerbel, M.C. McCoy, A.A. Mirin

Spectral Methods in Fluid Dynamics

C. Canuto, M.Y. Hussaini, A. Quarteroni, T.A. Zang

Computational Techniques for Fluid Dynamics 1. Second Edition

Fundamental and General Techniques. C. A. J. Fletcher

Computational Techniques for Fluid Dynamics 2. Second Edition

Specific Techniques for Different Flow Categories. C. A. J. Fletcher

Methods for the Localization of Singularities in Numerical Solutions of Gas Dynamics Problems. E. V. Vorozhtsov, N. N. Yanenko

Classical Orthogonal Polynomials of a Discrete Variable

A. F. Nikiforov, S. K. Suslov, V. B. Uvarov

Flux Coordinates and Magnetic Field Structure:

A Guide to a Fundamental Tool of Plasma Theory

W.D. D'haeseleer, W.N.G. Hitchon, J.D. Callen, J.L. Shohet

W.D. D'haeseleer W.N.G. Hitchon
J.D. Callen J.L. Shohet

Flux Coordinates and Magnetic Field Structure

A Guide to a Fundamental Tool of Plasma Theory

With 40 Figures

Springer-Verlag

Berlin Heidelberg New York London
Paris Tokyo Hong Kong Barcelona

Dr. William Denis D'haeseleer, Associate Professor (part time)

Katholieke Universiteit Leuven, School of Engineering,
Arenbergkasteel, Kard. Mercierlaan 94, B-3001 Leuven, Belgium
and

The NET Team, Max-Planck-Institut für Plasmaphysik,
Boltzmannstraße 2, D-8046 Garching bei München, Fed. Rep. of Germany

Dr. William Nicholas Guy Hitchon, Associate Professor

Dr. J. Leon Shohet, Professor

University of Wisconsin-Madison, Department of Electrical and Computer Engineering,
1415 Johnson Drive, Madison, WI 53706, USA

Dr. James D. Callen, D. W. Kerst Professor

University of Wisconsin-Madison,
Departments of Nuclear Engineering & Engineering Physics and Physics,
1500 Johnson Drive, Madison, WI 53706-1687, USA

Editors

R. Glowinski

Institut de Recherche d'Informatique
et d'Automatique (INRIA)
Domaine de Voluceau,
Rocquencourt, B. P. 105
F-78150 Le Chesnay, France

M. Holt

College of Engineering and
Mechanical Engineering
University of California
Berkeley, CA 94720, USA

P. Hut

The Institute of Advanced Study
School of Natural Sciences
Princeton, NJ 08540, USA

H. B. Keller

Applied Mathematics 101-50
Firestone Laboratory
California Institute of Technology
Pasadena, CA 91125, USA

J. Killeen

Lawrence Livermore Laboratory
P.O. Box 808
Livermore, CA 94551, USA

S. A. Orszag

Department of Mechanical and
Aerospace Engineering
Princeton University
Princeton, NJ 08544, USA

V. V. Rusanov

Keldysh Institute
of Applied Mathematics
4 Miusskaya pl.
SU-125047 Moscow, USSR

ISBN 3-540-52419-3 Springer-Verlag Berlin Heidelberg New York

ISBN 0-387-52419-3 Springer-Verlag New York Berlin Heidelberg

Library of Congress Cataloging-in-Publication Data. Flux coordinates and magnetic field structure : a guide to a fundamental tool of plasma theory / W. D. D'haeseleer . . . [et al.]. p. cm. — (Springer series in computational physics) Includes bibliographical references and index. ISBN (invalid) 3-540-52419-3 : 1. Plasma (Ionized gases) 2. Magnetic fields. I. D'haeseleer, W. D. (William Denis), 1957 – .II. Series. QC718.5.M3F48 1990 530.4'41—dc20 90-10333

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in other ways, and storage in data banks. Duplication of this publication or parts thereof is only permitted under the provisions of the German Copyright Law of September 9, 1965, in its current version, and a copyright fee must always be paid. Violations fall under the prosecution act of the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1991 · Printed in Germany

The use of registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Asco Trade Typesetting Ltd., Hong Kong
56/3140-543210 – Printed on acid-free paper

*Dedicated to the memory
of Michel Antoine D'haeseleer
and
ad majorem Dei gloriam*

Preface

During the course of our research work on the theory of transport phenomena in tokamaks and stellarators, we learned to appreciate the convenience and utility of the special types of coordinate systems in which the magnetic field lines appear as straight lines. These so-called *flux coordinates* sweep, so to speak, the complexity of the magnetic-field structure “under the rug”, and permit the user to concentrate on the (other) physical issues.

Although frequently used in theoretical calculations, at least in a formal way, it is striking that these coordinate systems are employed with little or no consistency, often in (over-)simplified form, and, regrettably, sometimes incorrectly. In addition, every researcher or group seems to have an individual preference concerning the type of flux-coordinate system used. It was rather surprising that there was no comprehensive review available that relates these flux-coordinate systems to each other in a rigorous but readable way.

Therefore, the idea grew to include a short survey of this subject as an appendix to one of our writings on transport. However, we soon discovered the unrealistic nature of our attempt and had to recognize that it was merely wishful thinking. In order to do a proper job, by which we mean producing a document that is comprehensive but still comprehensible for an audience of non-specialists, we realized that the quantity of material to be covered would by far surpass the volume of a reasonable appendix. This is the genesis of this book.

As we stated above, the target group of this monograph consists mainly of non-specialists. More specifically, we hope to reach an audience as varied as beginning graduate students in the field of plasmas and controlled fusion, as well as established researchers (especially experimentalists and those who undertake modeling in trying to interpret experimental results based on the available theories). Even for physicists with a more general interest in mathematical physics, or for applied mathematicians, this monograph might be useful in that it provides a self-contained treatment of curvilinear coordinates. The material on flux coordinates is then to be seen as an extensive example, illustrating the fundamental concepts.

In this spirit, we decided to concentrate our efforts on writing a document that is easily readable, but still quite rigorous. It aims to *teach* the reader. Therefore, we develop almost everything from first principles and cover the subject step by step, culminating in a fairly advanced level at the end of the treatise. In addition to the bare presentation of formulae and expressions, we give clarifications and provide comments and interpretations which help to put everything into context.

The book is divided into three parts.

In Part I we establish the foundation on which the remainder of the text is built. It comprises a first chapter on basics concerning vectors, differential geometry and curvilinear coordinates (Chap. 2), a second chapter deals with tensors (Chap. 3) and the last chapter of Part I (Chap. 4) deals with the basic concepts of the magnetic-field geometry in toroidal and open-ended devices. Chapters 2 and 4 in particular should be considered as prerequisites for the rest.

Part II treats the various distinct flux-coordinate systems. In Chaps. 5 through 9, we discuss Clebsch coordinates, the generic straight-field-line coordinate system, the Boozer toroidal flux coordinates, Hamada coordinates and the form of flux coordinates in axisymmetric tokamaks. In addition, we treat two coordinate systems that superficially resemble flux coordinates, but which are not flux coordinates: symmetry coordinates in a tokamak and so-called canonical coordinates.

In Part III of this monograph, we consider some topics that deserve attention, but which were tangential to the main issues in the previous chapters. Here, we treat "proper" toroidal coordinates (Chap. 10), the dynamic equilibrium of an ideal tokamak plasma (Chap. 11), the relationship between the proper length of a field line $\int dl/B$ and the specific volume of a flux surface $dV/d\Psi_{\text{tor}}$ (Chap. 12), the transformation properties of vector and tensor components (Chap. 13), and finally two alternative derivations for the divergence of a vector (Chap. 14).

There are a number of people who have been helpful in getting this monograph into its present shape.

We are most grateful to Mrs. Sarah Cohen, for whom it must have been a nightmare to type the various formulae with their variety of superscripts and subscripts. She retired soon after having finished this manuscript; we hope that we are not to blame for this decision.

We also would like to thank Springer-Verlag and its Physics Editor, Dr. H.-U. Daniel, for the professional handling of our request for the publication of this work. We also appreciate that we had the opportunity to update the manuscript with regard to recently (1988, 1989) published scientific literature. Also, we wish to express our gratitude to the referees for their useful comments.

One of us (W. D'H) wishes to thank Dr. J. A. Tataronis for numerous discussions on the subject of magnetic-field structure.

This work was supported by the US Department of Energy through Grants No. DE-FG02-87ER53216, No. DE-FG02-85ER53201 and No. DE-FG02-86ER53218, and in part by the Wisconsin Alumni Research Foundation.

We hope that the reader who works through all the steps given will reach a level where he/she can consider co- and contravariant representations to be quite simple. Whether we have succeeded in fulfilling this expectation must be judged by those patient and persistent enough to have worked through the book.

Garching and Madison,
November 1990

*W.D. D'haeseleer W.N.G. Hitchon
J.D. Callen J.L. Shohet*

Contents

1. Introduction and Overview	1
------------------------------------	---

Part I Fundamental Concepts

2. Vector Algebra and Analysis in Curvilinear Coordinates	7
2.1 Introduction	7
2.2 Reciprocal Sets of Vectors	8
2.3 Curvilinear Coordinates	10
2.3.1 Transformation to Curvilinear Coordinates	10
2.3.2 Tangent-Basis Vectors	11
2.3.3 Reciprocal-Basis Vectors	14
2.3.4 Covariant and Contravariant Components of a Vector	15
2.4 Covariant and Contravariant "Vectors"	17
2.5 Vector Relationships in Curvilinear Coordinates	20
2.5.1 The Metric Coefficients g_{ij} and g^{ij}	20
2.5.2 The Jacobian	23
a) The Jacobian of the Curvilinear Coordinate System	23
b) The Relationship Between J and g	25
2.5.3 The Dot and Cross-Products in Curvilinear Coordinates	26
a) Dot Products	26
b) Cross Products	27
2.5.4 The Differential Elements $dl(i)$, $dS(i)$, d^3R	28
a) The Differential Arc Length $dl(i)$	28
b) The Differential Area Element $dS(i)$	28
c) The Differential Volume Element d^3R	29
2.6 Vector Differentiation in Curvilinear Coordinates	30
2.6.1 The Covariant Derivative	30
a) Differentiation of Vector Components and Basis Vectors	30
b) The Covariant Derivative in Terms of the Metric Coefficients	32
c) The Christoffel Symbols	34
2.6.2 The Del Operator	35
a) Definition of the Del Operator	35

b) Gradient	35
c) Divergence	36
d) Curl	36
2.7 The Parallel and Perpendicular Components of a Vector	37
2.8 A Summary of Vector Related Identities	38
3. Tensorial Objects	40
3.1 Introduction	40
3.2 The Concept of Tensors; A Pragmatic Approach	40
3.3 Dot Product, Double Dot Product, Contraction	44
3.4 The Relationship Between Covariant, Contravariant and Mixed Components	47
3.5 Special Tensors	48
3.5.1 The Kronecker Delta and the Metric Tensor	48
3.5.2 Levi-Civita and Christoffel Symbols	50
a) Levi-Civita Symbols	50
b) Christoffel Symbols	52
3.6 Tensor and Dyadic Identities	52
3.7 Suggestions for Further Reading	54
4. Magnetic-Field-Structure-Related Concepts	55
4.1 The Equation of a Magnetic-Field Line	55
4.2 The Frozen-Flux Theorem	56
4.3 The Magnetic Field-Line Curvature	58
4.4 Magnetic Pressure and Magnetic Tension	63
4.5 Magnetic Surfaces	64
a) Toroidal Systems	65
b) Open-Ended Systems	67
4.6 Curvilinear Coordinate Systems in Confinement Systems with “Simple” Magnetic Surfaces	67
4.6.1 Toroidal Systems	68
a) “The Cylindrical-Toroidal” or “Elementary” Toroidal System	68
b) Generalized Cylindrical-Toroidal Coordinates; Flux Coordinates	70
c) An Illustrative Example	73
d) (ϱ, β, l) Coordinates in Toroidal Systems	73
4.6.2 Open-Ended Systems	76
4.7 Magnetic-Surface Labeling	77
4.7.1 Toroidal Systems	77
4.7.2 Open-Ended Systems	82
4.8 The Rotational Transform in Toroidal Systems	82
4.9 The Flux-Surface Average	85
4.9.1 Toroidal Systems	86
4.9.2 Open-Ended Systems	89
4.9.3 Properties of the Flux-Surface Average	90
4.10 The Magnetic Differential Equation	93

Part II Flux Coordinates

5. The Clebsch-Type Coordinate Systems	100
5.1 Stream Functions	100
5.2 Generic Clebsch Coordinates	100
5.3 Relationship to the Contra- and Covariant Formalism	103
5.4 Boozer-Grad Coordinates	107
6. Toroidal Flux Coordinates	116
6.1 Straight Field-Line Coordinates	116
6.2 Symmetry Flux Coordinates in a Tokamak	122
6.3 Interlude: Non-Flux Coordinates in Tokamaks	126
6.4 Straight Current-Density-Line Coordinates	129
6.5 Covariant \mathbf{B} Components and Their Relationship to the Boozer-Grad Form	131
6.5.1 The Vector Potential	131
6.5.2 Covariant \mathbf{B} Components	133
6.5.3 Relationship to the Boozer-Grad Form	140
6.6 Boozer’s Toroidal Flux Coordinates	142
6.7 Ideal-MHD-Equilibrium Conditions for Toroidally Confined Plasmas	146
6.8 Hamada Coordinates	150
7. Conversion from Clebsch Coordinates to Toroidal Flux Coordinates	156
7.1 The Generic Clebsch Coordinate System (ϱ, v, l)	156
7.2 Boozer-Grad Coordinates (ϱ, v, χ)	157
8. Establishment of the Flux-Coordinate Transformation; A Summary	160
8.1 Toroidal Systems	160
8.2 Open-Ended Systems	164
9. Canonical Coordinates or “Generalized Magnetic Coordinates” ..	165
9.1 Flux Coordinates Versus Canonical Coordinates	165
9.2 On the Existence of Flux Surfaces, Revisited	165
9.3 Flux Coordinates	166
9.4 Canonical Coordinates; The Field-Line Hamiltonian	170
9.5 Practical Evaluation of the Field-Line Hamiltonian	174
Part III Selected Topics <hr/>	
10. “Proper” Toroidal Coordinates	179
10.1 Introduction	179
10.2 Bipolar Coordinates: Intuitive Considerations	179

10.3 The Relationship of 3-D Spherical Coordinates to 2-D Polar Coordinates	181
10.4 "Proper" Toroidal Coordinates as a 3-D Version of Bipolar Coordinates	182
10.5 The Bipolar Coordinate System: A Detailed Analysis	183
11. The Dynamic Equilibrium of an Ideal Tokamak Plasma	186
11.1 Introduction	186
11.2 A Useful Identity in Tokamak Geometry	186
11.3 The Existence of an Electric Field in a Tokamak Plasma with Infinite Conductivity	187
11.4 Motion of the Plasma and the Flux Surfaces	190
11.4.1 Plasma Motion due to $E \times B$ Drift	190
11.4.2 Motion of Flux Surfaces Defined by the Poloidal Disk Flux Ψ_{pol}^d	191
11.4.3 Motion of Flux Surfaces Defined by the Poloidal Ribbon Flux Ψ_{pol}^r	192
11.4.4 Motion of Flux Surfaces Defined by the Toroidal Flux Ψ_{tor}	193
11.5 Constancy of the Rotational-Transform Flux Function	195
11.6 Remarks on the Evolution of a Finite-Resistivity Plasma	198
11.7 The Relationship Between Poloidal Disk and Ribbon Fluxes	199
12. The Relationship Between $\int dl/B$ and $dV/d\Psi_{\text{tor}}$	201
13. Transformation Properties of Vector and Tensor Components	204
13.1 Transformation of the Basis Vectors	204
13.2 Transformation of Vector Components	205
13.3 Transformation of Tensor Components	207
13.4 Transformation of Components of Special Tensors and Symbols	207
13.4.1 The Kronecker Delta	207
13.4.2 Metric Coefficients	208
13.4.3 Levi-Civita Symbols	208
13.4.4 Christoffel Symbols	209
14. Alternative Derivations of the Divergence Formula	211
14.1 A Straightforward Derivation of the Divergence Formula	211
14.2 Divergence-Formula Derivation Employing Christoffel Symbols	212
References	215
Subject Index	221

1. Introduction and Overview

This treatise on magnetic-flux coordinates is intended to present a self-contained treatment of what are the natural coordinate systems for the description of magnetically confined plasmas. *Flux coordinates* are carefully chosen curvilinear coordinates with the property that the equation of a *magnetic-field line*, expressed in those coordinates, takes the form of a *straight line*. Sometimes we shall also refer to these coordinates systems as *straight-B systems*.

There does not appear to be a comprehensive review of the subject and so this document will attempt to provide a comprehensible and fairly complete treatment. This should help to make theoretical plasma calculations more accessible to non-specialists and researchers new to the field. To achieve this, we shall develop the mathematical tools such as tensor analysis in curvilinear coordinates from first principles as far as possible.

There is little or no consistency in definitions, nomenclature, interpretations and even conclusions in the area of tensors, dyadics, curvilinear coordinates and flux coordinates. The use of the words covariant and contravariant vectors is especially confusing, but ideas such as tensors, dyads and dyadics are no better. In this monograph we have tried to retain the most logical and straightforward formulations or the ones put forward most frequently. We are convinced that this important area deserves some clarification.

The above does not imply that no useful treatments of magnetic-flux coordinates exist. Useful contributions containing interesting material on magnetic-flux coordinates, magnetic geometries and/or equilibrium have been written by Morozov and Solov'ev (1966), Shafranov (1966), Solov'ev and Shafranov (1970), Solov'ev (1975), Greene and Johnson (1962), Boozer (1980, 1981, 1982, 1983), Kieras (1982), Hirshman (1982) and Hazeltine and Meiss (1985). Some basic ideas are presented in textbook treatments by Bateman (1978), Miyamoto (1980), Balescu (1988) and White (1989). However, although much of the plasma-physics literature uses these coordinates and some papers provide a discussion of them, all assume considerable familiarity with the subject (or its mathematical foundations). Many times, simplified forms are given for axisymmetric systems, or vacuum fields are used.

Since this monograph is to provide *practical* formulae and concepts, it is not appropriate to dwell on the theoretical framework behind the vector and tensor formalism, originating in an interplay between linear algebra and differential geometry. Rather, the focus is on the geometrical properties of vectors and

tensors in curvilinear coordinates. Because magnetic-flux coordinates are located in the three-dimensional configuration space, we limit our discussion to three dimensions (3-D).

The book consists of three parts.

In Part I, containing Chaps. 2, 3 and 4, the mathematical groundwork is laid and the basic concepts concerning the magnetic geometry in magnetic-fusion devices are introduced. The chapter following this introduction, Chap. 2, will address the vector-related issues. Chapter 3 will generalize some ideas developed in Chap. 2 to tensors, because a vector can be considered as a very simple kind of tensor. In these two chapters, we shall treat vectors, dyads, triads (polyads) and tensors. The related, but for plasma physics unimportant, topic of quaternions will be skipped entirely here. (For a very elementary discussion on quaternions, see Kramer (1981), Chaps. 4 and 28; for more advanced treatments, see Morse and Feshbach (1953), Chap. 1, and Behnke (1974), Vol. 1, Chap. 8.) In Chap. 4, containing a preliminary discussion on magnetic geometries, the following subjects are considered: the equation of a magnetic-field line, the frozen-flux theorem, magnetic pressure and tension, field-line curvature, flux surfaces, types of coordinate systems in devices with “simple” flux surfaces, magnetic-surface labeling, the rotational transform, the flux-surface average and the magnetic differential equation.

Part II deals with the flux-coordinate systems. We treat each of the important flux-coordinate systems and, importantly, find the relationships between them. We start with the Clebsch system (Chap. 5), which is useful mainly in open-ended fusion devices and which has an intimate relationship with toroidal straight- \mathbf{B} coordinates. A relative of the Clebsch system, the “Boozer-Grad” coordinate system, is also treated in considerable detail. The Boozer-Grad system has the property that one of its coordinates reduces to the magnetic scalar potential in vacuum.

In Chapt. 6, the flux coordinates for general toroidal systems are handled. After having introduced the generic form for toroidal straight- \mathbf{B} coordinates in arbitrary toroidal confinement systems (with “good” magnetic surfaces), we specialize to the simplified case of an axisymmetric tokamak. In a next section, we discuss a special type of coordinate system that can always be constructed when an ignorable coordinate is present, but which is generally distinct from a flux-coordinate system. An important fact to recognize is that flux coordinates, in a general non-axisymmetric toroidal system, are not uniquely determined. Depending on how the extra degree of freedom is “consumed”, the so-called toroidal Boozer flux-coordinate system or the Hamada coordinate system emerges. The Boozer system has the property that the periodic part of the scalar magnetic potential vanishes. There exists a simple linear relationship between the Boozer angle variables and a streaming function of the Boozer-Grad system. That streaming function reduces to the scalar magnetic potential in vacuum. The Hamada system has the convenience that both the magnetic-field lines and the current-density lines are straight. A further advantage is that the Jacobian is a constant.

In Chap. 7, we comment on the conversion from Clebsch-type coordinates to toroidal angle flux coordinates. Both the generic Clebsch system and the Boozer-Grad systems are treated.

Chapter 8 deals with the establishment of the flux-coordinate transformation. Besides being a recapitulation of the methods scattered through the text, it gives us the opportunity to say a few words on equilibrium calculations.

A second non-flux-coordinate system, appearing in Part II of this monograph, is developed in Chap. 9. The so-called “canonical coordinates” have the property that the magnetic-field expression in those coordinates resembles that written in flux coordinates. For special cases, these generalized coordinates reduce to flux coordinates. We shall stress the similarities and differences between the two types of coordinates.

Some important but less crucial topics have been deferred to Part III, Selected Topics, in order not to interrupt the flow of reading in the earlier parts. In Chap. 10, we treat “proper” toroidal coordinates, which are a three-dimensional version of bipolar coordinates. In Chap. 11, we examine the time-dependent tokamak dynamics of an ideal plasma. Here we show that both the plasma and the flux surfaces shrink radially inward under the influence of a perpendicular electric field, which is in turn induced by the change in transformer flux. In Chap. 12, we study the relationship between the specific flux volume $dV/d\Psi_{\text{tor}}$ and the integral $\int dl/B$. Chapter 13 deals with the transformation properties of vector and tensor components. It will explain where the names covariant and contravariant come from. Finally, in Chap. 14, we consider two alternatives to the derivation of the divergence formula given in the text.

A few comments on the notation used are appropriate. Vectors are represented by ***boldface Roman italic*** symbols, e.g., \mathbf{A} . Second-order tensors or dyadics are denoted by ***boldface sanserif italic*** characters, e.g., \mathbf{A} .

These symbols are more elegant than the “arrow” or “over/under bar” notations: \vec{A} , \bar{A} or \underline{A} for vectors (which are first-order tensors) and \overrightarrow{A} , \overleftarrow{A} or \overline{A} for second-order tensors and dyadics. These notations have the advantage, though, that they reflect the tensorial order: n “arrow heads” or “over/under bars” for n -th-order tensors.

Because we ran out of reasonable typographical symbols, we shall use the “arrow” notation to denote the third-order tensors that we shall encounter in Chap. 3, e.g., $\overleftrightarrow{\mathbf{A}}$.

Fundamental Concepts

Before one is ready to attack the different flux-coordinate systems studied in this book, the reader must possess a reasonable background in vector and tensor quantities in general curvilinear coordinate systems, on the one hand, and in magnetic-field-structure related concepts on the other.

Chapters 2 and 3 deal with mathematical preliminaries. Chapter 2, on vectors, is a prerequisite for the remainder of the book; it introduces basic ideas such as co- and contravariant vector components, the Jacobian, the metric coefficients and vector differentiation (including Christoffel symbols). Chapter 3, treating tensors, is not really required to understand what follows, but it completes the discussion of the previous chapter. Tensors, dyads, dyadics and the fundamental or metric tensor are considered.

Chapter 4, then again, is essential. Here we consider a variety of properties of the magnetic field and its structure in toroidal and open-ended devices. These “magnetic preliminaries” give us the vocabulary to deal with the rest of the text. This chapter deals with the following subjects. First, the equation of a magnetic-field line is derived. It will be clear that the \mathbf{B} -field components to be used are of the contravariant type. Next, we prove the “frozen-flux” theorem. We show that in a plasma with infinite conductivity the magnetic field is frozen into the plasma. Then, we discuss the concepts of magnetic pressure and tension. These seem to be very useful concepts for qualitative arguments in magnetically-confined plasmas. In the next section, we carefully define the magnetic-field-line curvature. The “normal” and “geodesic” curvature with respect to the flux surfaces are discussed. As a next topic, flux surfaces are rigorously defined. We say a few words on their “existence” in actual devices. Then the stage is set for a qualitative discussion of curvilinear coordinates in toroidal devices with “simple” magnetic surfaces. Here, we point out that a Clebsch system can be rather confusing in toroidal devices and that an angular system is preferable. In the following section, the different ways to label magnetic-flux surfaces are explained. We concentrate on the poloidal and toroidal magnetic fluxes and the volume inside a flux surface. The section before last of Chap. 4 defines the concept of rotational transform and gives an expression in terms of the ratio of (the change of) poloidal and toroidal magnetic fluxes. Finally, we consider an important type of differential equation, known as the magnetic differential equation because it contains the operator $\mathbf{B} \cdot \nabla$. We sketch the solution procedure and discuss its solubility conditions.

2. Vector Algebra and Analysis in Curvilinear Coordinates

2.1 Introduction

In this chapter we derive and compile some fundamental relations concerning the 3-D vector algebra and calculus in curvilinear coordinates. More detailed treatments (containing additional material which is less important for our needs) can be found in standard works on vectors and tensors such as Borisenko and Tarapov (1968), Wrede (1963), and Spiegel (1959). Classic textbooks on mathematical physics such as Morse and Feshbach (1953), Margenau and Murphy (1976), Mathews and Walker (1970) and Butkov (1968), contain some instructive chapters on the subject matter. Also worth mentioning are the treatments in well-known physics and engineering books such as Symon (1971, Mechanics), Stratton (1941, Electromagnetic Theory), Milne-Thomson (1968, Hydrodynamics) and Bird, Stewart and Lightfoot (1960, Transport Phenomena).

As we try to develop a self-consistent formalism, we must be most careful with the *notation*. Only by consistent and straightforward usage of a correct notation can these mathematical tools prove their power.

In this Chapter, we take a pragmatic approach without spending too much space on abstract foundations on the one hand or elementary concepts on the other. The treatment will be quite rigorous but will assume that the reader knows what a vector, a basis, a curve, a surface, etc. are.

In Sect. 2.2, we introduce the important concept of reciprocal sets of basis vectors. This idea is then used in Sect. 2.3. to define two kinds of basis vectors in a general curvilinear-coordinate system, leading to the concept of covariant and contravariant components of a vector. In Sect. 2.4, co- and contravariant “vectors” are discussed. In the Euclidean space we work in, no such things exist; only *vectors* exist. Section 2.5 treats some properties of curvilinear-coordinate systems: the metric coefficients, the Jacobian, the form of the dot and cross products, and the differential arc length, area and volume. In Sect. 2.6, vector differentiation in curvilinear systems is put in perspective. We explain why “covariant derivative” might not be a good name and show how the Christoffel symbols arise naturally from differentiation of the basis vectors. These symbols are discussed here because they are occasionally used in plasma physics (see e.g., Stacey, 1981). In Sect. 2.6, we also write the proper forms of the gradient, divergence and curl in curvilinear coordinates. Section 2.7 consists of a few words on the perpendicular and parallel components of one vector with respect to another. To conclude the chapter, Sect. 2.8 summarizes the most important vector identities.

2.2 Reciprocal Sets of Vectors

The concept of reciprocal sets of vectors is a crucial building block for the vector description of vectors in curvilinear coordinates. Below we shall see that the usual basis vectors along which a vector is expanded are reciprocal vectors in the sense to be explained.

The sets of vectors $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are called *reciprocal sets of vectors* if:

$$\begin{aligned}\mathbf{A} \cdot \mathbf{a} &= \mathbf{B} \cdot \mathbf{b} = \mathbf{C} \cdot \mathbf{c} = 1 \\ \mathbf{A} \cdot \mathbf{b} &= \mathbf{A} \cdot \mathbf{c} = \mathbf{B} \cdot \mathbf{a} = \mathbf{B} \cdot \mathbf{c} = \mathbf{C} \cdot \mathbf{a} = \mathbf{C} \cdot \mathbf{b} = 0\end{aligned}\quad (2.2.1)$$

Each vector of one set is orthogonal to two vectors of the other set, and produces a dot product of unity with the remaining vector of that other set. These relations can only be satisfied if the sets $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are each comprised of linearly independent vectors (non-coplanar, nor parallel). This implies that their respective triple products $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ and $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$ are non-zero, meaning that the parallelepiped formed by their arrow heads and a common origin has a finite volume.

To find the vectors of one set $\mathbf{a}, \mathbf{b}, \mathbf{c}$ in terms of its reciprocal set $\mathbf{A}, \mathbf{B}, \mathbf{C}$, we have the inter-relationship:

$$\mathbf{a} = \frac{\mathbf{B} \times \mathbf{C}}{\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})}, \quad \mathbf{b} = \frac{\mathbf{C} \times \mathbf{A}}{\mathbf{B} \cdot (\mathbf{C} \times \mathbf{A})}, \quad \mathbf{c} = \frac{\mathbf{A} \times \mathbf{B}}{\mathbf{C} \cdot (\mathbf{A} \times \mathbf{B})}, \quad (2.2.2)$$

where $\mathbf{A} \cdot \mathbf{B} \times \mathbf{C} = \mathbf{B} \cdot \mathbf{C} \times \mathbf{A} = \mathbf{C} \cdot \mathbf{A} \times \mathbf{B} \neq 0$. Actually, (2.2.1) are necessary and sufficient conditions for $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and $\mathbf{A}, \mathbf{B}, \mathbf{C}$ to be reciprocal sets. Equation (2.2.2) also holds with the pairs \mathbf{a}, \mathbf{A} ; \mathbf{b}, \mathbf{B} ; and \mathbf{c}, \mathbf{C} interchanged. That (2.2.1) and (2.2.2) are equivalent can be shown as follows. From the properties $\mathbf{a} \cdot \mathbf{B} = \mathbf{a} \cdot \mathbf{C} = 0$ one concludes that \mathbf{a} must be perpendicular to \mathbf{B} and \mathbf{C} . For \mathbf{B} and \mathbf{C} being non collinear vectors, this means that \mathbf{a} is along $\pm(\mathbf{B} \times \mathbf{C})$, or $\mathbf{a} = K(\mathbf{B} \times \mathbf{C})$, where K is a constant. After this expression has been dot-multiplied by the vector \mathbf{A} , one finds that the constant K equals $(\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}))^{-1}$ because $\mathbf{a} \cdot \mathbf{A} = 1$. This gives the first expression in (2.2.2). The formulae for \mathbf{b} and \mathbf{c} can be proved similarly or obtained by cyclic permutation.

The fact that each set $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and $\mathbf{A}, \mathbf{B}, \mathbf{C}$ consists of vectors forming a trihedron with non-zero volume suggests that they can be utilized to express any 3-D vector as a linear combination of them. In other words, the sets $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and $\mathbf{A}, \mathbf{B}, \mathbf{C}$ qualify as legitimate sets of basis vectors in 3-D. Indeed, any vector \mathbf{W} can be written as

$$\mathbf{W} = (\mathbf{W} \cdot \mathbf{a})\mathbf{a} + (\mathbf{W} \cdot \mathbf{b})\mathbf{b} + (\mathbf{W} \cdot \mathbf{c})\mathbf{c} \quad (2.2.3a)$$

or

$$\mathbf{W} = (\mathbf{W} \cdot \mathbf{A})\mathbf{A} + (\mathbf{W} \cdot \mathbf{B})\mathbf{B} + (\mathbf{W} \cdot \mathbf{C})\mathbf{C} \quad (2.2.3b)$$

Application of the so-called “*bac-cab*” rule to the expression $\mathbf{x} \times (\mathbf{C} \times \mathbf{W}) \equiv (\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{W})$, i.e., with $\mathbf{A} \times \mathbf{B}$ considered as one vector \mathbf{x} , leads to

$$\begin{aligned}(\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{W}) &= \mathbf{C}(\mathbf{A} \times \mathbf{B} \cdot \mathbf{W}) - \mathbf{W}(\mathbf{A} \times \mathbf{B} \cdot \mathbf{C}) \\ &= \mathbf{C}(\mathbf{A} \cdot \mathbf{B} \times \mathbf{W}) - \mathbf{W}(\mathbf{A} \cdot \mathbf{B} \times \mathbf{C})\end{aligned}\quad (2.2.4)$$

The same rule performed on $(\mathbf{C} \times \mathbf{W}) \times (\mathbf{A} \times \mathbf{B}) \equiv \mathbf{y} \times (\mathbf{A} \times \mathbf{B})$ gives

$$\begin{aligned}(\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{W}) &= -(\mathbf{C} \times \mathbf{W}) \times (\mathbf{A} \times \mathbf{B}) \\ &= -\mathbf{A}(\mathbf{C} \times \mathbf{W} \cdot \mathbf{B}) + \mathbf{B}(\mathbf{C} \times \mathbf{W} \cdot \mathbf{A}) \\ &= -\mathbf{A}(\mathbf{B} \cdot \mathbf{C} \times \mathbf{W}) + \mathbf{B}(\mathbf{A} \cdot \mathbf{C} \times \mathbf{W})\end{aligned}\quad (2.2.5)$$

Equating the right-hand sides of both equations, and solving for \mathbf{W} in the second term of (2.2.4), we obtain

$$\mathbf{W} = \frac{\mathbf{A}(\mathbf{B} \cdot \mathbf{C} \times \mathbf{W})}{\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})} - \frac{\mathbf{B}(\mathbf{A} \cdot \mathbf{C} \times \mathbf{W})}{\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})} + \frac{\mathbf{C}(\mathbf{A} \cdot \mathbf{B} \times \mathbf{W})}{\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})} \quad (2.2.6)$$

or

$$\mathbf{W} = \mathbf{A}\mathbf{W} \cdot \frac{\mathbf{B} \times \mathbf{C}}{\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}} + \mathbf{B}\mathbf{W} \cdot \frac{\mathbf{C} \times \mathbf{A}}{\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}} + \mathbf{C}\mathbf{W} \cdot \frac{\mathbf{A} \times \mathbf{B}}{\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}} \quad (2.2.7)$$

With (2.2.2), this becomes

$$\mathbf{W} = \mathbf{A}(\mathbf{W} \cdot \mathbf{a}) + \mathbf{B}(\mathbf{W} \cdot \mathbf{b}) + \mathbf{C}(\mathbf{W} \cdot \mathbf{c}),$$

proving relation (2.2.3a). Thus, the expansion coefficients of a vector \mathbf{W} expressed along the basis set $\mathbf{A}, \mathbf{B}, \mathbf{C}$ can be found from the dot product of \mathbf{W} with the “complementary” vector of the set $\mathbf{a}, \mathbf{b}, \mathbf{c}$ which is reciprocal to the original basis set. As can be seen from (2.2.3b), we could also expand along the set $\mathbf{a}, \mathbf{b}, \mathbf{c}$. The expressions in (2.2.3) are of crucial importance for our upcoming development of non-orthogonal coordinate systems.

Specializing to the Cartesian coordinate system, the set of unit basis vectors \hat{i}, \hat{j} , and \hat{k} , is reciprocal to itself. (Unit vectors are denoted by “ $\hat{}$ ”.) This implies that in such a system, a vector \mathbf{W} can be written as

$$\mathbf{W} = (\mathbf{W} \cdot \hat{i})\hat{i} + (\mathbf{W} \cdot \hat{j})\hat{j} + (\mathbf{W} \cdot \hat{k})\hat{k} \quad (2.2.8)$$

or equivalently,

$$\mathbf{W} = \mathbf{W}(1)\hat{i} + \mathbf{W}(2)\hat{j} + \mathbf{W}(3)\hat{k}. \quad (2.2.9)$$

We used a bracketed notation for the indices rather than subscripts or superscripts. The notation will be adjusted after we have established the difference between subscripts and superscripts. Equation (2.2.9) is the “usual” expansion of a vector in its components in the Cartesian coordinate system.

2.3 Curvilinear Coordinates

This section on curvilinear coordinates contains the fundamental ideas on which the development of flux coordinates is based. At the end of this section we arrive at the important conclusion that a vector can be conveniently expanded along two reciprocal sets of basis vectors having a most natural relationship to the chosen coordinate system. This leads to the concepts of covariant and contravariant components.

We first establish the curvilinear transformation and define coordinate curves and coordinate surfaces. Then, a first set of basis vectors, namely the tangent-basis vectors, is introduced. The second set of basis vectors will be comprised of the gradients to the coordinate surfaces.

We assume a knowledge of differential geometry, as described by Stoker (1969), Goetz (1970), Guggenheim (1977), Spivak (1970), Buck (1965), Mathews and Walker (1970), or Spiegel (1959), among others.

2.3.1 Transformation to Curvilinear Coordinates

Consider a transformation $\mathbf{R}(u^1, u^2, u^3)$ by means of which any point, determined by the position vector \mathbf{R} of 3-D space, is expressed as a function of three parameters u^1 , u^2 and u^3 . When we expand \mathbf{R} in its components x , y , z in a Cartesian coordinate system, we can rewrite this transformation as

$$\begin{aligned} x &= x(u^1, u^2, u^3) \\ \mathbf{R}(u^1, u^2, u^3): \quad y &= y(u^1, u^2, u^3) \\ z &= z(u^1, u^2, u^3) . \end{aligned} \tag{2.3.1}$$

If the transformation is one-to-one, it can be inverted:

$$\begin{aligned} u^1 &= u^1(x, y, z) \\ u^2 &= u^2(x, y, z) \\ u^3 &= u^3(x, y, z) . \end{aligned} \tag{2.3.2}$$

The requirements for a transformation to be invertible are that the functions x , y , z of (2.3.1) must have continuous partial derivatives with respect to u^1 , u^2 and u^3 respectively, and that the determinant made up by the nine partial derivatives $\partial x / \partial u^1, \partial x / \partial u^2, \partial x / \partial u^3, \partial y / \partial u^1, \dots$ is not zero (at least in the domain under consideration). Below, we shall call this determinant the Jacobian of the transformation. (For a detailed discussion, see Buck (1965).) Then the correspondence between (x, y, z) and (u^1, u^2, u^3) is unique. In some practical cases the single-valuedness does not hold, e.g., for periodic functions on "closed" surfaces such as spheres, cylinders and tori; this problem can usually be circumvented.

Equation (2.3.2) shows that a certain point (determined by) \mathbf{R} with Cartesian coordinates (x, y, z) can also be described uniquely by the independent parameters u^1 , u^2 and u^3 , so they are *coordinates* as well. (The (u^1, u^2, u^3) coordinates will be called *curvilinear* as a generalization of "rectilinear".) The reason for using a *superscript notation* (which could at times be confused with powers) will become clear below. The position of the indices is important.

In the curvilinear system there are three naturally occurring families of *coordinate surfaces*, obtained when one coordinate u^i is held fixed while the other two are varied continuously. The equations for these surfaces are:

$$\begin{aligned} u^1 &= c^1 & (u^2, u^3 \text{ variable}) \\ u^2 &= c^2 & (u^1, u^3 \text{ variable}) \\ u^3 &= c^3 & (u^1, u^2 \text{ variable}) . \end{aligned}$$

c^1 , c^2 and c^3 are arbitrary constants. It will be assumed that a unique surface of each family goes through any given point determined by the position vector \mathbf{R} in keeping with the hypothesis that the endpoint of \mathbf{R} is uniquely determined by its coordinates u^1 , u^2 , and u^3 .

Analogously, three families of *coordinate curves* are produced when one coordinate u^i is allowed to vary while the other two, u^j and u^k , are held fixed. The equations are:

$$\begin{aligned} u^2 &= c^2 , & u^3 &= c^3 & (u^1 \text{ variable}) \\ u^3 &= c^3 , & u^1 &= c^1 & (u^2 \text{ variable}) \\ u^1 &= c^1 , & u^2 &= c^2 & (u^3 \text{ variable}) . \end{aligned}$$

The direction in which a variable point of the curve moves as u^i is increased is taken to be the positive direction along that coordinate curve. The general coordinate system is summarized in Fig. 2.1.

If the coordinate curves intersect at right angles, then the curvilinear system is called *orthogonal*.

Examples of simple curvilinear coordinate systems are the familiar cylindrical and spherical systems. These are orthogonal.

2.3.2 Tangent-Basis Vectors

Consider the system of generalized coordinates u^1 , u^2 and u^3 as shown in Fig. 2.2. We define a *basis* at a point P , determined by the position vector \mathbf{R} , as any set of vectors e_1 , e_2 and e_3 of fixed length pointing in the positive direction of the coordinate curves. In other words, the *basis vectors* e_1 , e_2 and e_3 are tangent vectors of the coordinate curves at the point P .

The curvilinear coordinates have been defined by the transformation $\mathbf{R}(u^1, u^2, u^3)$. Simple differential geometry considerations show that $\partial \mathbf{R} / \partial u^i$ is a

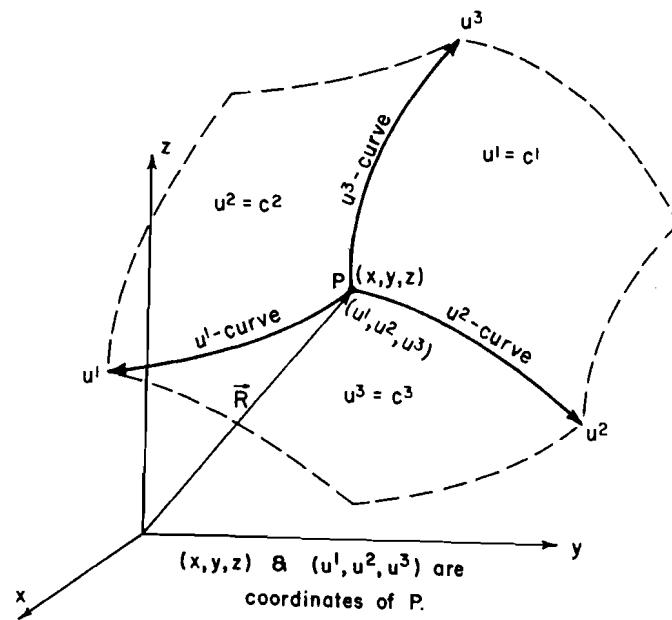


Fig. 2.1. General curvilinear coordinate system with coordinates (u^1, u^2, u^3) . The coordinate curves and coordinate surfaces are represented with respect to a Cartesian system with coordinates (x, y, z)

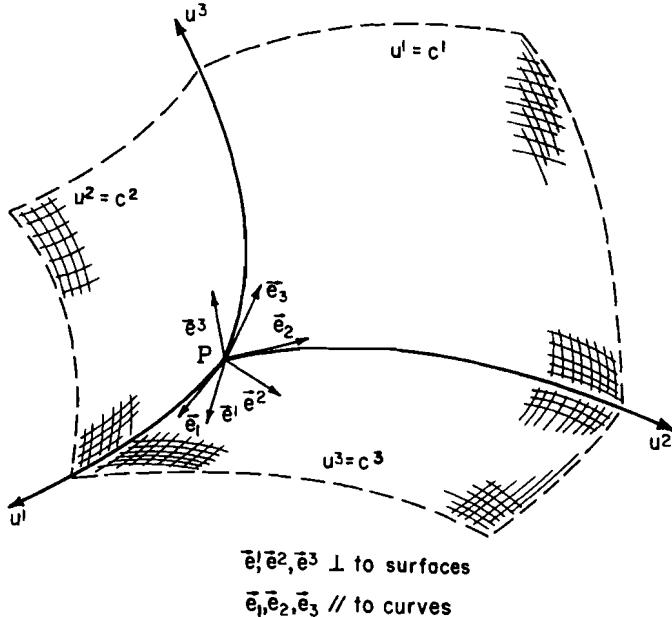


Fig. 2.2. Basis vectors e^i and e_j at point P . The contravariant basis vectors $e^i = \nabla u^i$ are perpendicular to the constant coordinate surfaces; the covariant basis vectors $e_j = \partial \mathbf{R} / \partial u^j$ are parallel to the coordinate curves

tangent vector of the u^i -coordinate curve. (A tangent vector to a one parameter curve $\mathbf{R} = \mathbf{R}(u)$ is $d\mathbf{R}/du$). It seems logical to choose these vectors as the *tangent-basis vectors*:

$$e_1 = \frac{\partial \mathbf{R}}{\partial u^1}; \quad e_2 = \frac{\partial \mathbf{R}}{\partial u^2}; \quad e_3 = \frac{\partial \mathbf{R}}{\partial u^3}. \quad (2.3.3)$$

The definition in (2.3.3) is consistent with the requirement that a set of basis vectors must be linearly independent; the triple product $e_1 \cdot e_2 \times e_3$ is non-zero.

This basis made up by the basis vectors e_1, e_2 and e_3 is said to be *local*, since, in general, it varies from point to point. (By "fixed length" in the definition of a basis, we mean that once the length of a basis vector at a point P_1 has been chosen, we always keep that length at P_1 ; at a point P_2 , the basis vector might have a different length.) Also, in general, these basis vectors are neither orthogonal nor of unit length, nor even dimensionless, since e.g., $\partial \mathbf{R} / \partial \theta$ has units of length. Only in coordinate systems with straight lines as coordinate "curves", i.e., with rectangular and oblique coordinate lines, do the basis vectors not vary from point to point.

If the variable u^i is taken to be the length along the coordinate curve, then the vector e_i is a *unit tangent vector*. In the more general case where u^i is just a parameter, one could define a unit tangent vector according to the obvious prescription:

$$\hat{e}_i \equiv \frac{e_i}{|e_i|} = \frac{\partial \mathbf{R}}{\partial u^i} \Bigg/ \left| \frac{\partial \mathbf{R}}{\partial u^i} \right| = \frac{e_i}{h_i} \quad (\text{no summation}). \quad (2.3.4)$$

The magnitude $|e_i| = |\partial \mathbf{R} / \partial u^i|$ is conventionally represented by the symbol h_i and is called a *scale factor* (or a metric coefficient).¹ Thus,

$$h_i = \left| \frac{\partial \mathbf{R}}{\partial u^i} \right| = |e_i|. \quad (2.3.5)$$

Although the h_i 's are usually mostly used for orthogonal systems, the above definition is completely general and valid for any coordinate system.

Note that the tangent-basis vector e_i has a *subscript* (i.e., the index is underneath). Although the coordinates u^i have superscripts, the fact that e_i has a subscript in (2.3.3) is suggestive because the index in

$$\frac{\partial \mathbf{R}}{\partial u^i}$$

is "underneath", i.e., below the "fraction-line".

¹ Sometimes these scale factors are called "Lamé factors" (especially in the Russian literature; see, e.g., Kovrzhnykh 1984). However, the name "Lamé factor" is also used in the context of elasticity theory, where it appears as the proportionality constant between the elasticity tensor and the strain tensor (Butkov 1968).

2.3.3 Reciprocal-Basis Vectors

The gradient, $\nabla\Phi$, of a function Φ is defined so that the differential $d\Phi$ is given by $d\Phi = \nabla\Phi \cdot d\mathbf{R}$. With $\Phi \equiv u^i$, this becomes

$$du^i = \nabla u^i \cdot d\mathbf{R} . \quad (2.3.6)$$

If we consider the position vector \mathbf{R} as a function of u^i , i.e., $\mathbf{R} = \mathbf{R}(u^1, u^2, u^3)$, we find from the chain rule that

$$d\mathbf{R} = \sum_{j=1}^3 \frac{\partial \mathbf{R}}{\partial u^j} du^j = \frac{\partial \mathbf{R}}{\partial u^j} du^j = e_j du^j . \quad (2.3.7)$$

Here we have used the definition of e_i given in (2.3.3) and in the last two forms of (2.3.7) the *summation convention* has been applied. The summation convention says that if a letter used as an index appears once as a superscript and once as a subscript on the same side of an equation, then that side of the equation should be considered to be summed over all values of that index. The summation sign is then superfluous. A repeated (or dummy) index implies summation and may be replaced by another letter; e.g. $e_j du^j = e_1 du^1 + e_2 du^2 + e_3 du^3 = e_k du^k$.

Substitution of $d\mathbf{R}$ as given by (2.3.7) into (2.3.6) leads to

$$du^i = \nabla u^i \cdot e_j du^j , \quad (2.3.8)$$

which can only hold if and only if

$$\nabla u^i \cdot e_j = \delta_j^i , \quad (2.3.9)$$

where δ_j^i is the Kronecker delta (which is unity if the indices coincide and zero otherwise).

The relationship in (2.3.9) states that the sets of vectors ∇u^i and e_j form reciprocal sets of vectors as defined in (2.2.1). Therefore, we define the set ∇u^i as the *reciprocal-basis vectors*:

$$e^i \equiv \nabla u^i \quad (2.3.10)$$

in which case (2.3.9) reads:

$$e^i \cdot e_j = \delta_j^i . \quad (2.3.11)$$

As was pointed out with respect to the e_j , the e^i are normally not of unit length nor are they perpendicular. While the basis vectors e_j are tangent to the u^j coordinate curves, the reciprocal basis vectors e^i are *perpendicular to the coordinate surfaces* $u^i = c^i$ (see Fig. 2.2).

The two sets of basis vectors e_j and e^i have been recognized as reciprocal sets of vectors. Consequently they satisfy the relationships (2.2.2), which allow any vector of one set to be calculated if the three vectors of the other set are known:

$$\nabla u^i = \frac{\frac{\partial \mathbf{R}}{\partial u^j} \times \frac{\partial \mathbf{R}}{\partial u^k}}{\frac{\partial \mathbf{R}}{\partial u^i} \cdot \left(\frac{\partial \mathbf{R}}{\partial u^j} \times \frac{\partial \mathbf{R}}{\partial u^k} \right)} = \frac{e_j \times e_k}{e_i \cdot (e_j \times e_k)} = e^i \quad (2.3.12)$$

(and similarly for ∇u^j and ∇u^k), and

$$\frac{\partial \mathbf{R}}{\partial u^i} = \frac{\nabla u^j \times \nabla u^k}{\nabla u^i \cdot (\nabla u^j \times \nabla u^k)} = \frac{e^j \times e^k}{e^i \cdot (e^j \times e^k)} = e_i \quad (2.3.13)$$

(and analogously for $\partial \mathbf{R}/\partial u^j$ and $\partial \mathbf{R}/\partial u^k$). Here i, j and k must be chosen such that (i, j, k) forms a cyclic permutation of $(1, 2, 3)$. This sentence will henceforth be abbreviated as “ i, j, k cyc 1, 2, 3”. The word “cyclic” is equivalent to the requirement that only “even” permutations of 1, 2, 3 are acceptable: 1, 2, 3 or 2, 3, 1, or 3, 1, 2. An ordering of numbers 1, 2, …, n is understood to be an even (odd) permutation of 1, 2, …, n if it results from the standard order by an even (odd) number of interchanges of pairs of numbers.

Alternatively, the reciprocal-basis vectors e^i are sometimes called *dual basis vectors* (Stoker 1969; Davis and Snider 1975). The reciprocal-basis vectors e^i are chosen to have a *superscript*, because ∇u^i does, consistently with e_i having a subscript.

When we use the term “basis vector” with no adjective, we mean the *tangent-basis vectors* e_i ; when the dual-basis vectors e^j are meant, we shall use the adjective “reciprocal”. Below, we will introduce other names for e_i and e^j , being related to the position of their indices i or j .

2.3.4 Covariant and Contravariant Components of a Vector

According to our earlier discussions, a vector can be written as a linear combination of the vectors from either of the (reciprocal) sets of basis vectors. From (2.2.3) a vector \mathbf{D} can be written as

$$\begin{aligned} \mathbf{D} &= \left(\mathbf{D} \cdot \frac{\partial \mathbf{R}}{\partial u^1} \right) \nabla u^1 + \left(\mathbf{D} \cdot \frac{\partial \mathbf{R}}{\partial u^2} \right) \nabla u^2 + \left(\mathbf{D} \cdot \frac{\partial \mathbf{R}}{\partial u^3} \right) \nabla u^3 \\ &= (\mathbf{D} \cdot e_1) e^1 + (\mathbf{D} \cdot e_2) e^2 + (\mathbf{D} \cdot e_3) e^3 \end{aligned} \quad (2.3.14a)$$

or

$$\mathbf{D} \equiv D_1 e^1 + D_2 e^2 + D_3 e^3 , \quad (2.3.15a)$$

which is an expansion along the set of reciprocal-basis vectors. Analogously, \mathbf{D} can be written as a linear combination of the tangent-basis vectors:

$$\begin{aligned} \mathbf{D} &= (\mathbf{D} \cdot \nabla u^1) \frac{\partial \mathbf{R}}{\partial u^1} + (\mathbf{D} \cdot \nabla u^2) \frac{\partial \mathbf{R}}{\partial u^2} + (\mathbf{D} \cdot \nabla u^3) \frac{\partial \mathbf{R}}{\partial u^3} \\ &= (\mathbf{D} \cdot e^1) e_1 + (\mathbf{D} \cdot e^2) e_2 + (\mathbf{D} \cdot e^3) e_3 \end{aligned} \quad (2.3.14b)$$

or,

$$\mathbf{D} \equiv D^1 \mathbf{e}_1 + D^2 \mathbf{e}_2 + D^3 \mathbf{e}_3 . \quad (2.3.15b)$$

In summary, we have

$$\mathbf{D} = D_i \mathbf{e}^i \quad \text{with} \quad D_i \equiv \mathbf{D} \cdot \mathbf{e}_i \quad (2.3.16a)$$

$$\mathbf{D} = D^i \mathbf{e}_i \quad \text{with} \quad D^i \equiv \mathbf{D} \cdot \mathbf{e}^i . \quad (2.3.16b)$$

The summation convention is implicit here.

One can look upon this expansion of (2.3.14) as follows: \mathbf{R} is the position vector pointing from the origin of a Cartesian coordinate system to the initial point of the vector we wish to describe, in this case the vector \mathbf{D} . The endpoint of \mathbf{R} is then the place where the two kinds of bases \mathbf{e}_i and \mathbf{e}^i are attached, and these allow us to find the components of \mathbf{D} by means of the dot products (2.3.16) (see Figs. 2.1 and 2.2).

The sub- or superscript notation replaces the bracket notation used in (2.2.9). There is more than one kind of component of a vector, depending on the choice of basis vectors.

The coefficients D_i appearing in (2.3.16a) are called the *covariant components* of the vector \mathbf{D} while the D^j in (2.3.16b) are called *contravariant components*. Equation (2.3.16a) shows that our notation is consistent; to find \mathbf{D} , one must sum over the sub- and superscript i ; the covariant component D_i results from the dot product of \mathbf{D} with the basis vector \mathbf{e}_i (the index is underneath on both sides of the equation $D_i = \mathbf{D} \cdot \mathbf{e}_i$).

From now on we adopt the following convention to name the components of a vector:

$$\text{covariant} \equiv \text{subscripts} \equiv ()_i$$

$$\text{contravariant} \equiv \text{superscripts} \equiv ()^j . \quad (2.3.17)$$

Thus an index *underneath* implies a *covariant* object, an index *above* denotes *contravariant* object. The reason for the words co- and contravariant has to do with the transformation properties of the components of a vector. This is discussed in Chap. 13. A simple mnemonic aid to distinguish these terms is as follows. Since there is an " R " in *superscript*, we call D^i a *contravariant component*. There is no " R " in *subscript*, hence the name *covariant*.

The functions $u^i(x, y, z)$ can have physical dimensions (units), as can the basis vectors. If the u^i are dimensionless, the \mathbf{e}_i and \mathbf{e}^i have dimensions of L and L^{-1} , respectively. Consequently, the *components of a vector do not necessarily have the same dimensions as the vector itself*.

For the special case of a Cartesian coordinate system it is clear from the definition formulae, (2.3.3), that the $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ basis vectors are the tangent-basis vectors. Moreover, they have a unit length since the coordinates are the "arc-lengths" along the coordinate axes. $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ thus form an orthonormal basis. On the other hand, since the relationship in (2.3.11) is exactly that obeyed by the $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ basis vectors, it follows that they are reciprocal basis vectors as well. (This is in agreement with our earlier observation that the set $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ is reciprocal to itself; see (2.2.8).) As a consequence, there is no distinction between co- and contravariant components of a vector in a Cartesian system: $\mathbf{W} \cdot \hat{\mathbf{e}}_1 = \mathbf{W} \cdot \hat{\mathbf{i}} = W(1)$, nor is there a need to make a distinction between sub- or superscripts in the component notation. Without any ambiguity, we can now rewrite (2.2.9) as

$\hat{\mathbf{W}} = W_1 \hat{\mathbf{i}} + W_2 \hat{\mathbf{j}} + W_3 \hat{\mathbf{k}}$

or

$$\hat{\mathbf{W}} = W_x \hat{\mathbf{i}} + W_y \hat{\mathbf{j}} + W_z \hat{\mathbf{k}} .$$

If we deal with a *position vector* \mathbf{R} , where $R_x = x$, $R_y = y$ and $R_z = z$, we obtain $\mathbf{R} = x \hat{\mathbf{i}} + y \hat{\mathbf{j}} + z \hat{\mathbf{k}}$.

2.4 Covariant and Contravariant "Vectors"

A lot of confusion about the terminology co- or contravariant vectors and/or components arises from the fact that both co- and contravariant vectors, and co- and contravariant components do exist, albeit in different contexts. Moreover, the wording "co- or contravariant vector" is often used to mean "co- or contravariant components of a vector". Here, we shall assert that when dealing with only one (Euclidean) vector space such as 3-D space, there exist only *vectors* (see e.g., Borisenko and Tarapov 1968). There are *no* contravariant and covariant vectors *as such* in the three-dimensional space: a vector is just a vector. If one digresses into abstract algebra, one can define contravariant vectors in one vector space and covariant vectors in *another* vector space, which is the dual of the first. References to this development are given at the end of Chap. 3.

A vector will always be identified by its *contravariant components* or by its *covariant components*. It must be stressed that, in principle, both representations contain all the information needed and are equally valid.

It is often said that the vector $d\mathbf{R}$ (the infinitesimal vector connecting the two position vectors \mathbf{R} and $\mathbf{R} + d\mathbf{R}$) is a contravariant vector. More specifically, the (misleading) statement would be that the coordinate differentials du^i form a contravariant vector. This statement should be modified to say that the "natural" and easiest way of writing a vector $d\mathbf{R}$ in components is by means of its contravariant components. These components are obtained by a simple chain-rule application, as can be seen from (2.3.7). Comparison of (2.3.7) with (2.3.16b) shows that the du^i do indeed represent the contravariant components of $d\mathbf{R}$. Since $d\mathbf{R}$ is the vector between two position vectors \mathbf{R} and $\mathbf{R} + d\mathbf{R}$ with coordinates (u^1, u^2, u^3) and $(u^1 + du^1, u^2 + du^2, u^3 + du^3)$, respectively, it is obvious why we used *superscripts* to label the coordinates u^i : contravariant implies above, and the

du^i are contravariant.

However, with equal validity, one could have chosen to write dR using its covariant components according to (2.3.16a)

$$dR = (dR)_i e^i = (dR)_i \nabla u^i. \quad (2.4.1)$$

Here, $(dR)_i$ is given by

$$\begin{aligned} (dR)_i &= dR \cdot e_i = dR \cdot \frac{\partial R}{\partial u^i} = du^j \frac{\partial R}{\partial u^j} \cdot \frac{\partial R}{\partial u^i} \\ &= du^j e_j \cdot e_i = e_i \cdot e_j du^j \\ &= g_{ij} du^j, \end{aligned} \quad (2.4.2)$$

where we have replaced the dot product of the basis vectors e_i and e_j by the symbol g_{ij} . (At this point, the symbol g_{ij} is a shorthand notation for $e_i \cdot e_j$; below, we shall give it a real meaning and stress its importance.) Thus, dR is just a vector, which can be written in terms of two sorts of components.

The confusion we alluded to above follows from the fact that some branches of science *define* a vector only by its components, without introducing a symbol for the vector itself. As a consequence, one (misleadingly) calls the quantity du^i a contravariant vector. However, here we employ the more correct term of contravariant components.

Similarly, it is sometimes said that the gradient of a function is a covariant vector. Once again this is an ambiguous statement and follows from the “normal” way of “decomposing” the gradient by means of chain-rule differentiation:

$$\nabla \Phi(u^1, u^2, u^3) = \nabla \Phi = \frac{\partial \Phi}{\partial u^i} \nabla u^i = \frac{\partial \Phi}{\partial u^i} e^i. \quad (2.4.3)$$

Therefore, we have that,

$$(\nabla \Phi)_i = \frac{\partial \Phi}{\partial u^i}, \quad (2.4.4)$$

showing that $(\partial \Phi / \partial u^i)$ are the covariant components of the gradient vector $\nabla \Phi$.

$\nabla \Phi$ can be written in terms of its contravariant components as:

$$\nabla \Phi = (\nabla \Phi)^i e_i$$

with

$$\begin{aligned} (\nabla \Phi)^i &= \nabla \Phi \cdot e^i = \nabla \Phi \cdot \nabla u^i = \frac{\partial \Phi}{\partial u^j} \nabla u^j \cdot \nabla u^i \\ &= \frac{\partial \Phi}{\partial u^j} e^j \cdot e^i = e^i \cdot e^j \frac{\partial \Phi}{\partial u^j} \\ &= g^{ij} \frac{\partial \Phi}{\partial u^j}. \end{aligned} \quad (2.4.5)$$

This time, we have replaced $(e^i \cdot e^j)$ by the symbol g^{ij} .

Calling $\partial \Phi / \partial u^i$ “the covariant vector” makes little sense; all one can say is that $\partial \Phi / \partial u^i$ are the covariant components, in this case with respect to the commonly chosen sets of bases $\partial R / \partial u^i$ and ∇u^i . The safest way (in the sense that the symbol is valid with respect to any kind of basis vectors) to denote covariant and contravariant components is by $(\nabla \Phi)_i$ and $(\nabla \Phi)^i$.

We shall, nevertheless, use the terms co- and contravariant vectors for certain types of vectors: the tangent and reciprocal basis vectors. It is common practice to call the tangent-basis vectors e_i “the covariant-basis vectors” (because they are labeled by a subscript) (Boozer 1980, 1981, 1982, 1983; Bateman 1978; Wrede 1963). Similarly the reciprocal basis vectors e^i are called “the contravariant-basis vectors” (because of the superscript). One should be aware of the fact that this is only for convenience, since the basis vectors e_i and e^i themselves can actually be written with covariant and contravariant components as well. Indeed, applying (2.3.16) we obtain for e_i the following: $e_i = (e_i)^j e_j = e_i \cdot e^j e_j = \delta_i^j e_j = 1 e_j$, and $e_i = (e_i)_j e^j = e_i \cdot e_j e^j = g_{ij} e^j$; analogously, for e^i , we would obtain $e^i = 1 e^i$ and $e^i = g^{ij} e_j$. Here we have again used the symbols g_{ij} and g^{ij} that we introduced above. Now, the surprising thing is that the “trivial” components (being equal to 1) for the e_i are the contravariant components and for the e^i are the covariant components! Nevertheless, one calls e_i the covariant and e^i the contravariant-basis vectors. This is thus even worse than for du^i and dR , or $\partial \Phi / \partial u^i$ and $\nabla \Phi$ discussed above; we are now advocating a “reverse logic”!² Since basis vectors are a privileged kind of vectors, one concentrates on the position of the indices and provides the basis vectors themselves with the names co- and contravariant.³ This is why we have been deliberately ambiguous below (2.3.17). We consider (2.3.17) as valid for everything that is labeled with subscripts and superscripts. Actually, there is a more profound reason for this policy. From the transformation laws discussed in Chap. 13, one can conclude that the description of e_i as the “covariant” basis vectors seems reasonable in the sense that they transform in the same way as the tangent-basis vectors, i.e., similarly to themselves. There it is explained that objects that transform similarly to the e_i are called covariant.

² For the same reason, our proposed nomenclature is not in agreement with the abstract-algebraic development of vectors in dual spaces. The tangent-basis vectors e_i are members of the “vector space of contravariant vectors” such that they are contravariant vectors viewed from that standpoint. On the other hand, the dual vector space (which is different from the first vector space), with the covariant vectors as elements, does contain the dual- or reciprocal-basis vectors. Hence, a reciprocal-basis vector is a covariant vector. However, when only one (Euclidean) vector space is involved, the “abstract algebraic” distinction between co- and contravariant vectors as such disappears. Only the distinction between co- and contravariant components of a vector can now be made. The fact that the basis vectors e_i and e^i are then merely vectors grants us the liberty to use the words co- and contravariant basis vectors according to the position of their indices.

³ There appears to be a typographical error in Bateman (1978), page 126, regarding this commonly used nomenclature in plasma physics.

2.5 Vector Relationships in Curvilinear Coordinates

2.5.1 The Metric Coefficients g_{ij} and g^{ij}

We now formally introduce the important metric coefficients g_{ij} and g^{ij} , which form the backbone of general curvilinear coordinate systems. They determine the differential arc length along a curve, they allow us to switch back and forth between the covariant and contravariant components of a vector, and they will provide us with a means to calculate the dot and cross products of two vectors.

The metric coefficients g_{ij} are defined as the dot product of the tangent-basis vectors:

$$g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = \frac{\partial \mathbf{R}}{\partial u^i} \cdot \frac{\partial \mathbf{R}}{\partial u^j}. \quad (2.5.1)$$

Similarly, the reciprocal or conjugate metric coefficients g^{ij} (also called simply metric coefficients) result from the dot product of the reciprocal-basis vectors:

$$g^{ij} = \mathbf{e}^i \cdot \mathbf{e}^j = \nabla u^i \cdot \nabla u^j. \quad (2.5.2)$$

From their definition, it follows that these coefficients are *symmetric*: $g_{ij} = g_{ji}$ and $g^{ij} = g^{ji}$.

When we deal with an orthogonal but not necessarily orthonormal coordinate system (such as a spherical, a cylindrical or a toroidal one) we find that $g_{ij} = 0$ if $i \neq j$ and similarly for g^{ij} . Thus *off-diagonal metric coefficients which are all zero imply an orthogonal coordinate system*.

Recall that the so called scale factors h_i were defined as $h_i = |\mathbf{e}_i|$. Therefore, (2.5.1) becomes

$$g_{ij} = h_i h_j \hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = h_i h_j \cos \varphi(\mathbf{e}_i, \mathbf{e}_j), \quad (2.5.3)$$

leading to

$$g_{ii} = h_i^2 \quad \text{or} \quad h_i = \sqrt{g_{ii}}. \quad (2.5.4)$$

This is the reason why the h_i are also often called metric coefficients.

The differential arc length along a curve is defined as:

$$dl = |\mathbf{dR}| = \sqrt{\mathbf{dR} \cdot \mathbf{dR}}. \quad (2.5.5)$$

The differential vector $d\mathbf{R}$ equals

$$d\mathbf{R} = \frac{\partial \mathbf{R}}{\partial u^1} du^1 + \frac{\partial \mathbf{R}}{\partial u^2} du^2 + \frac{\partial \mathbf{R}}{\partial u^3} du^3 = \frac{\partial \mathbf{R}}{\partial u^j} du^j = \mathbf{e}_j du^j. \quad (2.5.6)$$

so that the square of the arc length is given by

$$(dl)^2 = \sum_{i,j} du^i du^j \mathbf{e}_i \cdot \mathbf{e}_j. \quad (2.5.7)$$

Because of (2.5.1) this can be written as

$$(dl)^2 = g_{ij} du^i du^j \quad (2.5.8a)$$

in terms of the g_{ij} , or as

$$(dl)^2 = h_i h_j \cos \varphi(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) du^i du^j, \quad (2.5.8b)$$

expressed in terms of the h_i . These equations are often used to define the metric coefficients. Instead, we preferred to define the g_{ij} through (2.5.1), which seems to provide more “physical” insight.

Once the metric coefficients g_{ij} and g^{ij} are known, it turns out to be a trivial exercise to change covariant vector components into contravariant components, and vice versa. From the definitions in (2.3.16) we obtain:

$$\mathbf{D}_i = \mathbf{D} \cdot \mathbf{e}_i = D^j \mathbf{e}_j \cdot \mathbf{e}_i = D^j g_{ji}$$

or, utilizing the symmetry of g_{ij}

$$\mathbf{D}_i = g_{ij} D^j. \quad (2.5.9)$$

Similarly,

$$\mathbf{D}^k = \mathbf{D} \cdot \mathbf{e}^k = D_i \mathbf{e}^i \cdot \mathbf{e}^k = D_i g^{ik}$$

or

$$\mathbf{D}^k = g^{ki} \mathbf{D}_i. \quad (2.5.10)$$

Recall that (2.5.9) and (2.5.10) must be summed over j and i respectively.

Analogous expressions hold for the covariant and contravariant basis vectors as well. Following (2.3.14a) but with $\mathbf{D} \equiv \mathbf{e}_i$ leads to

$$\mathbf{e}_i = (\mathbf{e}_i \cdot \mathbf{e}_j) \mathbf{e}^j$$

or

$$\mathbf{e}_i = g_{ij} \mathbf{e}^j. \quad (2.5.11)$$

By the same token, (2.3.14b) gives then, for $\mathbf{D} \equiv \mathbf{e}^i$,

$$\mathbf{e}^i = (\mathbf{e}^i \cdot \mathbf{e}^j) \mathbf{e}_j$$

or

$$\mathbf{e}^i = g^{ij} \mathbf{e}_j. \quad (2.5.12)$$

Whenever covariant quantities are to be converted to contravariant components (and vice-versa), one multiplies them (allowing for the implicit summation) by the metric coefficients having the other index position.

A final interesting relationship concerns the metric coefficients g_{ij} , g^{ij} and the Kronecker delta δ_i^k . Consider (2.5.11).

$$e_i = g_{ij} e^j .$$

If we dot-multiply this equation with the reciprocal-basis vector e^k , we obtain

$$e_i \cdot e^k = g_{ij} e^j \cdot e^k$$

or

$$\delta_i^k = g_{ij} g^{jk} . \quad (2.5.13)$$

Upon interchanging i and k , (2.5.12) leads to an equivalent expression (since g_{ij} and g^{ij} are symmetric, this observation is trivial):

$$\delta_i^k = g^{kj} g_{ji} . \quad (2.5.14)$$

When we form a matrix with elements g_{ij} and denote it by $[g_{ij}]$, and similarly for g^{ik} , and recall the definition of the product of two matrices, then we can interpret (2.5.13) and (2.5.14) by saying that the product of the matrices $[g_{ij}]$ and $[g^{ik}]$ equals the identity matrix. (If $[C] = [A][B]$ then $C_{ij} = \sum_m A_{im} B_{mj}$.) Consequently, the coefficients g^{ij} (or g_{ij}) can be computed from the g_{ij} (or g^{ij}) by calculation of the inverse matrix. Thus, $[g^{ij}] = [g_{ij}]^{-1}$ and $[g_{ij}] = [g^{ij}]^{-1}$, or $g^{ij} = G^{ij}/g$ and $g_{ij} = G_{ij}/g^{-1} = G_{ij}g$. Here G_{ij} (or G^{ij}) represents the cofactor of the matrix element g^{ij} (or g_{ij}) while g is the determinant of the matrix $[g_{ij}]$. Thus

$$g = \det[g_{ij}] \quad \text{and} \quad g^{-1} = \det[g^{ij}] . \quad (2.5.15)$$

The expressions in (2.5.11–14), which imply summations, can be recalled by noting that the summation over an index, appearing above and below on the same side of the equation, can be thought of as a “saturation” of that index. As a consequence, that index disappears from the result. The indices in the result are the ones which we did not sum over, and are simply carried over to the other side of the equality sign, at the same position (above or below). This rule of thumb suggests that we can write (2.5.13) and (2.5.14) as follows:

$$g_i^k = g_{ij} g^{jk} = g^{kj} g_{ji} . \quad (2.5.16)$$

Indeed, as we shall discuss later in detail in Chap. 3. on tensor-related issues, the symbol g_i^k has the same values, 0 or 1, as the Kronecker delta δ_i^k :

$$g_i^k = \delta_i^k . \quad (2.5.17)$$

In the chapter on tensors we shall point out that in principle the only legitimate Kronecker delta must be written as δ_i^k and that (at least as long as we deal with covariant and contravariant quantities, i.e., with sub- and superscripts) one should be cautious in employing symbols like δ_{ij} or δ^{ij} .

2.5.2 The Jacobian

a) The Jacobian of the Curvilinear Coordinate System

Consider the curvilinear coordinate transformation introduced in (2.3.1) and (2.3.2). The Jacobian of the transformation (2.3.1) is by definition the determinant of the matrix produced by the nine partial derivatives $\partial x/\partial u^i$, $\partial y/\partial u^i$, $\partial z/\partial u^i$; $i = 1, 2, 3$. Its proper name is “the Jacobian of (x, y, z) with respect to the u^i ”. Thus we have

$$J \equiv J(u^1, u^2, u^3) \equiv J\left(\frac{x, y, z}{u^1, u^2, u^3}\right) \equiv \frac{\partial(x, y, z)}{\partial(u^1, u^2, u^3)} . \quad (2.5.18)$$

$$\equiv \begin{vmatrix} \frac{\partial x}{\partial u^1} & \frac{\partial x}{\partial u^2} & \frac{\partial x}{\partial u^3} \\ \frac{\partial y}{\partial u^1} & \frac{\partial y}{\partial u^2} & \frac{\partial y}{\partial u^3} \\ \frac{\partial z}{\partial u^1} & \frac{\partial z}{\partial u^2} & \frac{\partial z}{\partial u^3} \end{vmatrix} \equiv \begin{vmatrix} \frac{\partial x}{\partial u^1} & \frac{\partial y}{\partial u^1} & \frac{\partial z}{\partial u^1} \\ \frac{\partial x}{\partial u^2} & \frac{\partial y}{\partial u^2} & \frac{\partial z}{\partial u^2} \\ \frac{\partial x}{\partial u^3} & \frac{\partial y}{\partial u^3} & \frac{\partial z}{\partial u^3} \end{vmatrix} .$$

The last equality is a consequence of equal determinants of a matrix and its transpose. The notation $J(u^1, u^2, u^3)$ stresses that J is a function of the u^i .

If we recall that the triple product $\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}$ can be written as a determinant of the matrix with elements $A_x, A_y, A_z, B_x, B_y, B_z$, etc. in a Cartesian system and think of the transformation in (2.3.1) as $\mathbf{R}[x(u^1, u^2, u^3), y(u^1, u^2, u^3), z(u^1, u^2, u^3)]$, we can rewrite (2.5.18) in a more compact form:

$$J = \frac{\partial \mathbf{R}}{\partial u^1} \cdot \frac{\partial \mathbf{R}}{\partial u^2} \times \frac{\partial \mathbf{R}}{\partial u^3} . \quad (2.5.19)$$

All the even permutations of 1, 2, 3 are equally valid. The inverse transformation given by (2.3.2) has a Jacobian as well:

$$\mathcal{J} \equiv \mathcal{J}(x, y, z) \equiv \mathcal{J}\left(\frac{u^1, u^2, u^3}{x, y, z}\right) \equiv \frac{\partial(u^1, u^2, u^3)}{\partial(x, y, z)} . \quad (2.5.20)$$

$$= \begin{vmatrix} \frac{\partial u^1}{\partial x} & \frac{\partial u^1}{\partial y} & \frac{\partial u^1}{\partial z} \\ \frac{\partial u^2}{\partial x} & \frac{\partial u^2}{\partial y} & \frac{\partial u^2}{\partial z} \\ \frac{\partial u^3}{\partial x} & \frac{\partial u^3}{\partial y} & \frac{\partial u^3}{\partial z} \end{vmatrix} = \begin{vmatrix} \frac{\partial u^1}{\partial x} & \frac{\partial u^2}{\partial x} & \frac{\partial u^3}{\partial x} \\ \frac{\partial u^1}{\partial y} & \frac{\partial u^2}{\partial y} & \frac{\partial u^3}{\partial y} \\ \frac{\partial u^1}{\partial z} & \frac{\partial u^2}{\partial z} & \frac{\partial u^3}{\partial z} \end{vmatrix} .$$

Now $\partial u^i/\partial x$, $\partial u^i/\partial y$, $\partial u^i/\partial z$ are considered to be the Cartesian components of ∇u^i ,

and so we have equivalently:

$$\mathcal{J} = \nabla u^1 \cdot \nabla u^2 \times \nabla u^3 . \quad (2.5.21)$$

We recognize the vectors in (2.5.19) and (2.5.21) as the members of our bases: the tangent-basis vectors $e_i \equiv \partial \mathbf{R}/\partial u^i$, and the reciprocal-basis vectors $e^i \equiv \nabla u^i$. Hence, the Jacobians are identical to

$$J = e_1 \cdot e_2 \times e_3 ; \quad \mathcal{J} = e^1 \cdot e^2 \times e^3 . \quad (2.5.22a, b)$$

Since the e_i and e^i form reciprocal sets, we can prove the expected relationship between the two Jacobians \mathcal{J} and J , namely, that they are the inverses of each other:

$$\mathcal{J} = J^{-1} . \quad (2.5.23)$$

Starting from $\mathcal{J} = e^1 \cdot e^2 \times e^3$, and substituting the expression for the e^i as functions of the e_i , as given by (2.3.12), we obtain

$$\begin{aligned} \mathcal{J} &= \frac{(e_2 \times e_3) \cdot [(e_3 \times e_1) \times (e_1 \times e_2)]}{[e_1 \cdot e_2 \times e_3]^3} \\ &= \frac{(e_2 \times e_3)}{J^3} \cdot (e_1(e_2 \cdot e_3 \times e_1) - e_2(e_1 \cdot e_3 \times e_1)) \\ &= \frac{[e_1 \cdot (e_2 \times e_3)][e_2 \cdot e_3 \times e_1]}{J^3} = \frac{J^2}{J^3} = J^{-1} . \end{aligned}$$

(The distinction between superscripts and powers should be clear from the context.)

From now on, we shall always employ the symbol J for the Jacobian. The Jacobian of (2.5.21) or (2.5.22b) will be referred to as J^{-1} . It is customary to call J simply the Jacobian of the curvilinear coordinate system. Having said this, we can rewrite the “ $e_i - e^i$ back-and-forth” expressions of (2.3.12) and (2.3.13) more compactly:

$$e^i \equiv \nabla u^i = \frac{1}{J} \left(\frac{\partial \mathbf{R}}{\partial u^j} \times \frac{\partial \mathbf{R}}{\partial u^k} \right) = \frac{1}{J} (e_j \times e_k) \quad i, j, k \text{ cyc } 1, 2, 3 \quad (2.5.24)$$

and

$$e_i \equiv \frac{\partial \mathbf{R}}{\partial u^i} = J(\nabla u^j \times \nabla u^k) = J(e^j \times e^k) \quad i, j, k \text{ cyc } 1, 2, 3 \quad (2.5.25)$$

where

$$J = \frac{1}{\nabla u^1 \cdot \nabla u^2 \times \nabla u^3} = \frac{\partial \mathbf{R}}{\partial u^1} \cdot \frac{\partial \mathbf{R}}{\partial u^2} \times \frac{\partial \mathbf{R}}{\partial u^3} . \quad (2.5.26)$$

b) The Relationship Between J and g

In a previous section we defined the determinant g of the matrix of metric coefficients: $g = \det[g_{ij}]$. For $\mathbf{R} = \mathbf{R}(x, y, z)$, where (x, y, z) are Cartesian coordinates, the dot product $e_i \cdot e_j$ is readily computed:

$$g_{ij} = \frac{\partial \mathbf{R}}{\partial u^i} \cdot \frac{\partial \mathbf{R}}{\partial u^j} = \frac{\partial x}{\partial u^i} \frac{\partial x}{\partial u^j} + \frac{\partial y}{\partial u^i} \frac{\partial y}{\partial u^j} + \frac{\partial z}{\partial u^i} \frac{\partial z}{\partial u^j} . \quad (2.5.27)$$

It is a theorem of matrix algebra that the determinant of the product of two matrices $[A]$ and $[B]$ equals the product of the determinants of the matrices: $(\det[A])(\det[B]) = \det([A][B])$. Recall that if $[C] = [A][B]$, then

$$C_{ij} = \sum_{q=1}^3 A_{iq} B_{qj} .$$

If we now, for a moment, represent x by $R(1)$, y by $R(2)$, and z by $R(3)$, then (2.5.27) can be written as

$$g_{ij} = \sum_{q=1}^3 \frac{\partial R(q)}{\partial u^i} \frac{\partial R(q)}{\partial u^j} . \quad (2.5.28)$$

Thus the g_{ij} of (2.5.28) can be thought of as produced by the product of two matrices with elements $\partial R(q)/\partial u^i$ and $\partial R(q)/\partial u^j$,

$$g = \det[g_{ij}] = \left(\det \left[\frac{\partial R(q)}{\partial u^i} \right] \right) \left(\det \left[\frac{\partial R(q)}{\partial u^j} \right] \right)$$

$$= \begin{vmatrix} \frac{\partial x}{\partial u^1} & \frac{\partial x}{\partial u^2} & \frac{\partial x}{\partial u^3} \\ \frac{\partial y}{\partial u^1} & \frac{\partial y}{\partial u^2} & \frac{\partial y}{\partial u^3} \\ \frac{\partial z}{\partial u^1} & \frac{\partial z}{\partial u^2} & \frac{\partial z}{\partial u^3} \end{vmatrix} \cdot \begin{vmatrix} \frac{\partial x}{\partial u^1} & \frac{\partial y}{\partial u^1} & \frac{\partial z}{\partial u^1} \\ \frac{\partial x}{\partial u^2} & \frac{\partial y}{\partial u^2} & \frac{\partial z}{\partial u^2} \\ \frac{\partial x}{\partial u^3} & \frac{\partial y}{\partial u^3} & \frac{\partial z}{\partial u^3} \end{vmatrix} .$$

$$= \left(\frac{\partial \mathbf{R}}{\partial u^1} \cdot \frac{\partial \mathbf{R}}{\partial u^2} \times \frac{\partial \mathbf{R}}{\partial u^3} \right)^2 = (J)^2 . \quad (2.5.29)$$

Hence we arrive at the important result that the determinant of the “metric matrix”, g , is equal to the square of the Jacobian J :

$$g = (J)^2 \quad \text{or} \quad J = \sqrt{g} . \quad (2.5.30)$$

In an *orthogonal* coordinate system, (2.5.26) becomes

$$J = \sqrt{g} = h_1 \hat{e}_1 \cdot h_2 \hat{e}_2 \times h_3 \hat{e}_3 = h_1 h_2 h_3 \quad (2.5.31)$$

since $\hat{e}_1 \cdot \hat{e}_2 \times \hat{e}_3 = 1$. (Recall that \hat{e}_i is a unit vector.)

2.5.3 The Dot and Cross-Products in Curvilinear Coordinates

Now we turn to the computation of the dot and cross-products between two vectors. The existence of two convenient bases along which to expand a vector into its components has implications for the computation of the dot and cross products.

a) Dot Products

Two vectors \mathbf{A} and \mathbf{B} can be written in terms of their contravariant or covariant components:

$$\mathbf{A} = A^i \mathbf{e}_i = A_i \mathbf{e}^i \quad (2.5.32a)$$

$$\mathbf{B} = B^j \mathbf{e}_j = B_j \mathbf{e}^j . \quad (2.5.32b)$$

The dot product between two vectors can be “shifted” to their basis-vectors, giving rise to four possible combinations:

$$\mathbf{A} \cdot \mathbf{B} = A^i \mathbf{e}_i \cdot B_j \mathbf{e}^j = A^i B_j \mathbf{e}_i \cdot \mathbf{e}^j = A^i B_j \delta_i^j \quad (2.5.33a)$$

$$= A_i \mathbf{e}^i \cdot B^j \mathbf{e}_j = A_i B^j \mathbf{e}^i \cdot \mathbf{e}_j = A_i B^j \delta_i^j \quad (2.5.33b)$$

$$= A^i \mathbf{e}_i \cdot B^j \mathbf{e}_j = A^i B^j \mathbf{e}_i \cdot \mathbf{e}_j = A^i B^j g_{ij} \quad (2.5.33c)$$

$$= A_i \mathbf{e}^i \cdot B_j \mathbf{e}^j = A_i B_j \mathbf{e}^i \cdot \mathbf{e}^j = A_i B_j g^{ij} . \quad (2.5.33d)$$

In summary, the dot product of two vectors in general curvilinear coordinates is

$$\mathbf{A} \cdot \mathbf{B} = A^i B_i = A_i B^i = g_{ij} A^i B^j = g^{ij} A_i B_j . \quad (2.5.34)$$

We should note that the last two forms of (2.5.34) also follow from the first two (and vice versa) when we recall the $D_i \leftrightarrow D^j$ formulae previously derived in (2.5.9) and (2.5.10). In *orthogonal* coordinate systems $g_{ij} = g^{ij} = 0$ for $i \neq j$ and (2.5.34) reduces to

$$\mathbf{A} \cdot \mathbf{B} = A^i B_i = A_i B^i = (h_i)^2 A^i B^i = \frac{A_i B_i}{(h_i)^2} . \quad (2.5.35)$$

A summation over i is implied in all formulae in (2.5.35). If we restrict our attention further to an *orthonormal* coordinate system, where $h_i = 1$ and $A_i = A^i$, we retrieve the trivial expression $\sum A(i) B(i) \equiv \sum A_i B_i$.

The magnitude of a vector equals the square root of the dot product with itself:

$$\begin{aligned} A = |\mathbf{A}| &= \sqrt{\mathbf{A} \cdot \mathbf{A}} = \sqrt{A^i A_i} = \sqrt{A_i A^i} = \sqrt{g_{ij} A^i A^j} \\ &= \sqrt{g^{ij} A_i A_j} . \end{aligned} \quad (2.5.36)$$

We have essentially already applied the expressions derived above, when we determined the differential arc length along a curve in (2.5.8).

b) Cross Products

For the vector product, there are only two convenient combinations of basis vectors:

$$\mathbf{A} \times \mathbf{B} = A^i B^j \mathbf{e}_i \times \mathbf{e}_j = A_i B_j \mathbf{e}^i \times \mathbf{e}^j . \quad (2.5.37)$$

Here, as always, a summation over the indices i and j is understood (where we note that the terms involving $i = j$ give zero). The “ $\mathbf{e}_i - \mathbf{e}^j$ back-and-forth” formulae of (2.5.24) and (2.5.25) give

$$\mathbf{A} \times \mathbf{B} = J \sum'_{i,j} (\pm) A^i B^j \mathbf{e}^k = J^{-1} \sum'_{i,j} (\pm) A_i B_j \mathbf{e}_k . \quad (2.5.38)$$

The index k must be chosen such that i, j, k form a permutation of 1, 2, 3; an even permutation is given a “+” sign, an odd one is assigned a “−” sign. The prime on the summation sign indicates that the values $i = j$ must be omitted from the summation. This means that we can rewrite (2.5.38) as follows:

$$\begin{aligned} \mathbf{A} \times \mathbf{B} &= J \sum_k (A^i B^j - A^j B^i) \mathbf{e}^k \quad i, j, k \text{ cyc } 1, 2, 3 \\ &= \frac{1}{J} \sum_k (A_i B_j - A_j B_i) \mathbf{e}_k \quad i, j, k \text{ cyc } 1, 2, 3 . \end{aligned} \quad (2.5.39)$$

The change in summation index is made possible by the sentence “ i, j, k cyc 1, 2, 3.” This expression can be written more compactly with the help of the Levi-Civita symbols, ϵ^{ijk} and ϵ_{ijk} , which are defined as

$$\epsilon^{ijk} = \begin{cases} +1 & \text{if } i, j, k \text{ form an even permutation of } 1, 2, 3 \\ -1 & \text{if } i, j, k \text{ form an odd permutation of } 1, 2, 3 \\ 0 & \text{otherwise ,} \end{cases} \quad (2.5.40)$$

and similarly for ϵ_{ijk} .

This leads to:

$$(\mathbf{A} \times \mathbf{B})_k = \epsilon_{ijk} J A^i B^j = \epsilon_{ijk} \sqrt{g} A^i B^j \quad i, j, k \text{ cyc } 1, 2, 3 \quad (2.5.41)$$

for the first expression in (2.5.38) and

$$(\mathbf{A} \times \mathbf{B})^k = \frac{\epsilon^{ijk}}{J} A_i B_j = \frac{\epsilon^{ijk}}{\sqrt{g}} A_i B_j \quad i, j, k \text{ cyc } 1, 2, 3 \quad (2.5.42)$$

for the second expression. Note that in these expressions on the left- and right-hand side of the equality, components of the other kind appear, and there is an additional factor of $J = \sqrt{g}$ or $J^{-1} = (\sqrt{g})^{-1}$.

Expressions for the vector product in terms of components of the same sort are given by the expressions given in (2.5.9) and (2.5.10), i.e.,

$$(\mathbf{A} \times \mathbf{B})_i = g_{ik} (\mathbf{A} \times \mathbf{B})^k \quad \text{and} \quad (\mathbf{A} \times \mathbf{B})^i = g^{ik} (\mathbf{A} \times \mathbf{B})_k . \quad (2.5.43)$$

2.5.4 The Differential Elements $d\ell(i)$, $dS(i)$, d^3R

a) The Differential Arc Length $d\ell(i)$

The differential arc length along a coordinate curve u^i , which we denote by $d\ell(i)$ is, according to (2.5.5)

$$d\ell(\text{along } u^i) \equiv d\ell(i) = |d\mathbf{R}(i)| = \sqrt{d\mathbf{R}(i) \cdot d\mathbf{R}(i)} . \quad (2.5.44)$$

The general expression for the vector $d\mathbf{R}$ is given in (2.5.6). However, since the coordinate curve u^i is defined by the intersection of two surfaces $u^j = \text{constant}$ and $u^k = \text{constant}$, we must require that $du^j = du^k = 0$ for $j, k \neq i$. Therefore, we obtain

$$\begin{aligned} d\mathbf{R}(i) &= \frac{\partial \mathbf{R}}{\partial u^i} du^i = \mathbf{e}_i du^i \quad \text{no summation} \\ &= h_i du^i \hat{\mathbf{e}}_i = \sqrt{g_{ii}} du^i \hat{\mathbf{e}}_i \quad \text{no summation} . \end{aligned} \quad (2.5.45)$$

Consequently, (2.5.44) leads to

$$d\ell(i) = h_i du^i = \sqrt{g_{ii}} du^i , \quad \text{no summation} . \quad (2.5.46)$$

This is the usual definition of h_i . Note that the form $h_i du^i$ is also valid for non-orthogonal systems.

From the definition of g_{ij} in (2.5.1) and the $\mathbf{e}_i \leftrightarrow \mathbf{e}^i$ formulae in (2.5.25) we can derive an alternative expression for $d\ell(i)$:

$$d\ell(i) = J |\nabla u^i \times \nabla u^k| du^i \quad i, j, k \text{ cyc } 1, 2, 3 . \quad (2.5.47)$$

b) The Differential Area Element $dS(i)$

The differential element of area in the coordinate surface $u^i = \text{constant}$ (at a point P determined by the position vector \mathbf{R}) is easily computed if we recall that the magnitude of the vector product equals the area of the parallelogram made by the two vectors:

$$\begin{aligned} dS(\text{in } u^i = \text{constant}) &\equiv dS(i) = |d\mathbf{R}(j) \times d\mathbf{R}(k)| \\ &= \left| \frac{\partial \mathbf{R}}{\partial u^j} \times \frac{\partial \mathbf{R}}{\partial u^k} \right| du^j du^k , \quad \text{no summation over } j \text{ or } k . \end{aligned} \quad (2.5.48a)$$

Or

$$dS(i) = |\mathbf{e}_j \times \mathbf{e}_k| du^j du^k , \quad \text{no summation over } j \text{ or } k . \quad (2.5.48b)$$

Here again, i, j, k form a cyclic permutation of 1, 2, 3.

With the expression for $\mathbf{e}^i \equiv \nabla u^i$ as a function of \mathbf{e}_j and \mathbf{e}_k as given in (2.5.24) we can rewrite $dS(i)$ as

$$dS(i) = J |\nabla u^i| du^j du^k . \quad (2.5.49a)$$

This expression is useful when we need the differential area vector $d\mathbf{S}(i)$. Since we know that the gradient is perpendicular to the surface, we obtain

$$d\mathbf{S}(i) \equiv \frac{\nabla u^i}{|\nabla u^i|} dS = \pm J du^j du^k \nabla u^i , \quad \text{no summation} . \quad (2.5.49b)$$

The sign must be chosen according to the convention adopted for the normal to the surface.

We now rearrange the expression given in (2.5.48b). Making use of the vector identity $(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C})$, to obtain

$$\begin{aligned} dS(i) &= \sqrt{(\mathbf{e}_j \times \mathbf{e}_k) \cdot (\mathbf{e}_j \times \mathbf{e}_k)} du^j du^k \\ &= \sqrt{(\mathbf{e}_j \cdot \mathbf{e}_j)(\mathbf{e}_k \cdot \mathbf{e}_k) - (\mathbf{e}_j \cdot \mathbf{e}_k)(\mathbf{e}_j \cdot \mathbf{e}_k)} du^j du^k \\ &\quad \text{no summation over } j \text{ or } k . \end{aligned} \quad (2.5.50)$$

Subsequently we employ the definition of g_{ij} , and find

$$\begin{aligned} dS(i) &= \sqrt{g_{jj} g_{kk} - g_{jk}^2} du^j du^k , \\ &\quad \text{no summation over } j \text{ or } k ; \quad i, j, k \text{ cyc } 1, 2, 3 . \end{aligned} \quad (2.5.51)$$

The formulae for *orthogonal* systems are quite simple since the cross-metric coefficients g_{jk} are identically zero. Using (2.5.4), we have

$$\begin{aligned} dS(i) &= h_j h_k du^j du^k , \\ &\quad \text{no summation over } j \text{ or } k ; \quad i, j, k \text{ cyc } 1, 2, 3 . \end{aligned} \quad (2.5.52)$$

c) The Differential Volume Element d^3R

d^3R represents the differential-volume element at a point P determined by the position vector \mathbf{R} . It is “infinitesimal” in all three coordinates. The symbol dV is kept for a volume element that is infinitesimal in only one direction (e.g. in the spherical system d^3R would be $r^2 dr \sin \theta d\theta d\phi$, while dV is for instance $4\pi r^2 dr$). The volume element d^3R is the volume constructed by the differential vectors along each of the coordinate curves:

$$d^3R = d\mathbf{R}(1) \cdot d\mathbf{R}(2) \times d\mathbf{R}(3) = du^1 du^2 du^3 \mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3) \quad (2.5.53)$$

or

$$d^3R = J du^1 du^2 du^3 = \sqrt{g} du^1 du^2 du^3 . \quad (2.5.54)$$

This result is positive because by definition $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ form a right-handed system.

2.6 Vector Differentiation in Curvilinear Coordinates

2.6.1 The Covariant Derivative

a) Differentiation of Vector Components and Basis Vectors

Vector differentiation is a straightforward operation, as long as we keep in mind that, in curvilinear coordinates, not only the components of the vector but also the basis vectors (tangent and reciprocal) are position-dependent.

Consider a vector field $\mathbf{W} = \mathbf{W}(\mathbf{R}(u^1, u^2, u^3))$, i.e. the vector \mathbf{W} is a function of position as indicated by the position vector \mathbf{R} . \mathbf{W} can be expanded as usual:

$$\mathbf{W} = W^i \mathbf{e}_i \quad (2.6.1)$$

or

$$\mathbf{W} = W_i \mathbf{e}^i . \quad (2.6.2)$$

The W^i , W_i , \mathbf{e}_i and \mathbf{e}^i are all dependent upon (u^1, u^2, u^3) . The differential vector $d\mathbf{W}$ is given by

$$d\mathbf{W} = \frac{\partial \mathbf{W}}{\partial u^k} du^k . \quad (2.6.3)$$

The partial derivatives $\partial \mathbf{W}/\partial u^k$ follow from (2.6.1) or (2.6.2):

$$\frac{\partial \mathbf{W}}{\partial u^k} = \frac{\partial W^i}{\partial u^k} \mathbf{e}_i + W^i \frac{\partial \mathbf{e}_i}{\partial u^k} \quad (2.6.4)$$

or

$$\frac{\partial \mathbf{W}}{\partial u^k} = \frac{\partial W_i}{\partial u^k} \mathbf{e}^i + W_i \frac{\partial \mathbf{e}^i}{\partial u^k} . \quad (2.6.5)$$

For a single value of k , these expressions represent vectors, which can in turn be written with covariant or contravariant components:

$$\frac{\partial \mathbf{W}}{\partial u^k} = \left(\frac{\partial \mathbf{W}}{\partial u^k} \right)^j \mathbf{e}_j \equiv \left(\frac{\partial \mathbf{W}}{\partial u^k} \cdot \mathbf{e}^j \right) \mathbf{e}_j \quad (2.6.6)$$

or

$$\frac{\partial \mathbf{W}}{\partial u^k} = \left(\frac{\partial \mathbf{W}}{\partial u^k} \right)_j \mathbf{e}^j \equiv \left(\frac{\partial \mathbf{W}}{\partial u^k} \cdot \mathbf{e}_j \right) \mathbf{e}^j \quad (2.6.7)$$

where $(\partial \mathbf{W}/\partial u^k)^j$ and $(\partial \mathbf{W}/\partial u^k)_j$ represent the contravariant and covariant components of the vector $\partial \mathbf{W}/\partial u^k$, respectively. By combining (2.6.4) through (2.6.7), we can write the components of the partial derivative of \mathbf{W} with respect to u^k in four different ways:

$$\left(\frac{\partial \mathbf{W}}{\partial u^k} \right)^j = \frac{\partial W^i}{\partial u^k} \mathbf{e}_i \cdot \mathbf{e}^j + W^i \frac{\partial \mathbf{e}_i}{\partial u^k} \cdot \mathbf{e}^j \quad (2.6.8)$$

$$\left(\frac{\partial \mathbf{W}}{\partial u^k} \right)_j = \frac{\partial W_i}{\partial u^k} \mathbf{e}^i \cdot \mathbf{e}^j + W_i \frac{\partial \mathbf{e}^i}{\partial u^k} \cdot \mathbf{e}^j \quad (2.6.9)$$

for the contravariant components, and

$$\left(\frac{\partial \mathbf{W}}{\partial u^k} \right)_j = \frac{\partial W^i}{\partial u^k} \mathbf{e}_i \cdot \mathbf{e}_j + W^i \frac{\partial \mathbf{e}_i}{\partial u^k} \cdot \mathbf{e}_j \quad (2.6.10)$$

$$\left(\frac{\partial \mathbf{W}}{\partial u^k} \right)_j = \frac{\partial W_i}{\partial u^k} \mathbf{e}^i \cdot \mathbf{e}_j + W_i \frac{\partial \mathbf{e}^i}{\partial u^k} \cdot \mathbf{e}_j \quad (2.6.11)$$

for the covariant components.

Two of the four expressions above contain a Kronecker delta, and thus reduce to a simpler form. For this reason, only (2.6.8) and (2.6.11) are normally retained for the components of the partial derivative of \mathbf{W} :

$$\left(\frac{\partial \mathbf{W}}{\partial u^k} \right)^j = \frac{\partial W^j}{\partial u^k} + W^i \left\{ \frac{\partial \mathbf{e}_i}{\partial u^k} \cdot \mathbf{e}^j \right\} \quad (2.6.12)$$

and

$$\left(\frac{\partial \mathbf{W}}{\partial u^k} \right)_j = \frac{\partial W_j}{\partial u^k} + W_i \left\{ \frac{\partial \mathbf{e}^i}{\partial u^k} \cdot \mathbf{e}_j \right\} . \quad (2.6.13)$$

If we wish to employ the contravariant components of \mathbf{W} , we use the formula leading to the contravariant components of $\partial \mathbf{W}/\partial u^k$ as given in (2.6.12) and analogously for the covariant components, where we use (2.6.13).

An important observation is that the components of the derivative of \mathbf{W} are *not* equal to the derivatives of the corresponding components, since terms involving the change of the basis vectors must be added. The expressions in (2.6.12) and (2.6.13) are called the components of the *covariant derivative*. $d\mathbf{W}$ of (2.6.3), with (2.6.12) and (2.6.13) substituted is called the *intrinsic* or *absolute differential*.

The reason for the adjectives must be sought in abstract-vector analysis. There one defines a vector by its components without using a symbol for the vector itself; one suppresses explicit reference to a basis. With that convention, the derivative of the “vector” would “at first sight” be the derivative of the component. One then has to construct a “correction term” – the second term of (2.6.12) and (2.6.13) – in a formal and artificial way such that “the derivative” has the right transformation properties.⁴ The adjective indicates that one is after the proper derivative.

⁴ After reading Chap. 13, the reader will see that one needs the second term in (2.6.12) and (2.6.13) to make sure that the LHS of these equations are the components of a second-order tensor.

In our notation, a component of a intrinsic or absolute differential is represented by δW^j (or δW_j) and is equal to

$$\begin{aligned}\delta W^j &\equiv (dW)^j = \left(\frac{\partial W}{\partial u^k} \right)^j du^k = \frac{\partial W^j}{\partial u^k} du^k + W^i e^j \cdot \frac{\partial e_i}{\partial u^k} du^k \\ &= dW^j + \left\{ e^j \cdot \frac{\partial e_i}{\partial u^k} \right\} W^i du^k.\end{aligned}\quad (2.6.14)$$

In the “component \equiv vector” convention one would say that δW^j (or δW_j) is the intrinsic or absolute derivative itself. It is clear that by always referring explicitly to a basis we do not require any patchwork. Everything follows neatly from straightforward use of the common rules of calculus. It is only when one wishes to use the convention “component \equiv vector” that one must be careful to construct all the terms.

The components of the covariant-derivative vector ($\partial W/\partial u^k$) – we still think of k as a fixed index – are granted new symbols:

$$W_{,k}^j \equiv \left(\frac{\partial W}{\partial u^k} \right)^j = \frac{\partial W^j}{\partial u^k} + W^i \left\{ \frac{\partial e_i}{\partial u^k} \cdot e^j \right\} \quad (2.6.15)$$

and

$$W_{j,k} \equiv \left(\frac{\partial W}{\partial u^k} \right)_j = \frac{\partial W_j}{\partial u^k} + W_i \left\{ \frac{\partial e^i}{\partial u^k} \cdot e_j \right\}. \quad (2.6.16)$$

The comma in front of k on the LHS indicates that the differentiation is with respect to u^k . (In some texts, (Mathews and Walker 1970; Butkov 1968), a semi-colon is used instead of a comma for the components of the covariant derivative; the comma is then reserved for simple partial derivatives, i.e., the first terms in (2.6.15) and (2.6.16).)

b) The Covariant Derivative in Terms of the Metric Coefficients

The information about the chosen basis is largely contained in the metric coefficients g_{ij} . It can be expected that the “correction terms” in (2.6.15) and (2.6.16) are expressible in terms of these metric coefficients. Consider the correction term in (2.6.15): $W^i \{(\partial e_i/\partial u^k) \cdot e^j\}$. Keeping k still fixed, the vector $\partial e_i/\partial u^k$ can be expanded in two ways, along e_i or e^j as follows:

$$\frac{\partial e_i}{\partial u^k} = \left\{ \frac{\partial e_i}{\partial u^k} \cdot e^j \right\} e_j \equiv \left(\frac{\partial e_i}{\partial u^k} \right)^j e_j \quad (2.6.17)$$

or

$$\frac{\partial e_i}{\partial u^k} = \left[\frac{\partial e_i}{\partial u^k} \cdot e_j \right] e^j \equiv \left(\frac{\partial e_i}{\partial u^k} \right)_j e^j. \quad (2.6.18)$$

The expression in $\{ \}$ in (2.6.17) is what we need in our correction term. Since $e^j = g^{jn} e_n$, we see from equating (2.6.17) and (2.6.18) that

$$\left\{ \frac{\partial e_i}{\partial u^k} \cdot e^j \right\} = g^{jn} \left[\frac{\partial e_i}{\partial u^k} \cdot e_n \right] \quad (2.6.19a)$$

$$= g^{jn} \frac{1}{2} \left[e_n \cdot \frac{\partial e_i}{\partial u^k} + e_n \cdot \frac{\partial e_k}{\partial u^i} \right]. \quad (2.6.19b)$$

Here, we have used the property

$$\frac{\partial e_i}{\partial u^k} = \frac{\partial}{\partial u^k} \frac{\partial \mathbf{R}}{\partial u^i} = \frac{\partial}{\partial u^i} \frac{\partial \mathbf{R}}{\partial u^k} = \frac{\partial e_k}{\partial u^i}. \quad (2.6.20)$$

The right-hand side of (2.6.19) can be manipulated as follows:

$$\begin{aligned}g^{jn} \frac{1}{2} \left[\frac{\partial}{\partial u^k} (e_n \cdot e_i) + \frac{\partial}{\partial u^i} (e_n \cdot e_k) - e_i \cdot \frac{\partial e_n}{\partial u^k} - e_k \cdot \frac{\partial e_n}{\partial u^i} \right] \\ = g^{jn} \frac{1}{2} \left[\frac{\partial g_{ni}}{\partial u^k} + \frac{\partial g_{nk}}{\partial u^i} - e_i \cdot \frac{\partial e_k}{\partial u^n} - e_k \cdot \frac{\partial e_i}{\partial u^n} \right] \\ = g^{jn} \frac{1}{2} \left[\frac{\partial g_{ni}}{\partial u^k} + \frac{\partial g_{nk}}{\partial u^i} - \frac{\partial}{\partial u^n} (e_i \cdot e_k) \right].\end{aligned}$$

Hence, we obtain

$$\left\{ \frac{\partial e_i}{\partial u^k} \cdot e^j \right\} = \frac{1}{2} g^{jn} \left[\frac{\partial g_{ni}}{\partial u^k} + \frac{\partial g_{nk}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^n} \right]. \quad (2.6.21)$$

To find a similar correction term in (2.6.16) we observe that $e_j \cdot e^i = \delta_j^i$, so

$$\frac{\partial}{\partial u^k} (e_j \cdot e^i) = 0$$

and

$$\left\{ e_j \cdot \frac{\partial e^i}{\partial u^k} \right\} = - \left\{ e^i \cdot \frac{\partial e_j}{\partial u^k} \right\}. \quad (2.6.22)$$

The part of the correction term of (2.6.16) inside the brackets is almost the same as in (2.6.21) but with a minus sign, and the indices i and j interchanged. In summary,

$$W_{,k}^j = \frac{\partial W^j}{\partial u^k} + \frac{1}{2} g^{jn} \left[\frac{\partial g_{ni}}{\partial u^k} + \frac{\partial g_{nk}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^n} \right] W^i \quad (2.6.23)$$

and

$$W_{j,k} = \frac{\partial W_j}{\partial u^k} - \frac{1}{2} g^{in} \left[\frac{\partial g_{nj}}{\partial u^k} + \frac{\partial g_{nk}}{\partial u^j} - \frac{\partial g_{jk}}{\partial u^n} \right] W_i . \quad (2.6.24)$$

c) The Christoffel Symbols

The *Christoffel symbol of the second kind* is defined as

$$\left\{ \begin{matrix} j \\ i \ k \end{matrix} \right\} \equiv \left\{ e^j \cdot \frac{\partial e_i}{\partial u^k} \right\} = \frac{1}{2} g^{jn} \left[\frac{\partial g_{ni}}{\partial u^k} + \frac{\partial g_{nk}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^n} \right] . \quad (2.6.25)$$

This second Christoffel symbol is the j th contravariant component of the vector $\partial e_i / \partial u^k$ (for k fixed). Analogously, the *Christoffel symbol of the first kind* is the j th covariant component of the same vector $\partial e_i / \partial u^k$:

$$[j, ik] \equiv \left[e_j \cdot \frac{\partial e_i}{\partial u^k} \right] = \frac{1}{2} \left[\frac{\partial g_{ji}}{\partial u^k} + \frac{\partial g_{jk}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^j} \right] . \quad (2.6.26)$$

Some authors use $[ik, j]$. The second part of (2.6.26) follows from (2.6.19a) recognizing that $1/2 [\dots]$ of (2.6.25), without the g^{jn} , equals $[n, ik]$. The comma in $[j, ik]$ does not refer to the differentiating index here as can be seen from its definition. A better notation would perhaps have been $[j/ik]$.

The positions of the indices in the Christoffel symbols are very important. They are a shorthand for the $\{ \}$ and $[]$ in (2.6.17) and (2.6.18): delete the e 's but maintain the index positions. As usual, an upper index refers to a contravariant quantity, etc. Because of (2.6.20) we observe that the Christoffel symbols are symmetric in the indices appearing in $\partial e_i / \partial u^k$. Thus,

$$\left\{ \begin{matrix} j \\ i \ k \end{matrix} \right\} = \left\{ \begin{matrix} j \\ k \ i \end{matrix} \right\} \quad (2.6.27)$$

and

$$[j, ik] = [j, ki] . \quad (2.6.28)$$

The relationship between the two kinds of Christoffel symbols follows from (2.6.19a)

$$[i, jk] = g_{in} \left\{ \begin{matrix} n \\ j \ k \end{matrix} \right\} \quad (2.6.29)$$

$$\left\{ \begin{matrix} i \\ j \ k \end{matrix} \right\} = g^{in} [n, jk] . \quad (2.6.30)$$

Alternative symbols for the Christoffel symbols are $\Gamma_{jk}^i \equiv \left\{ \begin{matrix} i \\ j \ k \end{matrix} \right\}$ and $\Gamma_{i,jk} \equiv [i, jk]$. These Γ -symbols suggest a tensor character, which they do not have (see Chap. 3 or 13). Note that the so-called "Christoffel symbols" as defined by Stacey (1981) differ from those defined here and by other authors (Morse and Feshbach 1953, Spiegel 1959).

The introduction of the Christoffel symbols allows us to rewrite the components of the covariant derivative, given in (2.6.23) and (2.6.24) in a compact form:

$$W_{j,k}^i = \frac{\partial W_j^i}{\partial u^k} + \left\{ \begin{matrix} j \\ i \ k \end{matrix} \right\} W_i \quad (2.6.31)$$

and

$$W_{j,k} = \frac{\partial W_j}{\partial u^k} - \left\{ \begin{matrix} i \\ j \ k \end{matrix} \right\} W_i . \quad (2.6.32)$$

In all the foregoing, the summation convention was used.

The covariant derivative was obtained using the ordinary rules of differentiation. Thus the derivative of a sum is the sum of the derivatives and the derivative of a product is the sum of similar products where one of the factors, in turn, is replaced by its derivative.

By definition, the Christoffel symbols originate from the dependence on position of the basis vectors, so in a rectilinear coordinate system, the Christoffel symbols are zero.

2.6.2 The Del Operator

a) Definition of the Del Operator

The del operator, ∇ , is usually defined via the relationship $d\Phi = \nabla\Phi \cdot dR$. Expansion of $d\Phi$ via the chain rule leads to $d\Phi = (\partial\Phi/\partial u^i) du^i$, which can be written as a dot product using $\delta_j^i = e^i \cdot e_j$:

$$\frac{\partial\Phi}{\partial u^i} du^i = \frac{\partial\Phi}{\partial u^i} e^i \cdot e_j du^j = \frac{\partial\Phi}{\partial u^i} \nabla u^i \cdot e_j du^j . \quad (2.6.33)$$

Identification of this expression with $\nabla\Phi \cdot dR$ shows that we can formally introduce the *del operator* as

$$\nabla \equiv \nabla u^i \frac{\partial}{\partial u^i} \equiv e^i \frac{\partial}{\partial u^i} . \quad (2.6.34)$$

We now derive the formulae for the gradient of a scalar field, the divergence and the curl of a vector field, using the material of the previous sections.

b) Gradient

We operate with ∇ on a scalar function $\Phi(u^1, u^2, u^3)$ and simply juxtapose the two symbols with *no operator sign* in between them. We obtain:

$$\text{grad } \Phi \equiv \nabla\Phi = \nabla u^i \frac{\partial\Phi}{\partial u^i} = \frac{\partial\Phi}{\partial u^i} e^i \quad (2.6.35)$$

or

$$(\nabla\Phi)_i = \frac{\partial\Phi}{\partial u^i} . \quad (2.6.36)$$

Equation (2.6.36) is the covariant component of the gradient vector and does not necessarily have the same units as $\nabla\Phi$, since the reciprocal-basis vectors can have units themselves.

c) Divergence

The simplest derivation of the divergence makes use of the properties $\nabla \times \nabla\Phi \equiv 0$ and $\nabla \cdot (\nabla \times A) \equiv 0$ (S.P. Hirshman 1982). We first note that the $e_i \leftrightarrow e^j$ transformation formulae of (2.5.25) can be written in a curl form:

$$\frac{e_i}{J} = (\nabla u^j \times \nabla u^k) = \nabla \times (u^j \nabla u^k) \quad i, j, k \text{ cyc } 1, 2, 3 . \quad (2.6.37)$$

Because e_i/J is the curl of a vector, its divergence is zero:

$$\nabla \cdot \left(\frac{e_i}{J} \right) = \nabla \cdot [\nabla \times (u^j \nabla u^k)] \equiv 0 . \quad (2.6.38)$$

Therefore, we have for the *divergence* of a vector A :

$$\begin{aligned} \nabla \cdot A &= \nabla \cdot (A^j e_j) = \nabla \cdot \left(J A^j \frac{e_j}{J} \right) = J A^j \nabla \cdot \left(\frac{e_j}{J} \right) + \frac{e_j}{J} \cdot \nabla (J A^j) \\ &= \frac{e_j}{J} \cdot \nabla (J A^j) , \end{aligned}$$

which upon using the definition of the ∇ -operator, (2.6.34), becomes

$$\nabla \cdot A = \frac{1}{J} \frac{\partial}{\partial u^i} (J A^i) = \frac{1}{\sqrt{g}} \frac{\partial}{\partial u^i} (\sqrt{g} A^i) . \quad (2.6.39)$$

For completeness, we discuss two alternative derivations of the formula in (2.6.39) in Chap. 14. It concerns those that are encountered most frequently in the literature but quite often with unsatisfactory notations.

d) Curl

The curl of a vector field is obtained as the cross-product of the ∇ -operator with the vector A , $\nabla \times A$ or sometimes represented by $\text{rot } A$.

First, we rewrite (2.5.25), with the constraint “ $i, j, k, \text{cyc } 1, 2, 3$ ” absorbed into the formula:

$$e^i \times e^j = \epsilon^{ijk} \frac{e_k}{J} . \quad (2.6.40)$$

The curl of A is then

$$\nabla \times A = \nabla \times (A_j e^j) = A_j (\nabla \times e^j) + \nabla A_j \times e^j = \frac{\partial A_j}{\partial u^i} e^i \times e^j . \quad (2.6.41)$$

Here we have used $\nabla \times \nabla u^j \equiv 0$ and the definition of the del operator. Upon using (2.6.40) we obtain

$$\nabla \times A = \frac{\epsilon^{ijk}}{J} \frac{\partial A_j}{\partial u^i} e_k = \frac{\epsilon^{ijk}}{\sqrt{g}} \frac{\partial A_j}{\partial u^i} e_k . \quad (2.6.42)$$

More explicitly,

$$\begin{aligned} \nabla \times A &= \frac{1}{J} \sum_k \left(\frac{\partial A_j}{\partial u^i} - \frac{\partial A_i}{\partial u^j} \right) e_k \quad i, j, k \text{ cyc } 1, 2, 3 \\ &= \frac{1}{\sqrt{g}} \sum_k \left(\frac{\partial A_j}{\partial u^i} - \frac{\partial A_i}{\partial u^j} \right) e_k \quad i, j, k \text{ cyc } 1, 2, 3 . \end{aligned} \quad (2.6.43)$$

This derivation may also be performed via the covariant derivative:

$$\begin{aligned} \nabla \times A &= e^i \frac{\partial}{\partial u^i} \times A = e^i \times \frac{\partial A}{\partial u^i} = e^i \times e^j \left(\frac{\partial A}{\partial u^i} \right)_j \\ &= e^i \times e^j A_{j,i} = \frac{1}{J} \sum'_{ij} (\pm) A_{j,i} e_k \quad i, j, k \text{ cyc } 1, 2, 3 . \end{aligned} \quad (2.6.44)$$

The prime on \sum again implies $i \neq j$. We used (2.5.37) and (2.5.38) for the cross product. This implies:

$$\nabla \times A = J^{-1} \sum_k (A_{j,i} - A_{i,j}) e_k \quad i, j, k \text{ cyc } 1, 2, 3 . \quad (2.6.45)$$

After substitution of the expressions for $A_{j,i}$ and $A_{i,j}$ according to (2.6.32) and recognition that the Christoffel symbols cancel each other because of their symmetry properties, we retrieve (2.6.43).

2.7 The Parallel and Perpendicular Components of a Vector

Sometimes one wants to find the parallel or perpendicular components of a vector V with respect to a certain given vector, say B (e.g., the magnetic field).

The dot product is obtained via the orthogonal projection of the one vector onto the other. Consequently, it should be clear that the *parallel* component of V (denoted by $V_{||}$) with respect to (w.r.t.) a vector B is simply:

$$V_{\parallel} = \frac{\mathbf{V} \cdot \mathbf{B}}{B} = \mathbf{V} \cdot \hat{\mathbf{B}}, \quad (2.7.1)$$

$\hat{\mathbf{B}}$ being the unit vector along \mathbf{B} .

On the other hand, the component of \mathbf{V} perpendicular to \mathbf{B} , can be derived from the definition of the vector product:

$$\mathbf{V}_{\perp} = -\hat{\mathbf{B}} \times (\hat{\mathbf{B}} \times \mathbf{V}). \quad (2.7.2)$$

Indeed, $|\hat{\mathbf{B}} \times \mathbf{V}|$ is the magnitude of the perpendicular component (w.r.t. $\hat{\mathbf{B}}$) and $-\hat{\mathbf{B}} \times (\hat{\mathbf{B}} \times \mathbf{V})$ is the vector \perp to $\hat{\mathbf{B}}$, lying in the plane formed by the vectors $\hat{\mathbf{B}}$ and \mathbf{V} , which means that it must be in the direction of $\pm \mathbf{V}_{\perp}$.

Thus, any vector \mathbf{V} can be written in terms of its parallel and perpendicular components with respect to another vector \mathbf{B} as:

$$\mathbf{V} = V_{\parallel} \hat{\mathbf{B}} + \mathbf{V}_{\perp} = \hat{\mathbf{B}}(\hat{\mathbf{B}} \cdot \mathbf{V}) - \hat{\mathbf{B}} \times (\hat{\mathbf{B}} \times \mathbf{V}), \quad (2.7.3)$$

where $\hat{\mathbf{B}} = \mathbf{B}/B$.

It is useful to know how the \parallel and \perp labels can “shift” in dot products (\parallel and \perp are with respect to an arbitrary vector \mathbf{B} , which can coincide with \mathbf{A} or \mathbf{C}). We obtain:

$$\mathbf{A}_{\perp} \cdot \mathbf{C} = \mathbf{A}_{\perp} \cdot \mathbf{C}_{\perp} = \mathbf{A} \cdot \mathbf{C}_{\perp} \quad (2.7.4)$$

$$\mathbf{A}_{\parallel} \cdot \mathbf{C} = \mathbf{A}_{\parallel} \mathbf{C}_{\parallel} = \mathbf{A} \cdot \mathbf{C}_{\parallel}. \quad (2.7.5)$$

2.8 A Summary of Vector Related Identities

Here we summarize the most important vector identities used in plasma physics, of the NRL Plasma Formulary (Book 1983). More formulae involving vectors and tensors can be found at the end of Chap. 3.

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} \times \mathbf{C} &= \mathbf{A} \times \mathbf{B} \cdot \mathbf{C} = \mathbf{B} \cdot \mathbf{C} \times \mathbf{A} = \mathbf{B} \times \mathbf{C} \cdot \mathbf{A} \\ &= \mathbf{C} \cdot \mathbf{A} \times \mathbf{B} = \mathbf{C} \times \mathbf{A} \cdot \mathbf{B} \end{aligned} \quad (V.1)$$

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{C} \times \mathbf{B}) \times \mathbf{A} = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}) \quad (V.2)$$

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) + \mathbf{B} \times (\mathbf{C} \times \mathbf{A}) + \mathbf{C} \times (\mathbf{A} \times \mathbf{B}) = 0 \quad (V.3)$$

$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C}) \quad (V.4)$$

$$(\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \times \mathbf{B} \cdot \mathbf{D})\mathbf{C} - (\mathbf{A} \times \mathbf{B} \cdot \mathbf{C})\mathbf{D} \quad (V.5)$$

$$\nabla(fg) = \nabla(gf) = f\nabla g + g\nabla f \quad (V.6)$$

$$\nabla \cdot (f\mathbf{A}) = f\nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla f \quad (V.7)$$

$$\nabla \times (f\mathbf{A}) = f\nabla \times \mathbf{A} + \nabla f \times \mathbf{A} \quad (V.8)$$

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B} \quad (V.9)$$

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = \mathbf{A}(\nabla \cdot \mathbf{B}) - \mathbf{B}(\nabla \cdot \mathbf{A}) + (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} \quad (V.10)$$

$$\mathbf{A} \times (\nabla \times \mathbf{B}) = (\nabla \mathbf{B}) \cdot \mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} \quad (V.11)$$

$$\nabla(\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A} \quad (V.12)$$

$$\nabla^2 f = \nabla \cdot \nabla f \quad (V.13)$$

$$\nabla^2 \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla \times \nabla \times \mathbf{A} \quad (V.14)$$

$$\nabla \times \nabla f = 0 \quad (V.15)$$

$$\nabla \cdot \nabla \times \mathbf{A} = 0 \quad (V.16)$$

$$\nabla \cdot \mathbf{R} = 3, \quad \mathbf{R} = \text{position vector}. \quad (V.17)$$

3. Tensorial Objects

3.1 Introduction

A brief discussion of second-order tensorial objects such as dyads and dyadics is necessary in a comprehensive treatment of flux coordinates. The reason is twofold. Firstly, we regularly deal with dyadic objects in studying plasma equilibrium, even in scalar pressure cases. We frequently come across expressions such as $\mathbf{A} \cdot \mathbf{BC}$ or $\mathbf{A} \cdot \nabla \mathbf{B}$, where \mathbf{BC} and $\nabla \mathbf{B}$ are dyads. An important example is the curvature along a magnetic-field line: $\hat{\mathbf{B}} \cdot \nabla \hat{\mathbf{B}}$. Secondly, tensor analysis in curvilinear coordinates is merely an extension of vector analysis; more accurately, vectors are a special case of tensors: vectors are first-order tensors.

Tensors are of considerable importance in general plasma physics. Examples of tensors used are the above mentioned dyads of the form \mathbf{AB} or $\nabla \mathbf{B}$ needed for a description of the curvature and fluid inertial terms $\mathbf{v} \cdot \nabla \mathbf{v}$ among others, the pressure and viscous-stress tensors, the electromagnetic stress tensor, the dyadics utilized in the derivation of the Fokker-Planck equation, etc.

This is not the place to give a detailed exposition on tensors, but we can explain the main ideas. We shall take a pragmatic viewpoint and stress the similarities with and generalizations of vectors.

3.2 The Concept of Tensors; A Pragmatic Approach

The simplest way to introduce tensors is via a dot-product operation between vectors. A *dyad* or *second-order “simple” tensor*, \mathbf{AB} , is defined by

$$\mathbf{X} \cdot \mathbf{AB} \equiv (\mathbf{X} \cdot \mathbf{A})\mathbf{B} \quad (3.2.1)$$

or

$$\mathbf{AB} \cdot \mathbf{Y} \equiv \mathbf{A}(\mathbf{B} \cdot \mathbf{Y}) . \quad (3.2.2)$$

\mathbf{X} and \mathbf{Y} are two arbitrary vectors.

The result of (3.2.1) and (3.2.2) is a vector as is clear from their RHS. The LHS states that this vector can be considered as being produced by the dot product of a vector and a dyad. The notation for a dyad is a simple juxtaposition of the two vectors making up the dyad, without a symbol in between. In more

abstract mathematical treatments, the symbol \otimes is used and is called the tensor product or the outer product (sometimes it is called the exterior product, although this name is mostly reserved for yet another product, namely, that related to Grassmann Algebra). Thus, $\mathbf{AB} \equiv \mathbf{A} \otimes \mathbf{B}$ in that notation. Note that the Russian notations for dot and cross products are also juxtapositions, distinguished by different brackets: $(\mathbf{AB}) \equiv$ what we write as $\mathbf{A} \cdot \mathbf{B}$ and $[\mathbf{AB}] \equiv$ what we write as $\mathbf{A} \times \mathbf{B}$.

From the definition, (3.2.1) and (3.2.2), it follows that the dot product of a dyad with a vector is (in general) *not commutative*. The result of $\mathbf{X} \cdot \mathbf{AB} = (\mathbf{X} \cdot \mathbf{A})\mathbf{B}$ is a vector along \mathbf{B} , whereas $\mathbf{AB} \cdot \mathbf{X} = \mathbf{A}(\mathbf{B} \cdot \mathbf{X})$ is a vector proportional to \mathbf{A} . These equations also explain why *the order of the vectors making up a dyad is important* i.e.,

$$\mathbf{AB} \neq \mathbf{BA}. \quad (3.2.3)$$

The reason is self-explanatory; $\mathbf{X} \cdot \mathbf{AB} = (\mathbf{X} \cdot \mathbf{A})\mathbf{B} \propto \mathbf{B}$, whereas $\mathbf{X} \cdot \mathbf{BA} = (\mathbf{X} \cdot \mathbf{B})\mathbf{A} \propto \mathbf{A}$.

We now expand the vectors \mathbf{A} and \mathbf{B} along the basis vectors $\mathbf{e}_i \equiv \partial \mathbf{R} / \partial u^i$ and $\mathbf{e}^j = \nabla u^j$; $\mathbf{A} = A^i \mathbf{e}_i = A_j \mathbf{e}^j$ with $A^i = \mathbf{A} \cdot \mathbf{e}^i$ and $A_j = \mathbf{A} \cdot \mathbf{e}_j$, and $\mathbf{B} = B^k \mathbf{e}_k = B_i \mathbf{e}^i$ with $B^k = \mathbf{B} \cdot \mathbf{e}^k$ and $B_i = \mathbf{B} \cdot \mathbf{e}_i$. The summation convention must be used over repeated indices. This leads to four possible forms of the dyad:

$$\mathbf{AB} = A^i \mathbf{e}_i B^k \mathbf{e}_k = A^i B^k \mathbf{e}_i \mathbf{e}_k \quad (3.2.4a)$$

$$\mathbf{AB} = A^i \mathbf{e}_i B_l \mathbf{e}^l = A^i B_l \mathbf{e}_i \mathbf{e}^l \quad (3.2.4b)$$

$$\mathbf{AB} = A_j \mathbf{e}^j B^k \mathbf{e}_k = A_j B^k \mathbf{e}^j \mathbf{e}_k \quad (3.2.4c)$$

$$\mathbf{AB} = A_j \mathbf{e}^j B_l \mathbf{e}^l = A_j B_l \mathbf{e}^j \mathbf{e}^l. \quad (3.2.4d)$$

The second form of (3.2.4a-d) follows trivially from the definition, (3.2.1) and (3.2.2).

Sometimes, we use a general symbol $A(m) B(n)$ to denote the four variants of \mathbf{A} 's and \mathbf{B} 's when we make statements that are valid for all of them. In three dimensions, there are nine possible combinations $A(m) B(n)$ in each of the four cases above. The coefficients $A(m) B(n)$, i.e., $A^i B^k$, $A^i B_l$, $A_j B^k$, $A_j B_l$ are called the components of the dyad for the basis vector sets \mathbf{e}_i and \mathbf{e}^j . The dyads $\mathbf{e}_i \mathbf{e}_k$ are called “basis-dyads” or “basis-tensors” in analogy with basis vectors. There are nine “basis-dyads” of each sort (in the sense of (3.2.4a-d) in a three-dimensional space).

We introduce first a single symbol for the dyad \mathbf{AB}

$$\mathbf{D} \equiv \mathbf{AB}. \quad (3.2.5)$$

Second-order tensorial quantities are represented by a bold face, sanserif, italic notation. (Other notations instead of \mathbf{D} are \vec{D} , $\vec{\vec{D}}$ or $\underline{\underline{D}}$; the two “arrow heads” or the two bars refer to the second-order character.) Depending on the index positions, we have the following definitions:

$$D^{ik} \equiv A^i B^k \equiv \text{contravariant components of } \mathbf{AB} (\equiv \mathbf{D}) \quad (3.2.6a)$$

$$D_{jl} \equiv A_j B_l \equiv \text{covariant components of } \mathbf{AB} (\equiv \mathbf{D}) \quad (3.2.6b)$$

$$\begin{aligned} D_{i,l}^i &\equiv A^i B_l \\ D_j^k &\equiv A_j B^k \equiv \text{mixed components of } \mathbf{AB} (\equiv \mathbf{D}). \end{aligned} \quad (3.2.6c)$$

The dots in the symbols $D_{i,l}^i$ and D_j^k remind us that i and j are the first indices in each case. The dot is necessary because otherwise, D_i^l could be interpreted as either the components of $\mathbf{D} = D_i^l e^i e^l$ or as the components of $\mathbf{D} = D_i^l e^l e^i$. But since, in general, $e^l e^i$ is different from $e^i e^l$, the D_i^l are also different, creating the need to indicate which basis vector must come first.

The components $D(m, n)$ can be ordered as a square array. They are then said to form a 3×3 matrix. In our co- and contravariant formalism, there are four different matrices belonging to the same dyad, at least for the choice $e_i = \partial R / \partial u^i$ and $e^j = \nabla u^j$. When we change basis vectors, the components of the dyad will change as will the matrices. As is the case for vectors, tensors (or dyads) in Euclidean space are no longer covariant, contravariant or mixed. All we have are dyads or tensors which can be expressed in covariant, contravariant or mixed components, depending on the choice of basis vectors.

We now examine \mathbf{D} from another viewpoint. The second-order tensorial object \mathbf{D} that we referred to as a dyad or “simple” second-order tensor has a particular set of components which, in a certain sense, is special. The nine components, say D_{jl} , can be written as the mixed product of two groups of three numbers: (A_1, A_2, A_3) and (B_1, B_2, B_3) , such that $D_{jl} = A_j B_l$, $j = 1, 2, 3$ and $l = 1, 2, 3$. An analogous statement can be made for the nine contravariant, and 18 mixed components of that same second-order tensor \mathbf{D} , as shown in (3.2.6). This means that \mathbf{D} is a tensor which is produced by two vectors, \mathbf{AB} .

We now treat more general second-order tensors. Consider a second-order tensor \mathbf{F}

$$\mathbf{F} = F^{ik} e_i e_k = F_{jl} e^j e^l = F_{i,l}^i e^i e^l = F_j^k e^j e_k. \quad (3.2.7)$$

For a general second-order tensor, it is *not* possible to find two sets of three numbers in the sense explained above for \mathbf{D} . In other words, a general second-order tensor cannot be written as a single dyad. It can, however, be written as a sum of at least three dyads (in 3-D).

Let us show this for a tensor written with covariant components. Choose any three vectors \mathbf{A} , \mathbf{B} and \mathbf{C} with the constraints that the 3×3 matrix formed by their components A_i , B_j , and C_k has a non-zero determinant. We then must find three vectors \mathbf{X} , \mathbf{Y} and \mathbf{Z} such that

$$\mathbf{F} = \mathbf{XA} + \mathbf{YB} + \mathbf{YC}. \quad (3.2.8)$$

When \mathbf{F} is written in terms of its covariant components F_{jl} and all the vectors appearing on the RHS of (3.2.8) are written using their covariant components as

well, we can equate the coefficients of the same “basis-dyads” $e^j e^l$ on the LHS and RHS:

$$F_{jl} = X_j A_l + Y_j B_l + Z_j C_l. \quad (3.2.9)$$

In matrix language, the RHS is the matrix product of the matrices formed by the components A_i , B_j , and C_k and the transpose of that formed by the components X_i , Y_m , Z_n :

$$[\mathbf{F}] = [\chi]^T [\alpha], \quad (3.2.10)$$

Here, $\chi_{11} = X_1$, $\chi_{12} = X_2$, $\chi_{13} = X_3$, $\chi_{21} = Y_1$, etc., and $\alpha_{11} = A_1$, $\alpha_{12} = A_2$, $\alpha_{13} = A_3$, $\alpha_{21} = B_1$, etc. Since we required that $\det[\alpha] \neq 0$ from the outset, the matrix equation, (3.2.10), can be inverted, and the coefficients of $[\chi]$ can be found.

To write \mathbf{F} with more than three dyads is straightforward. For instance, any dyad can be replaced by the sum of a symmetric and an antisymmetric dyad: $\mathbf{D} \equiv (\mathbf{D} + \mathbf{D}^T)/2 + (\mathbf{D} - \mathbf{D}^T)/2$. Here, T stands for transpose. Also, if we absorb the coefficients in the second form of (3.2.7) into other vectors, so that e.g., $a = F_{11} e^1$, $b = e^1$, $c = e^1$, $d = F_{12} e^2$, $e = \sqrt{F_{13}} e^3$, ..., we can write \mathbf{F} as a sum of nine dyads

$$\mathbf{F} = ab + cd + ef + \dots \quad (3.2.11)$$

Because a general second-order tensor can be written as a sum of (at least three) dyads, it is also called a *dyadic*. It is clear that a dyad is a dyadic, but not all dyadics are dyads. The expressions in (3.2.4) and (3.2.7) are called the *nonion* form of the dyad or the dyadic to denote the “nine-component form”.

The words “second-order tensor” and “dyadic” refer to one and the same “geometrical object”, although they might be regarded as different representations for the same entity. If one deals with the indexed-component form, one could refer to a “tensor”; if, on the other hand, one considers the “geometrical object” as a sum of dyads, one could refer to a “dyadic”. This is a matter of taste. Since there is no real difference between the two, we will use the words *dyadic* and *second order tensor* as *synonyms*.

We have encountered two other interpretations of dyadics, which we think have no ground. First, Morse and Feshbach (1953) state that the relationship between the components of a dyadic and those of a second-order tensor is analogous to that between the physical components of a vector and its covariant or contravariant components. They essentially state that a dyadic is to a tensor what a real “vector” is to a co- or contravariant “vector”. Since they subsequently treat dyadics as if they are common second-order tensors, and since we found no other evidence in the literature to back up this notion, we do not accept this interpretation. Another (restricted) definition is given by Drew (1961), who asserts that dyads and dyadics refer to tensor analysis in Cartesian coordinate systems. Again, we found no other reference stating the same. This interpretation seems too restricted and is actually contradicted by most applications in plasma

physics (e.g., the pressure tensor is really a dyad, and the viscous-stress tensor is a dyadic; however, we almost never express them in Cartesian components).

Higher-order tensors can be defined in a “similar” way. A third-order “simple” tensor or a triad \mathbf{ABC} is defined as

$$\mathbf{Y} \cdot \mathbf{ABC} = (\mathbf{Y} \cdot \mathbf{A})\mathbf{BC}, \quad (3.2.12)$$

where \mathbf{BC} is a dyad, a by now familiar object. Triadics or general third-order tensors can be written as sums of triads. For fourth-order tensorial objects, we talk about tetrads and tetradiics. For the n -th order case, we refer to polyads and polyadics.

A vector can be considered as a first-order adic or first-order tensor, and a scalar as a zeroth-order tensor.

3.3 Dot Product, Double Dot Product, Contraction

The dot product of a vector and a tensor is merely an extension of the definition of a dyad.

In this section, we shall stress the index notation, so by straightforward expansions in components:

$$\begin{aligned} \mathbf{X} \cdot \mathbf{F} &= X_m e^m \cdot F^{ik} e_i e_k = X_m F^{ik} e^m \cdot e_i e_k = X_i F^{ik} e_k = Y^k e_k = \mathbf{Y} \\ &= X^n e_n \cdot F_j^k e^j e_k = X^n F_j^k e_n \cdot e^j e_k = X^j F_j^k e_k = Y^k e_k = \mathbf{Y} \\ &= X^n e_n \cdot F_{jl} e^j e^l = X^n F_{jl} e_n \cdot e^j e^l = X^j F_{jl} e^l = Y_l e^l = \mathbf{Y} \\ &= X_m e^m \cdot F_{il} e_i e^l = X_m F_{il} e^m \cdot e_i e^l = X_i F_{il} e^l = Y_l e^l = \mathbf{Y} \end{aligned} \quad (3.3.1a)$$

for $\mathbf{X} \cdot \mathbf{F} = \mathbf{Y}$, and

$$\begin{aligned} \mathbf{F} \cdot \mathbf{X} &= F^{ik} e_i e_k \cdot X_m e^m = F^{ik} X_m e_i e_k \cdot e^m = F^{ik} X_k e_i = Z^i e_i = \mathbf{Z} \\ &= F_{il}^i e_i e^l \cdot X^n e_n = F_{il}^i X^n e_i e^l \cdot e_n = F_{il}^i X^l e_i = Z^i e_i = \mathbf{Z} \\ &= F_{jl} e^j e^l \cdot X^m e_m = F_{jl} X^m e^j e^l \cdot e_m = F_{jl} X^l e^j = Z_j e^j = \mathbf{Z} \\ &= F_j^k e^i e_k \cdot X_n e^n = F_j^k X_n e^i e_k \cdot e^n = F_j^k X_k e^i = Z_i e^i = \mathbf{Z} \end{aligned} \quad (3.3.1b)$$

for $\mathbf{F} \cdot \mathbf{X} = \mathbf{Z}$.

Here we used the fact that $e^i \cdot e_j = \delta_j^i$. Note that we always performed the dot product between a basis vector with a subscript and one with a superscript. That way we were able to avoid using the $g_{ij} = e_i \cdot e_j$. In component form, the dot product amounts to a *saturation* over one upper index and one lower index:

$$\begin{aligned} (\mathbf{X} \cdot \mathbf{F})^k &= X_i F^{ik} = X^j F_j^k = Y^k \\ (\mathbf{X} \cdot \mathbf{F})_l &= X^j F_{jl} = X_i F_{il} = Y_l \end{aligned} \quad (3.3.1c)$$

and

$$\begin{aligned} (\mathbf{F} \cdot \mathbf{X})^i &= F^{ik} X_k = F_{il}^i X^l = Z^i \\ (\mathbf{F} \cdot \mathbf{X})_j &= F_{jl} X^l = F_j^k X_k = Z_j \end{aligned} \quad (3.3.1d)$$

It is clear that the index which is the furthest away from where the dot “would have been” is kept unchanged. The index saturation is over the indices closest to the “dot”: the first one of \mathbf{F} in $\mathbf{X} \cdot \mathbf{F}$, and the second one in $\mathbf{F} \cdot \mathbf{X}$. The saturation of the indices is formally called the *contraction* of a tensor: the summation over two identical indices, one upper and one lower. The contraction of a tensor results in a tensor of a rank two lower than that of the original tensor. The dot product of a vector with a tensor can be considered as the contraction of a third-order tensor. If we represent the tensors by their components, we obtain, e.g.,

$$P_i^{ik} \equiv X_i F^{ik} = Y^k \quad (3.3.2)$$

and

$$P_{jl}^j \equiv X^j F_{jl} = Y_l.$$

The dot product between two second-order tensors is likewise *not* commutative. It can be considered as the contraction of a fourth-order tensor; the saturation is over the middle indices, and the result is a second-order tensor. We have for $\mathbf{E} = \mathbf{F} \cdot \mathbf{V}$

$$\begin{aligned} F_{ij} V^{jk} &= F_i^j V_j^k = E_i^k \\ F^{ij} V_{jk} &= F_{.j}^i V_{.k}^j = E_{.k}^i \\ F_i^{ij} V_{jk} &= F_{ij} V_{.k}^j = E_{ik} \\ F_{.j}^i V^{jk} &= F^{ij} V_{.k}^k = E^{ik}. \end{aligned} \quad (3.3.3)$$

Considered as the contraction of a fourth-order tensor, we would have, for instance,

$$Q_{ij}^{jk} \equiv F_{ij} V^{jk} = E_i^k. \quad (3.3.4)$$

Consider now the dot product of a tensor with two vectors, one on its left and one on its right:

$$\begin{aligned} \mathbf{A} \cdot \mathbf{F} \cdot \mathbf{X} &= \mathbf{A} \cdot (\mathbf{F} \cdot \mathbf{X}) = \mathbf{A} \cdot \mathbf{Z} = A^i (F_{ij} X^j) = A^i Z_i \\ &= (\mathbf{A} \cdot \mathbf{F}) \cdot \mathbf{X} = \mathbf{B} \cdot \mathbf{X} = (A^i F_{ij}) X^j = B_j X^j = \text{number}. \end{aligned} \quad (3.3.5)$$

Here we have used the covariant components of \mathbf{F} , and we have set $\mathbf{F} \cdot \mathbf{X} = \mathbf{Z}$ and $\mathbf{A} \cdot \mathbf{F} = \mathbf{B}$. For later reference, we write

$$\mathbf{A} \cdot \mathbf{F} \cdot \mathbf{X} = A^i X^j F_{ij} = X^j A^i F_{ij} = F_{ij} X^j A^i = \text{number}. \quad (3.3.6)$$

Because of the non-commutativity of the vector-tensor dot product, the vectors \mathbf{A} and \mathbf{X} cannot be interchanged, in general. However, since the dot product of vectors is commutative, we have that:

$$\mathbf{A} \cdot \mathbf{F} \cdot \mathbf{X} = (\mathbf{A} \cdot \mathbf{F}) \cdot \mathbf{X} = \mathbf{X} \cdot (\mathbf{A} \cdot \mathbf{F}) \quad (3.3.7a)$$

$$\mathbf{A} \cdot \mathbf{F} \cdot \mathbf{X} = \mathbf{A} \cdot (\mathbf{F} \cdot \mathbf{X}) = (\mathbf{F} \cdot \mathbf{X}) \cdot \mathbf{A} . \quad (3.3.7b)$$

It is convenient to write the consecutive dot product in a more compact way, namely, as the *double dot* product of two tensors. Given two second-order tensors \mathbf{P} and \mathbf{R} , we define the double dot product as:

$$\mathbf{P} : \mathbf{R} \equiv \sum_{ij} P^{ij} R_{ji} \equiv \sum_{ij} P_{ij} R^{ji} \equiv \sum_{ij} P_i^j R_j^i \equiv \mathbf{R} : \mathbf{P} . \quad (3.3.8)$$

Here, $P_i^j R_j^i$ stands for $P_i^j R_j^i$ or $P_{.i}^j R_j^i$. It is important to note that the sequence of indices is in a certain sense symmetric ("reflected" about the position where the double dot would have been).

The literature is not consistent in defining the double product. Most authors use the definition as given in (3.3.8): Krall and Trivelpiece (1973); Bird, Stewart and Lightfoot (1960); Shkarofsky, Johnston and Bachynski (1966); Jancel and Kahan (1966); Chapman and Cowling (1973); Ferziger and Kaper (1972); Milne-Thomson (1968). One important deviation from this convention is used in the standard textbook on classical mechanics by Goldstein (1981), where the tensor indices are "repeated" ($ij|ij$) instead of "reflected" ($ij|ji$). When the indices are "repeated", the position of the "dot" in the mixed-tensor components is different: $P_i^j R_j^i$ or $P_{.i}^j R_j^i$; one "dot" above, and one underneath. Van Bladel (1964) also uses this latter convention. On page 57 of Morse and Feshbach (1953), the double dot is defined as in our (3.3.8), but in (1.6.29) of M&F, the "Goldstein" convention is used.

From (3.3.6), we see that we can write for (3.3.5):

$$\mathbf{A} \cdot \mathbf{F} \cdot \mathbf{X} \equiv \mathbf{X} \mathbf{A} : \mathbf{F} \equiv \mathbf{F} : \mathbf{X} \mathbf{A} . \quad (3.3.9)$$

Observe that the first double-dot expression also resembles the last form in (3.3.7a): the vector \mathbf{A} remains closest to the tensor \mathbf{F} if it operates on the left. Similarly, the last form of (3.3.7b) looks like the last expression in (3.3.9) and keeps the vector \mathbf{X} closest to \mathbf{F} if it is on the right.

For a "simple" tensor or dyad of the form \mathbf{ST} , (3.3.9) leads to:

$$\mathbf{X} \mathbf{A} : \mathbf{ST} = \mathbf{A} \cdot \mathbf{ST} \cdot \mathbf{X} = (\mathbf{A} \cdot \mathbf{S})(\mathbf{T} \cdot \mathbf{X}) \quad (3.3.10a)$$

$$\mathbf{ST} : \mathbf{XA} = \mathbf{T} \cdot \mathbf{XA} \cdot \mathbf{S} = (\mathbf{T} \cdot \mathbf{X})(\mathbf{A} \cdot \mathbf{S}) . \quad (3.3.10b)$$

According to the convention used here, a first dot product is performed between the two vectors next to the double dot symbol; the other dot product is then between the two other vectors. Following the Goldstein (1981) or Van Bladel (1964) convention, we would have instead of (3.3.10) $\mathbf{XA} : \mathbf{ST} = \mathbf{X} \cdot \mathbf{ST} \cdot \mathbf{A} =$

$(\mathbf{X} \cdot \mathbf{S})(\mathbf{T} \cdot \mathbf{A})$, which is clearly a different result because \mathbf{X} and \mathbf{A} have been interchanged.

The above discussion should make clear how single and double dot products of tensors must be interpreted. Generalizations for three or more dot products are readily made on the basis of the index definition of (3.3.8).

The dot product of a vector or a tensor with a dyad involving the vector operator *del*, or the divergence of a tensor can straightforwardly be computed with the rules given so far. A few important rules should be kept in mind, however:

- 1) The *order* of the vectors in a dyad is important; \mathbf{AV} is different from \mathbf{VA} .
- 2) The dot product of vector and dyad or dyadic is *not* commutative.
- 3) the del operator is most conveniently expressed along a contravariant basis: $\nabla \equiv e^i \partial / \partial u^i$.
- 4) The basis vectors are position dependent.

In addition, we mention the obvious facts:

- 5) The del operator operates on everything on its right, except when brackets indicate otherwise.
- 6) The product rule of differentiation is often a very handy tool.

Regarding cross products of vectors and tensors, we again expand each in components along a basis and perform the operations between the basis vectors. To keep the order of the indices right, the Levi-Civita symbol will be helpful. Similar rules apply to expressions of the form $\nabla \times (\text{tensor})$ or $\mathbf{A} \times \nabla \mathbf{B}$, etc., as long as we keep in mind the rules involving the ∇ operator, mentioned above.

3.4 The Relationship Between Covariant, Contravariant and Mixed Components

In (3.3.6), we stated that $\mathbf{A} \cdot \mathbf{F} \cdot \mathbf{X} = A^i X^j F_{ij}$. If we now take the numbers A^i equal to zero, except the i -th component which we set equal to 1, and similarly for the X^j , where we take only the j -th component equal to 1 and the others zero, we have $\mathbf{A} = A^i \mathbf{e}_i \rightarrow \mathbf{e}_i$ and $\mathbf{X} = X^j \mathbf{e}_j \rightarrow \mathbf{e}_j$. This leads to

$$F_{ij} \equiv \mathbf{e}_i \cdot \mathbf{F} \cdot \mathbf{e}_j . \quad (3.4.1)$$

Analogously, we have for the contravariant and mixed components, respectively,

$$F^{kl} \equiv \mathbf{e}^k \cdot \mathbf{F} \cdot \mathbf{e}^l \quad (3.4.2)$$

$$F_m^n \equiv \mathbf{e}_m \cdot \mathbf{F} \cdot \mathbf{e}^n \quad (3.4.3)$$

$$F_{.m}^n \equiv \mathbf{e}^n \cdot \mathbf{F} \cdot \mathbf{e}_m . \quad (3.4.4)$$

Consider now $\mathbf{F} = F_{ij} \mathbf{e}^i \mathbf{e}^j$ and dot multiply on the left by \mathbf{e}^k and on the right by

e^l . Because of (3.4.2), we obtain:

$$\begin{aligned} F^{kl} &= e^k \cdot F \cdot e^l = F_{ij} e^k \cdot e^i e^j \cdot e^l \\ &= F_{ij} (e^k \cdot e^i) (e^j \cdot e^l) , \end{aligned} \quad (3.4.5)$$

which, after using the definition of the reciprocal metric coefficients g^{ij} , becomes

$$F^{kl} = g^{ki} g^{lj} F_{ij} . \quad (3.4.6)$$

By an analogous procedure, we can write:

$$F_i^n = g^{jn} F_{ij} \quad (3.4.7)$$

and

$$F_{ij} = g_{im} g_{jn} F^{mn} . \quad (3.4.8)$$

3.5 Special Tensors

3.5.1 The Kronecker Delta and the Metric Tensor

The Kronecker delta δ_i^j is in fact a second-order tensor I . Strictly, δ_i^j are the mixed components of the *identity tensor* I (also called the “idemfactor”). I is defined by

$$A \cdot I \equiv A \equiv I \cdot A . \quad (3.5.1)$$

An even stronger statement is: δ_i^j are the only legitimate components of the identity tensor (as long as we work in the covariant-contravariant formalism). The symbols δ_{ij} and δ^{ij} , where $\delta_{ij} = \delta^{ij} = 1$ for $i = j$ and 0 for $i \neq j$, which could be thought of as covariant and contravariant components in fact do *not* qualify as such. This can easily be checked by means of the transformation rules of Chap. 13. We can use the symbols δ_{ij} and δ^{ij} to indicate that something is 1 when the two indices coincide and zero when they are different. However, these symbols can only be used in one particular coordinate system. The 1, 0 character of the symbols δ_{ij} and δ^{ij} considered as potential components of a tensor is spoiled when we go to another coordinate system.

A second important second-order tensor is the *metric tensor* g with covariant components the metric coefficients $g_{ij} \equiv e_i \cdot e_j$ and with contravariant components the reciprocal metric coefficients $g^{ij} \equiv e^i \cdot e^j$. The transformation properties show that both sorts of components are legitimate.

We now encounter a very important and interesting observation. Since we work in one single Euclidean space, the *metric tensor* is identical with the *identity tensor*. This particular tensor will be called the *fundamental tensor* (FT). It has the following properties:

$$(FT)_{ij} \equiv g_{ij} \quad (3.5.2a)$$

$$(FT)^{ij} \equiv g^{ij} \quad (3.5.2b)$$

$$(FT)_j^i \equiv \delta_j^i . \quad (3.5.2c)$$

The equivalence can be understood from (3.4.7), where we set F_{ij} on the RHS equal to g_{ij} . We have then

$$g_i^n = g^{jn} g_{ij} . \quad (3.5.3a)$$

On the other hand, we have that $e_i = g_{ij} e^j$. Dot multiplying this expression on both sides by e^n leads to

$$\delta_i^n = g_{ij} g^{jn} . \quad (3.5.3b)$$

Therefore, we have from comparing (3.5.3a) and (3.5.3b):

$$\delta_i^n = g_i^n ; \quad (3.5.4)$$

which says that the mixed components of the metric (or fundamental) tensor are identical to the Kronecker delta. Because I is equivalent to g , there should be a similar relationship to (3.5.1) involving g . We, therefore, propose a formal relationship:

$$A \cdot g = A = g \cdot A , \quad (3.5.5)$$

given that the A in the dot product is written in terms of the “other” components than the A in the middle. Indeed, we have

$$\begin{aligned} A^m e_m \cdot g_{ij} e^i e^j &= A^m g_{ij} e_m \cdot e^i e^j = A^m g_{ij} \delta_m^i e^j \\ &= A^i g_{ij} e^j = A_j e^j = A \end{aligned} \quad (3.5.6a)$$

and

$$\begin{aligned} A_n e^n \cdot g^{ij} e_i e_j &= A_n g^{ij} e^n \cdot e_i e_j = A_n g^{ij} \delta_i^n e_j \\ &= A_i g^{ij} e_j = A^j e_j = A , \end{aligned} \quad (3.5.6b)$$

whereas, for comparison, we can write for (3.5.1):

$$\begin{aligned} A^m e_m \cdot \delta_j^i e_i e^j &= A^m \delta_j^i e_m \cdot e_i e^j = A^m \delta_j^i g_{mi} e^j \\ &= A^m g_{mj} e^j = A_j e^j = A \end{aligned} \quad (3.5.6c)$$

or

$$A_n e^n \cdot \delta_j^i e_i e^j = A_n \delta_j^i e^n \cdot e_i e^j = A_n \delta_j^i \delta_i^n e^j = A_i e^i = A . \quad (3.5.6d)$$

The most convenient form of the fundamental tensor is the “sum-of-dyad” form:

$$\begin{aligned} (FT) \equiv g \equiv I &\equiv e_1 e^1 + e_2 e^2 + e_3 e^3 \equiv e_i e^i \\ &\equiv e^1 e_1 + e^2 e_2 + e^3 e_3 \equiv e^i e_i . \end{aligned} \quad (3.5.7b)$$

From this form, we obtain its coefficients via (3.4.1) through (3.4.4):

$$(\mathbf{FT})_{kl} \equiv \mathbf{e}_k \cdot (\mathbf{FT}) \cdot \mathbf{e}_l = \mathbf{e}_k \cdot \mathbf{e}_i \mathbf{e}^i \cdot \mathbf{e}_l = g_{ki} \delta_l^i \equiv g_{kl} (\equiv \delta_{kl}) \quad (3.5.8a)$$

$$(\mathbf{FT})^{kl} \equiv \mathbf{e}^k \cdot (\mathbf{FT}) \cdot \mathbf{e}^l = \mathbf{e}^k \cdot \mathbf{e}_i \mathbf{e}^i \cdot \mathbf{e}^l = \delta_i^k g^{il} \equiv g^{kl} (\equiv \delta^{kl}) \quad (3.5.8b)$$

$$(\mathbf{FT})_l^k \equiv \mathbf{e}^k \cdot (\mathbf{FT}) \cdot \mathbf{e}_l = \mathbf{e}^k \cdot \mathbf{e}_i \mathbf{e}^i \cdot \mathbf{e}_l = \delta_i^k \delta_l^i \equiv \delta_l^k (\equiv g_l^k) \quad (3.5.8c)$$

The symbols between brackets on the extreme RHS of these expressions may be used if they are interpreted properly, i.e., the co- or contravariant deltas are *not* generally one or zero, and the g_i^k are either one or zero. The same results are obtained if we use $(\mathbf{FT}) \equiv \mathbf{e}^i \mathbf{e}_i$ instead of $\mathbf{e}_i \mathbf{e}^i$.

3.5.2 Levi-Civita and Christoffel Symbols¹

a) Levi-Civita Symbols

It is important to draw attention to the fact that the *Levi-Civita symbols* used before are *not* the components of a third-order tensor. These symbols were defined as

$$\varepsilon_{ijk} = \frac{\mathbf{e}_i \cdot \mathbf{e}_j \times \mathbf{e}_k}{\sqrt{g}} \quad (3.5.9a)$$

and

$$\varepsilon^{ijk} = \sqrt{g} \mathbf{e}^i \cdot \mathbf{e}^j \times \mathbf{e}^k. \quad (3.5.9b)$$

They are +1 for i, j, k equal to an even permutation of 1, 2, 3; -1 for an odd permutation and zero otherwise. If these symbols *were* the components of a tensor, then we would have that the tensor $\vec{\varepsilon}$ could be written as $\vec{\varepsilon} = \varepsilon_{lmn} \mathbf{e}^l \mathbf{e}^m \mathbf{e}^n = \varepsilon^{lmn} \mathbf{e}_l \mathbf{e}_m \mathbf{e}_n$. We now show that this statement is incorrect.

In analogy with (3.4.2), we can write for the contravariant components of a third-order tensor \vec{F} :

$$F^{ijk} \equiv \mathbf{e}^j \cdot (\mathbf{e}^i \cdot \vec{F} \cdot \mathbf{e}^k) \equiv (\mathbf{e}^i \cdot \vec{F} \cdot \mathbf{e}^k) \cdot \mathbf{e}^j. \quad (3.5.10)$$

Between the brackets we need the two outer indices of F^{ijk} , here i and k in that order. Since the quantity between the brackets is a vector, commutativity applies, and the position of \mathbf{e}^j can be either of those indicated. Similarly, we have that

$$F_{ijk} \equiv \mathbf{e}_j \cdot (\mathbf{e}_i \cdot \vec{F} \cdot \mathbf{e}_k) \equiv (\mathbf{e}_i \cdot \vec{F} \cdot \mathbf{e}_k) \cdot \mathbf{e}_j. \quad (3.5.11)$$

According to (3.5.10), we would have for the contravariant components of $\vec{\varepsilon} = \varepsilon^{lmn} \mathbf{e}_l \mathbf{e}_m \mathbf{e}_n$:

$$(\varepsilon)^{ijk} \equiv \varepsilon^{lmn} \mathbf{e}^j \cdot (\mathbf{e}^i \cdot \mathbf{e}_l \mathbf{e}_m \mathbf{e}_n \cdot \mathbf{e}^k) = \varepsilon^{lmn} \delta_l^i \delta_m^k \mathbf{e}^j \cdot \mathbf{e}_n$$

$$= \varepsilon^{lmn} \delta_l^i \delta_m^k \delta_n^j = \varepsilon^{ijk} \quad (3.5.12a)$$

which is what we expect, being equal to 1, -1, or zero. However, if we try to compute the contravariant components from the form $\vec{\varepsilon} = \varepsilon_{lmn} \mathbf{e}^l \mathbf{e}^m \mathbf{e}^n$, we get instead

$$\begin{aligned} (\varepsilon)^{ijk} &= \varepsilon_{lmn} \mathbf{e}^j \cdot (\mathbf{e}^i \cdot \mathbf{e}^l \mathbf{e}^m \mathbf{e}^n \cdot \mathbf{e}^k) = \varepsilon_{lmn} g^{il} g^{nk} \mathbf{e}^j \cdot \mathbf{e}^m \\ &= \varepsilon_{lmn} g^{il} g^{nk} g^{jm} = \varepsilon^{ijk} g^{-1} = \frac{\varepsilon^{ijk}}{g}, \end{aligned} \quad (3.5.12b)$$

which is different from (3.5.12a). Here, $g = \det(g_{ij})$ and $\det(g^{ij}) = g^{-1}$. We used the fact that $\sum_{lmn} g^{il} g^{jm} g^{kn} \varepsilon_{lmn}$, with ε_{lmn} the Levi-Civita symbol, is equal to the determinant g for i, j, k an even permutation of 1, 2, 3; $-g$ for i, j, k an odd permutation, and zero if any two indices are equal. Hence, $\vec{\varepsilon}$ cannot be a tensor with components ε^{ijk} or ε_{ijk} . Similar results are obtained for covariant components $(\varepsilon)_{ijk}$. Note that one should not convert $\varepsilon_{lmn} g^{il} g^{jm} g^{kn}$ into ε^{ijk} in analogy to (3.4.6). This would lead to the wrong result because the ε_{lmn} is a symbol equal to +1, -1 or 0 and is not a covariant component of $\vec{\varepsilon}$. In Chap. 13, an alternative proof for the non-tensor character of ε_{ijk} and ε^{ijk} is given via transformation equations.

One could define a tensor called the *permutation tensor*, which is closely related to the Levi-Civita symbol:

$$\vec{E} \equiv \frac{\varepsilon^{lmn}}{\sqrt{g}} \mathbf{e}_l \mathbf{e}_m \mathbf{e}_n \quad (3.5.13a)$$

or

$$\vec{E} \equiv \varepsilon_{lmn} \sqrt{g} \mathbf{e}^l \mathbf{e}^m \mathbf{e}^n. \quad (3.5.13b)$$

As in (3.5.12a) and (3.5.12b), we have

$$E^{ijk} = \frac{\varepsilon^{lmn}}{\sqrt{g}} \delta_l^i \delta_m^j \delta_n^k = \frac{\varepsilon^{ijk}}{\sqrt{g}} \quad (3.5.14a)$$

or

$$E^{ijk} = \varepsilon_{lmn} \sqrt{g} g^{il} g^{jm} g^{kn} = \varepsilon^{ijk} \sqrt{g} g^{-1} = \frac{\varepsilon^{ijk}}{\sqrt{g}} \quad (3.5.14b)$$

and

$$E_{ijk} = \frac{\varepsilon^{lmn}}{\sqrt{g}} g_{il} g_{jm} g_{kn} = \frac{\varepsilon_{ijk}}{\sqrt{g}} g = \varepsilon_{ijk} \sqrt{g} \quad (3.5.14c)$$

or

¹ Can be skipped without loss of continuity.

$$E_{ijk} = \epsilon_{lmn} \sqrt{g} \delta_i^l \delta_j^m \delta_k^n = \epsilon_{ijk} \sqrt{g} . \quad (3.5.14d)$$

So we have agreement: ϵ^{lmn}/\sqrt{g} and $\epsilon_{lmn}\sqrt{g}$ are respectively the contravariant and covariant components of \vec{E} . In comparison with (3.5.9), we have for the components of \vec{E} :

$$E_{ijk} = e_i \cdot e_j \times e_k = \begin{cases} +\sqrt{g} \\ -\sqrt{g} \\ 0 \end{cases} \quad (3.5.15a)$$

and

$$E^{ijk} = e^i \cdot e^j \times e^k = \begin{cases} +1/\sqrt{g} \\ -1/\sqrt{g} \\ 0 \end{cases} . \quad (3.5.15b)$$

The transformation properties of E_{ijk} and E^{ijk} are given in Chap. 13.

b) Christoffel Symbols

The Christoffel symbols discussed in Chap. 2 are *not* the components of a third-order tensor as the alternative Γ -notation might suggest. Recall that the Γ -notation identified Γ_{jk}^i with $\{_j^i_k\}$ and $\Gamma_{i,jk}$ with $[i,jk]$. Therefore, the Γ notation should be avoided. The easiest way to prove that $\{_j^i_k\}$ and $[i,jk]$ are not tensor components is via the transformation properties. This is discussed in Chap. 13.

3.6 Tensor and Dyadic Identities

In this section we give a summary of the most important identities involving vectors and second-order tensors. Most of these formulae can be found in the more pragmatic texts mentioned in the next section. However, this collection is mainly based on Shkarofsky, Johnston and Bachynski (1966); Bird, Stewart and Lightfoot (1960) and Wimmel (1982). The last is especially recommended for its tensor formulae.

$$\mathbf{I} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{I} = \mathbf{A} \quad (T.1)$$

$$\mathbf{I} \cdot \mathbf{AB} = \mathbf{AB} \quad (T.2)$$

$$\mathbf{I} \cdot \mathbf{VB} = \mathbf{VB} \quad (T.3)$$

$$\mathbf{AB} \cdot \mathbf{CD} = (\mathbf{B} \cdot \mathbf{C})(\mathbf{A} \cdot \mathbf{D}) = \mathbf{D} \cdot \mathbf{AB} \cdot \mathbf{C} \quad (T.4)$$

$$\mathbf{F} \cdot \mathbf{AB} = (\mathbf{F} \cdot \mathbf{A}) \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{F} \cdot \mathbf{A} \quad (T.5)$$

$$\mathbf{I} \cdot \mathbf{AB} = \mathbf{A} \cdot \mathbf{B} \quad (T.6)$$

$$\mathbf{I} \cdot \mathbf{VB} = \mathbf{V} \cdot \mathbf{B} \quad (T.7)$$

$$\mathbf{I} \cdot \mathbf{F} = F_i^i = \text{tr } \mathbf{F} \quad (T.8)$$

$$\mathbf{V} \cdot (\mathbf{aI}) = \mathbf{Va} \quad (T.9)$$

$$\mathbf{V} \cdot \mathbf{AB} = \mathbf{A} \cdot \mathbf{VB} + (\mathbf{V} \cdot \mathbf{A})\mathbf{B} \quad (T.10)$$

$$\mathbf{I} \times \mathbf{A} = \mathbf{A} \times \mathbf{I} \quad (T.11)$$

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{I} = \mathbf{I} \times (\mathbf{A} \times \mathbf{B}) = \mathbf{BA} - \mathbf{AB} \quad (T.12)$$

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = \mathbf{C} \times (\mathbf{B} \times \mathbf{A}) = \mathbf{B}(\mathbf{C} \cdot \mathbf{A}) - \mathbf{A}(\mathbf{C} \cdot \mathbf{B}) = \mathbf{C} \cdot (\mathbf{AB} - \mathbf{BA}) \quad (T.13)$$

$$\mathbf{A} \times (\mathbf{BC}) = (\mathbf{A} \times \mathbf{B})\mathbf{C} \quad (T.14)$$

$$\mathbf{V} \cdot (\mathbf{I} \times \mathbf{C}) = \mathbf{V} \times \mathbf{C} \quad (T.15)$$

$$\mathbf{V} \times (\mathbf{A} \times \mathbf{B}) = \mathbf{V} \cdot (\mathbf{BA} - \mathbf{AB}) \quad (T.16)$$

$$\mathbf{A} \times (\mathbf{V} \times \mathbf{B}) = (\mathbf{VB}) \cdot \mathbf{A} - \mathbf{A} \cdot \mathbf{VB} \quad (T.17)$$

$$\begin{aligned} \mathbf{V}(\mathbf{A} \cdot \mathbf{B}) &= \mathbf{A} \times (\mathbf{V} \times \mathbf{B}) + \mathbf{B} \times (\mathbf{V} \times \mathbf{A}) + \mathbf{A} \cdot \mathbf{VB} + \mathbf{B} \cdot \mathbf{VA} \\ &= (\mathbf{VA}) \cdot \mathbf{B} + (\mathbf{VB}) \cdot \mathbf{A} \end{aligned} \quad (T.18)$$

$$\mathbf{V}(\mathbf{A} \times \mathbf{B}) = \mathbf{VA} \times \mathbf{B} - \mathbf{VB} \times \mathbf{A} \quad (T.19)$$

$$\mathbf{V} \times (\mathbf{AB}) = (\mathbf{V} \times \mathbf{A})\mathbf{B} - \mathbf{A} \times \mathbf{VB} \quad (T.20)$$

$$\mathbf{A} \times \mathbf{VB} - (\mathbf{A} \times \mathbf{VB})^T = [(\mathbf{V} \cdot \mathbf{B})\mathbf{A} - \mathbf{VB} \cdot \mathbf{A}] \times \mathbf{I} \quad (T.21)$$

$$\mathbf{VB} \times \mathbf{A} - (\mathbf{VB} \times \mathbf{A})^T = [(\mathbf{V} \cdot \mathbf{B})\mathbf{A} - \mathbf{A} \cdot \mathbf{VB}] \times \mathbf{I} \quad (T.22)$$

$$(\mathbf{V} \cdot \mathbf{B})\mathbf{A} - \mathbf{VB} \cdot \mathbf{A} = -(\mathbf{A} \times \mathbf{V}) \times \mathbf{B} \quad (T.23)$$

$$(\mathbf{V} \cdot \mathbf{B})\mathbf{A} - \mathbf{A} \cdot \mathbf{VB} = \mathbf{A} \times (\mathbf{V} \times \mathbf{B}) - (\mathbf{A} \times \mathbf{V}) \times \mathbf{B} \quad (T.24)$$

$$(\mathbf{A} \times \mathbf{V}) \cdot \mathbf{B} = \mathbf{A} \cdot (\mathbf{V} \times \mathbf{B}) \quad (T.25)$$

$$\mathbf{A} \times \mathbf{VB} + (\mathbf{VB} \times \mathbf{A})^T = [\mathbf{A} \cdot (\mathbf{V} \times \mathbf{B})]\mathbf{I} - (\mathbf{V} \times \mathbf{B})\mathbf{A} \quad (T.26)$$

$$\mathbf{VB} \times \mathbf{A} + (\mathbf{A} \times \mathbf{VB})^T = [\mathbf{A} \cdot (\mathbf{V} \times \mathbf{B})]\mathbf{I} - \mathbf{A}(\mathbf{V} \times \mathbf{B}) \quad (T.27)$$

$$\begin{aligned} \mathbf{A} \times \mathbf{VB} + \mathbf{VB} \times \mathbf{A} &= \mathbf{I} \times [(\mathbf{V} \cdot \mathbf{B})\mathbf{A} - \mathbf{VB} \cdot \mathbf{A}] \\ &\quad + [\mathbf{A} \cdot (\mathbf{V} \times \mathbf{B})]\mathbf{I} - \mathbf{A}(\mathbf{V} \times \mathbf{B}) \\ &= \mathbf{I} \times [(\mathbf{V} \cdot \mathbf{B})\mathbf{A} - \mathbf{A} \cdot \mathbf{VB}] \\ &\quad + [\mathbf{A} \cdot (\mathbf{V} \times \mathbf{B})]\mathbf{I} - (\mathbf{V} \times \mathbf{B})\mathbf{A} \end{aligned} \quad (T.28)$$

$$(\mathbf{A} \times \mathbf{C}) \cdot \mathbf{F} = \mathbf{A} \cdot (\mathbf{C} \times \mathbf{F}) = -\mathbf{C} \cdot (\mathbf{A} \times \mathbf{F}) \quad (T.29)$$

$$\mathbf{F} \cdot (\mathbf{A} \times \mathbf{C}) = (\mathbf{F} \times \mathbf{A}) \cdot \mathbf{C} = -(\mathbf{F} \times \mathbf{C}) \cdot \mathbf{A} \quad (T.30)$$

$$\mathbf{A} \cdot \mathbf{F} \times \mathbf{C} = (\mathbf{A} \cdot \mathbf{F}) \times \mathbf{C} = \mathbf{A} \cdot (\mathbf{F} \times \mathbf{C}) = -\mathbf{C} \times (\mathbf{A} \cdot \mathbf{F}) \quad (T.31)$$

$$\mathbf{A} \times \mathbf{F} \cdot \mathbf{C} = \mathbf{A} \times (\mathbf{F} \cdot \mathbf{C}) = (\mathbf{A} \times \mathbf{F}) \cdot \mathbf{C} = -(\mathbf{F} \cdot \mathbf{C}) \times \mathbf{A} \quad (T.32)$$

$$\mathbf{A} \cdot \nabla \mathbf{B} \times \mathbf{C} + \mathbf{C} \times \nabla \mathbf{B} \cdot \mathbf{A} = \mathbf{C} \times [\mathbf{A} \times (\nabla \times \mathbf{B})] \quad (T.33)$$

$$\mathbf{A} \cdot \nabla \mathbf{B} \times \mathbf{C} - \mathbf{C} \cdot \nabla \mathbf{B} \times \mathbf{A} = [(\nabla \cdot \mathbf{B}) \mathbf{I} - \nabla \mathbf{B}] \cdot (\mathbf{A} \times \mathbf{C}) \quad (T.34)$$

$$\mathbf{A} \times \nabla \mathbf{B} \cdot \mathbf{C} - \mathbf{C} \times \nabla \mathbf{B} \cdot \mathbf{A} = (\mathbf{A} \times \mathbf{C}) \cdot [(\nabla \cdot \mathbf{B}) \mathbf{I} - \nabla \mathbf{B}] \quad (T.35)$$

$$\mathbf{A} \cdot \nabla \mathbf{B} \cdot \mathbf{C} - \mathbf{C} \cdot \nabla \mathbf{B} \cdot \mathbf{A} = (\mathbf{A} \times \mathbf{C}) \cdot (\nabla \times \mathbf{B}) \quad (T.36)$$

$$\nabla \mathbf{R} = \mathbf{I}, \quad \mathbf{R} = \text{position vector} . \quad (T.37)$$

3.7 Suggestions for Further Reading

This chapter was a minimal discussion of tensors. For more on this powerful “technique”, we must refer to the vast literature in mathematics and mathematical physics. It is difficult to suggest one single reference because none of those known to us treats tensors correctly *and* generally enough *and* appropriately for plasma physics.

A substantial number of physics textbooks devote a chapter or an appendix to tensors or polyadics. These books cover fields like plasma physics, mechanics, electromagnetics, transport, elasticity theory and fluid dynamics. In particular, we would like to mention Symon (1971); Goldstein (1981); Bird, Stewart and Lightfoot (1960); Milne-Thomson (1968); Stratton (1941); Van Bladel (1964); Krall and Trivelpiece (1973); Seshadri (1973); Shkarofsky, Johnston and Bachynski (1966); Chapman and Cowling (1973); Ferziger and Kaper (1972); Jancel and Kahan (1966); Woodson and Melcher (1968).

In the more practical mathematics and mathematical-physics books, tensors are often defined by means of their transformation laws; in some of them, dyads and dyadics are discussed (or parenthetically referred to). Dual spaces are not mentioned. In this category, we recommend Wrede (1963); Borisenko and Tarapov (1968); Morse and Feshbach (1953); Mathews and Walker (1970); Margenau and Murphy (1976); Butkov (1968); Spiegel (1959); Menzel (1961); Hoffmann (1977).

On the subject of tensor calculus in the context of differential geometry and its relationship with differential forms and Grassmann algebra, see Stoker (1969); Guggenheim (1977); Spivak (1970); Choquet-Bruhat (1982); Behnke (1974); Flanders (1963); Arnold (1978); Abraham and Marsden (1978); Deschamps (1970); Schutz (1980); Misner, Thorne and Wheeler (1973); Burke (1985).

4. Magnetic-Field-Structure-Related Concepts

4.1 The Equation of a Magnetic-Field Line

A magnetic field line (or “line of force”) is by definition a curve whose tangent at any point is parallel to the magnetic field-vector \mathbf{B} . This property can be translated into mathematical terms as follows:

$$\mathbf{B} \propto d\mathbf{R} \quad \text{or} \quad \mathbf{B} = c \, d\mathbf{R} , \quad (4.1.1)$$

where c is a proportionality constant and $d\mathbf{R}$ is a differential vector tangent to the field line (\mathbf{R} is the position vector from a chosen origin to the point in question on the curve). Equivalently,

$$\mathbf{B} \times d\mathbf{R} = 0 . \quad (4.1.2)$$

Using the contravariant components of both $\mathbf{B}(B^1, B^2, B^3)$ and $d\mathbf{R}(du^1, du^2, du^3)$, either of the above equations leads to:

$$\frac{B^1}{du^1} = \frac{B^2}{du^2} = \frac{B^3}{du^3} = c . \quad (4.1.3)$$

If we recall that $B^i = \mathbf{B} \cdot \nabla u^i$, we can also write:

$$\frac{\mathbf{B} \cdot \nabla u^1}{du^1} = \frac{\mathbf{B} \cdot \nabla u^2}{du^2} = \frac{\mathbf{B} \cdot \nabla u^3}{du^3} . \quad (4.1.4)$$

If we parameterize the field-line curve by the arc length l , the tangent vector $d\mathbf{R}/dl$ is a unit vector, commonly denoted by $\hat{\mathbf{t}}$ or $\hat{\mathbf{B}} \equiv \mathbf{B}/B$:

$$\frac{\mathbf{B}}{B} = \frac{d\mathbf{R}}{dl} \Big|_{\text{along } \mathbf{B}} . \quad (4.1.5)$$

From (4.1.1), then, the constant c equals B/dl , so that the *equation of a magnetic-field line* is, in component form:

$$\frac{B}{dl} = \frac{B^1}{du^1} = \frac{B^2}{du^2} = \frac{B^3}{du^3} \quad (4.1.6)$$

or

$$\frac{B}{dl} = \frac{\mathbf{B} \cdot \nabla u^1}{du^1} = \frac{\mathbf{B} \cdot \nabla u^2}{du^2} = \frac{\mathbf{B} \cdot \nabla u^3}{du^3}. \quad (4.1.7)$$

The form of (4.1.1) implies that a choice for contravariant components of $d\mathbf{R}$, $du^i \equiv d\mathbf{R} \cdot \nabla u^i$, requires the use of the contravariant components of \mathbf{B} , the B^i .

4.2 The Frozen-Flux Theorem

The frozen-flux theorem states that the magnetic flux threading an open surface in a perfectly conducting (MHD) plasma is constant. Even when the surface moves with the plasma, the flux remains the same. This is equivalent to saying that the flux is frozen within the plasma or that the field lines move with the plasma. This theorem is proven in most plasma-physics textbooks. We will follow closely the concise but clear discussion given by Schmidt (1979).

The starting point to prove the theorem is a combination of Ampère's law, Faraday's law and Ohm's law:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \quad (4.2.1)$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} \quad (4.2.2)$$

$$\mathbf{J} = \sigma(\mathbf{E} + \mathbf{V} \times \mathbf{B}). \quad (4.2.3)$$

Here, \mathbf{V} represents the plasma-fluid velocity, σ is the electrical conductivity, and the other symbols have the usual meanings. For consistency with MHD, we neglected the displacement current in Ampère's law, (4.2.1). This statement is equivalent to $\mathbf{V} \cdot \mathbf{J} = 0$. This is a particular version of charge conservation. The equations are written in MKS units.

After substituting (4.2.1) for \mathbf{J} into Ohm's law, (4.2.3), and solving for the electric field \mathbf{E} , which in turn is substituted into (4.2.2), we obtain an equation for $\partial \mathbf{B} / \partial t$:

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \left[\frac{\nabla \times \mathbf{B}}{\mu_0 \sigma} - \mathbf{V} \times \mathbf{B} \right]. \quad (4.2.4)$$

With the help of the vector identity $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$, this becomes

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{B} - \nabla \left(\frac{1}{\mu_0 \sigma} \right) \times (\nabla \times \mathbf{B}) + \nabla \times (\mathbf{V} \times \mathbf{B}), \quad (4.2.5)$$

since $\mathbf{V} \cdot \mathbf{B} = 0$. Equation (4.2.5) is a diffusion equation for the magnetic field in a moving plasma (see, e.g., Schmidt, 1979, or Golant, Zhilinsky and Sakharov, 1980).

We are interested in the perfectly conducting limit, where $\sigma \rightarrow \infty$:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{V} \times \mathbf{B}). \quad (4.2.6)$$

We now integrate this equation over an arbitrary open but fixed surface S inside the plasma. With Stokes' law, we obtain

$$\iint_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} = \iint_S \nabla \times (\mathbf{V} \times \mathbf{B}) \cdot d\mathbf{S} = \oint_{\Gamma} \mathbf{V} \times \mathbf{B} \cdot dl. \quad (4.2.7)$$

dl represents a differential length element along the boundary curve Γ of S (see Fig. 4.1). The last form of (4.2.7) can be rewritten using the triple-product identity $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = -\mathbf{c} \cdot (\mathbf{b} \times \mathbf{a})$, while the $\partial / \partial t$ operator can be moved outside the first integral. The surface S is stationary, and $\partial / \partial t$ is taken at constant \mathbf{R} . With the magnetic flux Ψ through S defined as usual, $\Psi = \iint S \mathbf{B} \cdot d\mathbf{S}$, we can write

$$\frac{\partial \Psi}{\partial t} + \oint_{\Gamma} \mathbf{B} \cdot (\mathbf{V} \times dl) \equiv 0. \quad (4.2.8)$$

To interpret this integral relation, we consider Fig. 4.1. The open surface S is fixed in space, regardless of the plasma velocity. Since the velocity \mathbf{V} represents the plasma velocity, it is clear that the open surface S' is the equivalent of surface S , but now "carried" with the plasma. Between the open surfaces S (at time $t = 0$), and S' (at time $t = \Delta t$), there is a "side"-surface created by the moving boundary Γ' . A differential area element of this side-surface is $|\mathbf{V} \times dl| / \Delta t$ (cross-hatched in the figure). The first term of (4.2.8) is the time rate of change of the magnetic flux through the fixed open surface S . The second term represents the increment of flux through the side-surface swept out by the moving boundary Γ' . Because \mathbf{B} is a solenoidal field, we know that the total flux through the closed surface $S + S_{\text{side}} + S'$ must be zero. Since (4.2.8) states that the change in flux through $S + S_{\text{side}}$ is zero, we must have that the change of flux through the moving surface S' is zero as well. This implies that the total time derivative of the flux through a (moving) surface (such as S') must be zero (see, e.g., Jackson, 1975, Chap. 6):

$$\frac{d\Psi}{dt} \equiv 0. \quad (4.2.9)$$

The surface S (or S') in the above proof was arbitrary. It follows that magnetic-field lines are tied to the plasma, which makes the abstract concept of a field line more concrete in an ideal MHD plasma. This allows us to speak about the motion of field lines (see Schmidt 1979). The motion of magnetic-field lines has been studied by Newcomb (1958).

The notion of moving field lines is used throughout fusion-plasma physics, even when the fluid is not perfectly conducting, and the plasma and the field lines are no longer tied together. However, we must be careful with the motion of field

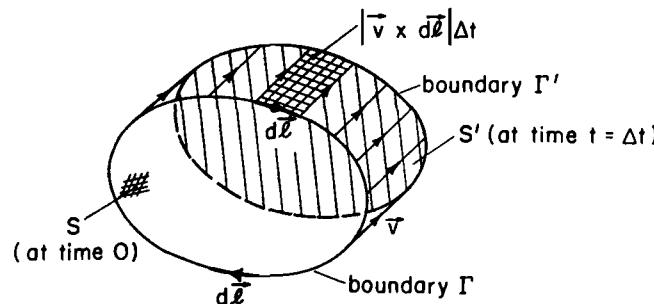


Fig. 4.1. The motion of a boundary curve Γ , which is tied to the plasma. See text

lines when a parallel electric field is present (Newcomb, 1958; Kerst, private communication). Perpendicular E fields lead to perpendicular $E \times B$ fluid motion and do not give rise to problems.

4.3 The Magnetic Field-Line Curvature

Consider a magnetic-field line parameterized by its arc length l . The vector dR/dl , where R is the position vector of a point on the curve, is a unit tangent vector. It is denoted by \hat{t} or $\hat{\ell}$ or $\hat{B} \equiv B/B$. (We assume for convenience that the arc length which parameterizes the field line increases in the direction in which B points. Otherwise, we would have that $\hat{B} = -\hat{l}$.) The dot product of \hat{t} with itself is unity, $\hat{t} \cdot \hat{t} = 1$, so that $\hat{t} \cdot d\hat{t}/dl = 0$. Thus the vector $d\hat{t}/dl$ is perpendicular to the tangent vector \hat{t} . We now define:

$$\frac{d\hat{t}}{dl} \Big|_{\text{along } \hat{t}} \equiv \kappa \hat{n} \equiv \frac{1}{R_c} \hat{n} \equiv \kappa \equiv \frac{d\hat{B}}{dl} \Big|_{\text{along } \hat{B}} . \quad (4.3.1)$$

Here, \hat{n} is the principal normal to the field line. The parameter κ is called the *curvature*, and the vector κ the *curvature vector*. The quantity $R_c = 1/\kappa$ is the *local radius of curvature*. That (4.3.1) is in agreement with intuition can be shown geometrically from the definition:

$$\kappa = \lim_{\Delta l \rightarrow 0} [\hat{t}(l + \Delta l) - \hat{t}(l)]/\Delta l = \lim_{\Delta l \rightarrow 0} \frac{d\hat{t}}{dl} .$$

Sometimes one also uses the *radius-of-curvature vector* R_c , pointing in the negative \hat{n} direction. It is defined by

$$\kappa = -\frac{R_c}{R_c^2} . \quad (4.3.2)$$

The derivative $d/dl|_{\text{along } \hat{B}}$ is the directional derivative and is equal to the dot product of \hat{B} with ∇ :

$$\hat{B} \cdot \nabla \equiv \frac{\partial}{\partial l} = \frac{d}{dl} \Big|_{\text{along } \hat{B}} . \quad (4.3.3)$$

From (4.3.1) and (4.3.3), we can write the curvature in the form most commonly used in plasma physics:

$$\kappa \equiv (\hat{B} \cdot \nabla) \hat{B} = \hat{B} \cdot \nabla \hat{B} . \quad (4.3.4)$$

The second form reminds us that it is basically the dot product of a vector \hat{B} with a dyad $\nabla \hat{B}$. Note that the dot-product rule from Chap. 3, on dyads, is apparently “violated”. The result of $A \cdot (CD) = (A \cdot C)D$ is normally a vector proportional to the vector not affected by the dot, here D . Because of the special character of the operator ∇ , $(\hat{B} \cdot \nabla) \hat{B}$ is not along the vector \hat{B} , but orthogonal to \hat{B} according to (4.3.1).

An alternative expression for κ can be found by applying the *bac-cab* rule to $\hat{B} \times (\nabla \times \hat{B})$, which yields $\hat{B} \times (\nabla \times \hat{B}) = \nabla(\hat{B} \cdot \hat{B})/2 - (\hat{B} \cdot \nabla)\hat{B} = -(\hat{B} \cdot \nabla)\hat{B}$. Hence, we obtain

$$\kappa = \hat{B} \cdot \nabla \hat{B} = -\hat{B} \times (\nabla \times \hat{B}) . \quad (4.3.5)$$

There is yet another commonly used expression for the curvature vector, where it is related to the gradient of $|B|$. However, this expression is only valid for low β plasmas. With J set equal to $\nabla \times B/\mu_0$, and with the help of the vector identity $A \times (\nabla \times B) = (\nabla B) \cdot A - A \cdot \nabla B$, we obtain from the equilibrium equation $\nabla p = J \times B$:

$$\nabla p = \frac{1}{\mu_0} (\nabla \times B) \times B = \frac{1}{\mu_0} (B \cdot \nabla) B - \frac{1}{2\mu_0} \nabla(B)^2 . \quad (4.3.6)$$

The “2” in $(B)^2$ represents a square, not a superscript. When the first term on the right-hand side is written as $B = B\hat{B}$, it becomes equal to

$$\frac{B^2}{\mu_0} (\hat{B} \cdot \nabla) \hat{B} + \frac{B\hat{B}}{\mu_0} \cdot (\nabla B) \hat{B} . \quad (4.3.7)$$

Recall that the order of the dyad in this second term must be kept, first ∇ and then \hat{B} . Then

$$\nabla p = \frac{B^2}{\mu_0} \kappa + \hat{B} \left(\hat{B} \cdot \nabla \frac{B^2}{2\mu_0} \right) - \nabla \frac{B^2}{2\mu_0} . \quad (4.3.8)$$

We define the parallel and perpendicular (to \hat{B}) ∇ operators as

$$\nabla_{\parallel} = \hat{B}(\hat{B} \cdot \nabla) \text{ and } \nabla_{\perp} = \nabla - \nabla_{\parallel} = \nabla - \hat{B}(\hat{B} \cdot \nabla) = -\hat{B} \times (\hat{B} \times \nabla) ,$$

respectively. Since $\hat{B} \cdot \nabla p = 0$, we see that $\nabla p = \nabla_{\perp} p$, so that (4.3.8) can be written

as

$$\nabla_{\perp} \left(p + \frac{B^2}{2\mu_0} \right) = \frac{B^2}{\mu_0} \kappa . \quad (4.3.9)$$

In low-pressure plasmas, when $\beta \equiv p/(B^2/2\mu_0) \ll 1$, we can neglect the plasma pressure p in comparison with the magnetic pressure $B^2/2\mu_0$, and obtain

$$\nabla_{\perp} \frac{B^2}{2} \simeq B^2 \kappa . \quad (4.3.10)$$

Alternatively,

$$\kappa \simeq \frac{\nabla_{\perp} B}{B} = \nabla_{\perp} \ln B, \text{ for } \beta \ll 1 . \quad (4.3.11)$$

This expression shows that the curvature (in low β plasmas) is effectively the reciprocal of the scale length for (perpendicular) variation of $|B|$.

So far, we have been dealing with the curvature of a curve (e.g., a magnetic-field line) in isolation. When the curve under consideration lies within a surface, it is customary to define curvature components with respect to the surface. The component of κ normal to the surface is called the *normal curvature*, whereas the component of κ tangent to the surface is referred to as the *geodesic curvature* (Goetz 1970; Guggenheim 1977).

As we will discuss below, in magnetic fusion the magnetic-field lines are often chosen to be coordinate curves. As indicated on Fig. 4.2, we identify the

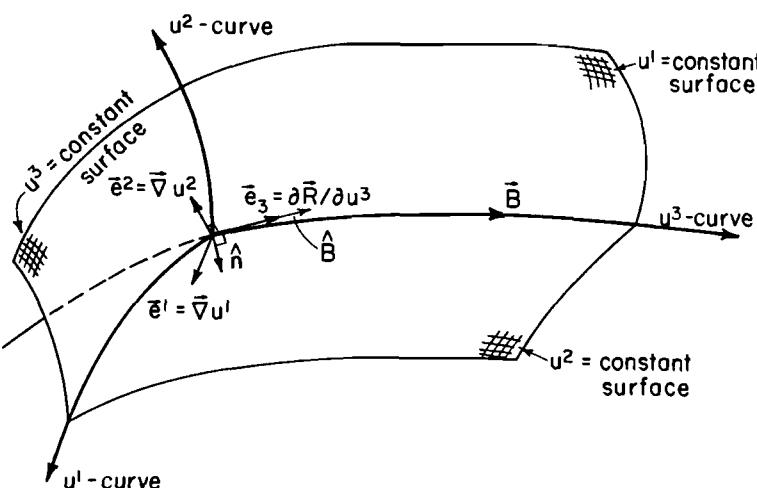


Fig. 4.2. The basis vectors e^1 , e^2 , and e_3 in relation to the principal normal to the u^3 -curve. The vectors \hat{n} , $\nabla u^1 = e^1$ and $\nabla u^2 = e^2$ are perpendicular to \vec{B} and lie in the same plane

third coordinate curve, the u^3 -curve with a magnetic-field line. This lies at the intersection of the coordinate surfaces $u^1 = \text{constant}$ and $u^2 = \text{constant}$. By construction, the tangent-basis vector e_3 (which is proportional to \vec{B}) is perpendicular to the reciprocal-basis vectors $e^1 = \nabla u^1$ and $e^2 = \nabla u^2$. To make the discussion somewhat less abstract, one could take u^3 as the length l along the field line.

“Natural” surfaces to which magnetic-field lines belong are magnetic-flux surfaces (which will be defined rigorously below). We could think of the surfaces $u^1 = \text{constant}$ as flux surfaces. For some applications in fusion-plasma physics, it is desirable to know the components of the curvature κ perpendicular and tangent to the flux surface $u^1 = \text{constant}$. The curvature lies along the principal normal, \hat{n} , to the field line. The vectors, \hat{n} , ∇u^1 and ∇u^2 are perpendicular to \vec{B} and lie in the same plane as each other.

We want to decompose \hat{n} into a component along ∇u^1 (i.e., a component perpendicular to the (flux) surface $u^1 = \text{constant}$), and a component tangent to the $u^1 = \text{constant}$ surface. ∇u^2 is (in general) not tangential to the $u^1 = \text{constant}$ surface (see Fig. 4.2). Also, the tangent-basis vector, $e_2 = \partial R / \partial u^2$, which is tangential to the $u^1 = \text{constant}$ surface, is not acceptable for our purposes, since it has a component along \vec{B} (in general, tangent-basis vectors are not orthogonal to each other, $e_i \cdot e_j \neq 0$). This is equivalent to saying that e_2 does not lie in the plane to which κ and ∇u^1 belong. As mentioned above, ∇u^2 does lie in the plane perpendicular to \vec{B} , and all we must do is find the component of ∇u^2 parallel to the ∇u^1 direction and subtract it from ∇u^2 . Then we have a vector both tangent to the flux surface and perpendicular to \vec{B} . The parallel and perpendicular components of the vector ∇u^2 with respect to a vector normal to the $u^1 = \text{constant}$ surface, $\hat{N} \equiv \nabla u^1 / |\nabla u^1|$ are given by

$$(\nabla u^2)_{\parallel} = \hat{N} (\hat{N} \cdot \nabla u^2) \quad (4.3.12a)$$

$$(\nabla u^2)_{\perp} = -\hat{N} \times (\hat{N} \times \nabla u^2) . \quad (4.3.12b)$$

When the curvature vector κ is expanded in its covariant components as usual, we have

$$\kappa = \kappa_1 \nabla u^1 + \kappa_2 \nabla u^2 ; \quad \kappa_i \equiv \kappa \cdot e_i , \quad (4.3.13)$$

since $\kappa_3 \equiv \kappa \cdot e_3 = 0$, recalling that $e_3 \propto \vec{B}$. After setting ∇u^2 in (4.3.13) equal to $(\nabla u^2)_{\parallel} + (\nabla u^2)_{\perp}$ of (4.3.12), we obtain:

$$\kappa = [\kappa_1 |\nabla u^1| + \kappa_2 (\hat{N} \cdot \nabla u^2)] \hat{N} + \kappa_2 (\hat{N} \times \nabla u^2) \times \hat{N} . \quad (4.3.14)$$

The first term is the *normal-curvature vector* (along the normal \hat{N} to the surface $u^1 = \text{constant}$). The second term is the *geodesic-curvature vector*. In summary, we have

$$\kappa_N = \kappa_1 |\nabla u^1| + \kappa_2 |\nabla u^2| \cos \varphi (\nabla u^1, \nabla u^2) , \quad \text{normal curvature} \quad (4.3.15a)$$

$$\kappa_G = \kappa_2 |\nabla u^2| \sin \varphi (\nabla u^1, \nabla u^2) , \quad \text{geodesic curvature} . \quad (4.3.15b)$$

These are the correct definitions of normal and geodesic curvature. Sometimes, however, this terminology is used loosely, and κ_1 in (4.3.13) is called the normal curvature, while κ_2 is designated as the geodesic curvature. In magnetic geometries with a high degree of symmetry, these distinctions become irrelevant.

Below we will identify u^1 with a radial variable and represent it by symbols such as ρ , Ψ , or α . u^2 will be considered as an angle-like variable and will be labeled β or v . Choosing $u^1 = \alpha$, and $u^2 = \beta$, (4.3.15a, b) are then in a more familiar form:

$$\kappa_N = \kappa_\alpha |\nabla\alpha| + \kappa_\beta |\nabla\beta| \cos \hat{\alpha}(\nabla\alpha, \nabla\beta) \quad (4.3.16a)$$

$$\kappa_G = \kappa_\beta |\nabla\beta| \sin \hat{\alpha}(\nabla\alpha, \nabla\beta). \quad (4.3.16b)$$

In some cases it may be convenient to use the tangent vector \hat{t} and the principal normal \hat{n} to a curve as basis vectors of a coordinate system. The third basis vector is the binormal \hat{b} and is defined by

$$\hat{b} \equiv \hat{t} \times \hat{n}. \quad (4.3.17)$$

Often the curve on which this frame is attached is the magnetic axis (to be defined rigorously below). See, for instance, Shafranov (1968), Solov'ev and Shafranov (1970) and Solov'ev (1975). This type of system has gained renewed interest in fusion-plasma physics, since the development of the helical-axis stellarators and heliacs.

The coordinate frame $\hat{t}, \hat{n}, \hat{b}$ changes from point to point along its “base”-curve and is known as the moving Frenet-Serret frame. (Other names are F-S trihedron, F-S trihedral or F-S triad. Note, however, that we used the word “triad” with a different meaning as a third-order “simple” tensor.) It is necessary to introduce a second scalar function τ , known as the *torsion*, in order to relate the change in \hat{t}, \hat{n} and \hat{b} via the F-S formulae:

$$\frac{d\hat{t}}{dl} = \kappa \hat{n} \quad (4.3.18a)$$

$$\frac{d\hat{n}}{dl} = -\kappa \hat{t} + \tau \hat{b} \quad (4.3.18b)$$

$$\frac{d\hat{b}}{dl} = -\tau \hat{n}. \quad (4.3.18c)$$

The first expression, (4.3.18a) is our definition of the curvature κ encountered earlier in (4.3.1). The last equation defines the torsion τ . To show (4.3.18c), we first take the derivative of (4.3.17) and apply (4.3.18a). Hence, $d\hat{b}/dl = \hat{t} \times d\hat{n}/dl$, implying that $d\hat{b}/dl \perp \hat{t}$. Furthermore, from $\hat{b} \cdot \hat{b} = 1$, we have that $\hat{b} \cdot d\hat{b}/dl = 0$ or $d\hat{b}/dl \perp \hat{b}$. Therefore, $d\hat{b}/dl$ is perpendicular to both \hat{t} and \hat{b} , meaning that it is directed along $-\hat{n}$, since $\hat{t}, \hat{n}, \hat{b}$ are mutually orthogonal and form a right-handed system. τ is the proportionality constant. The middle F-S equation is easily proved by taking the derivative of $\hat{n} = \hat{b} \times \hat{t}$ and subsequently substituting (4.3.18a) and (4.3.18c). Then, from $\hat{b} = \hat{t} \times \hat{n}$ and $\hat{t} = \hat{n} \times \hat{b}$, (4.3.18b) follows.

4.4 Magnetic Pressure and Magnetic Tension

It is instructive to investigate how one can assign certain forces to magnetic-field lines embedded in a current-carrying medium. This allows us to elucidate commonly used concepts such as magnetic pressure and magnetic tension (Schmidt 1979; Cross 1989).

The Lorentz force density f (force per unit volume) in a current-carrying magnetized medium is given by

$$f = \frac{dF}{dV} = \mathbf{J} \times \mathbf{B}. \quad (4.4.1)$$

However, since the current density is $\text{curl } \mathbf{B}$, we can rewrite this in terms of \mathbf{B} only:

$$f = \frac{1}{\mu_0} (\nabla \times \mathbf{B}) \times \mathbf{B} = \frac{1}{\mu_0} (\mathbf{B} \cdot \nabla) \mathbf{B} - \frac{1}{2\mu_0} \mathbf{B} \nabla^2 \mathbf{B}. \quad (4.4.2)$$

The RHS is identical to (4.3.6). Following Schmidt (1979), we decompose f into its components along the Frenet-Serret trihedron tied to a field-line curve. The manipulations needed have already been performed in (4.3.7) and (4.3.8). With the RHS of (4.3.8) and (4.3.3), we have

$$f = \frac{B^2}{\mu_0} \kappa \hat{n} + \hat{t} \frac{\partial}{\partial l} \left(\frac{B^2}{2\mu_0} \right) - \nabla \left(\frac{B^2}{2\mu_0} \right), \quad (4.4.3)$$

where we have replaced \hat{B} by \hat{t} (we still assume that $\hat{B} = +\hat{l}$) for notational aesthetics. Recall that κ represents the curvature $\kappa = 1/R_c$, where R_c is the local radius of curvature, \hat{n} is the principal normal and $\kappa \equiv (\hat{B} \cdot \nabla) \hat{B} = \kappa \hat{n}$.

Expansion of ∇ along the Frenet-Serret trihedron by means of the partial derivatives $\partial/\partial l$, $\partial/\partial n$, and $\partial/\partial b$ gives

$$\nabla \left(\frac{B^2}{2\mu_0} \right) = \hat{t} \frac{\partial}{\partial l} \left(\frac{B^2}{2\mu_0} \right) + \hat{n} \frac{\partial}{\partial n} \left(\frac{B^2}{2\mu_0} \right) + \hat{b} \frac{\partial}{\partial b} \left(\frac{B^2}{2\mu_0} \right). \quad (4.4.4)$$

Here, \hat{b} is the binormal and *not* the unit vector along \mathbf{B} , as is sometimes used in the literature. This implies that the component along \mathbf{B} (or \hat{t}) in the expression for f vanishes:

$$f = 0\hat{t} + \left[\frac{B^2}{\mu_0} \kappa - \frac{\partial}{\partial n} \left(\frac{B^2}{2\mu_0} \right) \right] \hat{n} - \frac{\partial}{\partial b} \left(\frac{B^2}{2\mu_0} \right) \hat{b}. \quad (4.4.5)$$

We can interpret this result in the following way. In the plane perpendicular to the field line (i.e., the normal plane which contains the vectors \hat{n} and \hat{b}), the force density can be thought of as a superposition of a *pressure*, as if the magnetic field were a fluid with the “magnetic pressure” $B^2/(2\mu_0)$, and a *tension* as if the field lines were elastic cords each under a tension B/μ_0 , pressing against

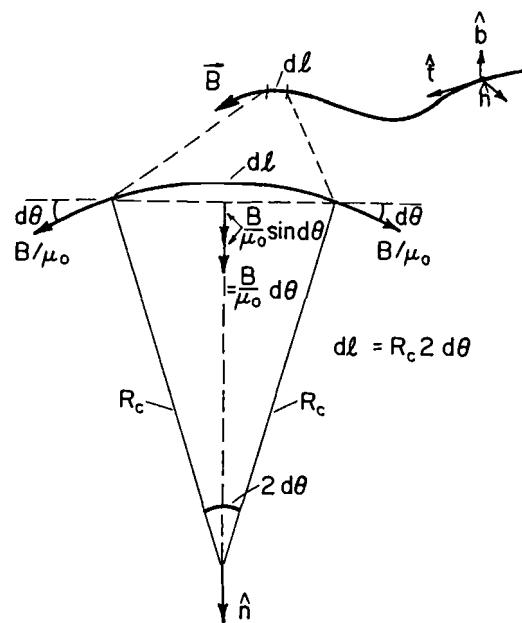


Fig. 4.3. Computation of the tension on a field line. See text

the fluid (Schmidt 1979, page 89). The part of the force density representing pressure is the negative gradient of the magnetic pressure:

$$-(\hat{n}\partial/\partial n + \hat{b}\partial/\partial b)(B^2/(2\mu_0)).$$

The tension needs more explanation. In this context, “tension” must be interpreted as a *force* pulling tangentially on a single field line. (In elasticity, by contrast, “tension” stands for *pulling-force-per-unit area* (Symon 1971).) The force density representing this “tension” is $B^2\kappa/\mu_0 = B^2/(\mu_0 R_c)$. Conventionally, the number of field lines per unit area is taken to be equal to the local value of $|B| = B$ so that the area associated with a single field line is $1/B$. Therefore, the force per unit length (of one field line) in the \hat{n} direction equals $(B^2/(\mu_0 R_c))(1/B) = B/(\mu_0 R_c)$. How this force per unit length is related to a tension of B/μ_0 is made plausible in Fig. 4.3. The force in the \hat{n} direction due to an elemental length dl equals $2Bd\theta/\mu_0$ (for $\sin d\theta \approx d\theta$). With $dl = R_c 2d\theta$, this is then $B/(\mu_0 R_c)dl$, leading to the above force per unit length (due to one field line) $B/(\mu_0 R_c)$.

4.5 Magnetic Surfaces

Any surface that is traced out by a series of magnetic-field lines is called a *magnetic surface*. In practice, however, the name is reserved for “closed” surfaces of a cylindrical or toroidal type. In that context, a magnetic surface is also called a *flux surface*.

a) Toroidal Systems

In a toroidal system, a magnetic surface (or flux surface) is a surface which is ergodically covered by non-closing magnetic-field lines. The structure of magnetic fields and surfaces in various devices has been discussed instructively by Morozov and Solov'ev (1966). The *existence* of magnetic surfaces has been investigated in the above review paper and by many others. We shall say a few words on the existence question in Chap. 9, where a literature survey is included.

The equation of a field line can be cast in a Hamiltonian form. Hamiltonian theory guarantees the existence of magnetic surfaces in devices with a symmetry direction: an ideal, axisymmetric tokamak; a helically symmetric straight stellarator; etc. In non-symmetric devices, on the other hand, magnetic surfaces do not exist rigorously everywhere because ergodic regions of finite volume exist (see Chap. 9). In practice, we can get away with considering approximate magnetic surfaces provided the perturbations away from symmetry are not too large. In principle, one should either measure the surfaces experimentally or follow field lines on a computer to demonstrate the existence of suitable surfaces. Note that in a tokamak, a toroidal (or parallel) plasma current is necessary for the existence of magnetic surfaces; in a stellarator, no such current is required, and it is meaningful to consider vacuum magnetic surfaces.

Between the magnetic surfaces traced out by non-closing field lines, we can define surfaces that contain field lines that close upon themselves after one or several transits around the machine (we have to “construct” or define these surfaces, since they are not traced out by a single field line). Such surfaces are called *rational surfaces*, because the ratio of the number of poloidal transits (the short way around the torus) to the number of toroidal transits (the long way around) of a field line is a rational number. The other surfaces are called *irrational surfaces*. Since the set of real numbers is dense, every irrational number can be approximated infinitesimally closely by a rational number and vice versa. The same holds for magnetic surfaces. However, the rational numbers are denumerable whereas the irrationals are not; there are infinitely more irrationals than rationals. Nevertheless, the important point is that every rational surface can be approximated very accurately by a nearby irrational surface, and the other way around.

Suitable magnetic surfaces are usually taken to be a requirement for adequate plasma equilibrium as described by the MHD equilibrium equation $\mathbf{J} \times \mathbf{B} = \nabla p$. These magnetic surfaces are constant pressure surfaces, since $\mathbf{B} \cdot \nabla p = 0$, and the current-density lines lie on these surfaces because $\mathbf{J} \cdot \nabla p = 0$. A consequence of the simple force balance, $\mathbf{J} \times \mathbf{B} = \nabla p$, and the current-continuity equation, $\nabla \cdot \mathbf{J} = 0$, is the “current-closure condition” on rational surfaces: the closed loop integral $\oint dl/B$ must be constant on rational surfaces, no matter on which field line we start (we shall discuss this condition below). Violation of this condition leads to the formation of an island-chain configuration and possibly to ergodic behavior of field lines in a volume in the vicinity of the rational surfaces.

In axisymmetric devices, the loop integral is constant and no problems arise. In non-axisymmetric devices, however, $\oint dl/B$ is not constant on rational surfaces, and according to ideal MHD the rational surfaces are unstable and break up, leading to the formation of fiber-like structures with many magnetic axes (i.e., the creation of magnetic islands) (Solov'ev and Shafranov 1970).

It is questionable, however, whether the simple MHD equilibrium model based on $\mathbf{J} \times \mathbf{B} = \nabla p$ is not too crude an approximation. The model neglects potentially important effects such as resistivity, non-scalar pressure (including viscosity), finite flow velocities (inertial forces) and finite Larmor radius effects. In reality, one must solve a time evolution problem; the system is likely to evolve to a self-consistent stationary state that probably has some magnetic surfaces, but not those predicted by ideal MHD. It can be expected that some of the requirements of the idealized model will be washed out by dissipative and other effects.

From a pragmatic point of view, we assume that (even in non-axisymmetric systems) a suitable set of nested toroidal magnetic surfaces is present, whose cross-section in a poloidal plane forms a set of smooth closed curves. This "imaginary" structure is actually the zero-th order approximation to the real structure. The magnetic surfaces in ideal tokamaks are symmetric about the major axis; the toroidal direction is ignorable. In stellarators, the surfaces have no symmetry direction and can take many forms. As an example, consider a surface of an $l = 2$ stellarator (Fig. 4.4). The toroidal magnetic surface is produced approximately by making an elliptic curve rotate in the poloidal direction, while moving toroidally along the minor axis.

The degenerate magnetic surface with the limiting value of zero volume is called the *magnetic axis*. It is formed by a field line that we assume closes upon itself after one toroidal transit. It is a line around which the magnetic surfaces are nested. A magnetic axis does not necessarily have to be circular, but can have, for example, a helical form ($l = 1$ stellarator, helical-axis stellarator). (The scope of our treatment is limited to situations with only one magnetic axis; thus, we disregard magnetic islands.)

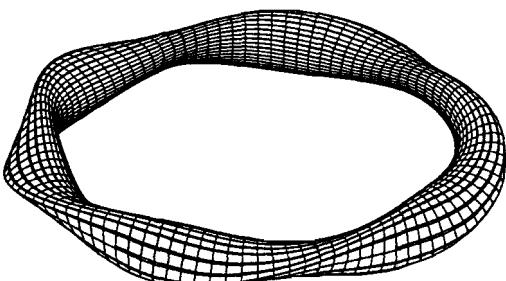


Fig. 4.4. Three-dimensional picture of a flux surface of an $l = 2$ stellarator. The elliptical cross section rotates poloidally when traversing toroidally. (Provided by H. Wobig and J. Kisslinger, Max-Planck-Institut für Plasmaphysik, Garching bei München, Fed. Rep. Germany.)

b) Open-Ended Systems

Magnetic-mirror devices do not possess magnetic surfaces that are traced out by one single field line. This leads, in principle, to a considerable degree of arbitrariness. One must *define* the shape of the surfaces, and a natural choice seems to be that of a circular cross-section in the center of the device, i.e., at $z = 0$ if the device is bounded axially by $z = \pm L$. For a simple baseball mirror, this circle fans out into an ellipse at one end, at $z = +L$, and into another ellipse rotated by 90-degrees at $z = -L$. In tandem mirrors with thermal barriers, the magnetic-surface structure is essentially a connection of a regular cylinder and baseball-coil generated surfaces.

4.6 Curvilinear Coordinate Systems in Confinement Systems with "Simple" Magnetic Surfaces

In this section we undertake a qualitative exploration of the kind of curvilinear coordinate systems suitable for the description of magnetically confined plasmas. We start the section with a brief reminder of the terminology in general curvilinear systems. Then we attack successively coordinates suitable for toroidal confinement systems on the one hand and for open-ended devices on the other. In this first category, we begin with a discussion of the cylindrical-toroidal hybrid coordinate system commonly used in elementary considerations, which we shall call the "elementary" toroidal system. It is based on "ideal" or "perfect" tori, by which are meant tori characterized by *circles* upon intersection with vertical and horizontal planes. Then we describe a generalization of that hybrid coordinate system. The main coordinate surfaces will now be the toroidal magnetic surfaces, "toroids" (which are topologically equivalent to regular tori). The usual concept of angle is generalized as well in order to find a coordinate system in which the magnetic-field lines appear as straight lines. To clarify the issue, we present a simple example next. To end the exposition on toroidal systems, we treat the (ρ, β, l) system, which has the arc length along a magnetic-field line as its third coordinate. This system is later identified as a particular type of Clebsch system. The disadvantages of its use in toroidal devices are stressed. The section is concluded with a note that Clebsch-type systems are most appropriate for open-ended systems.

A coordinate system is a tool to locate a point uniquely. In three-dimensional space, it consists of three suitable families of surfaces such that a point is defined by the intersection of three surfaces, one for each family. The *coordinate curves* are the intersection lines of two *coordinate surfaces*. We label the three coordinates by u^1, u^2 and u^3 ; the coordinate surfaces are labeled by $u^i = \text{constant}$ while we designate a coordinate curve by its running variable u^j (where it is understood that this curve lies on the intersection of $u^i = \text{constant}$ and $u^k = \text{constant}$, where $i \neq k, i \neq j$). A point $P(c^1, c^2, c^3)$ is completely determined by the

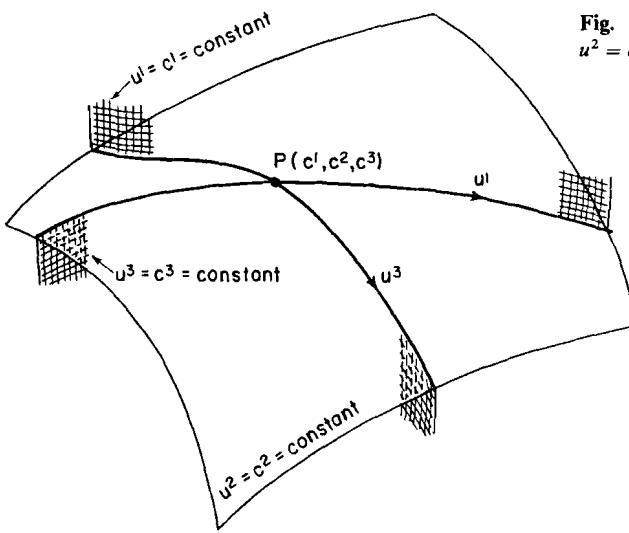


Fig. 4.5. Position location on a $u^2 = c^2 = \text{constant}$ surface

intersection of the three coordinate surfaces, $u^1 = c^1$, $u^2 = c^2$, $u^3 = c^3$ or equivalently by the intersection of two coordinate curves lying in the appropriate coordinate surface. For example, Fig. 4.5 shows that P is at the crossing point of the u^1 curve ($u^2 = c^2$ and $u^3 = c^3$) and the u^3 curve ($u^1 = c^1$ and $u^2 = c^2$), where evidently both curves lie in the $u^2 = c^2$ surface.

We shall use only *right-handed* coordinate systems. This implies that the triple product of the basis vectors $e_i \cdot e_j \times e_k$ or $e^i \cdot e^j \times e^k$ is positive as long as i, j and k are cyclic permutations of 1, 2, 3. It is clear that the *order* in which we take coordinates as well as their orientation is important: there is a first, a second and a third coordinate. (Of course, the order can be changed at will, but often then the orientation of a coordinate must be changed.)

4.6.1 Toroidal Systems

In toroidal plasma confinement the coordinate systems are based on toroidal topology. "Proper" toroidal coordinates, being the three-dimensional version of bipolar coordinates, have been used only occasionally in fusion-plasma equilibrium calculations; see, e.g., Shafranov (1960) and Fielding and Hitchon (1980). These coordinates are discussed briefly in Chap. 10.

a) The "Cylindrical-Toroidal" or "Elementary" Toroidal System

The best known coordinate system is the "pseudo-toroidal" or "hybrid cylindrical-toroidal" or the "elementary" toroidal coordinate system. This coordinate system is essentially a "toroidalized" cylindrical system with coordinates

$$(R, \zeta_c, z),$$

as shown in Fig. 4.6. The cylindrical angle is measured from the x -axis towards

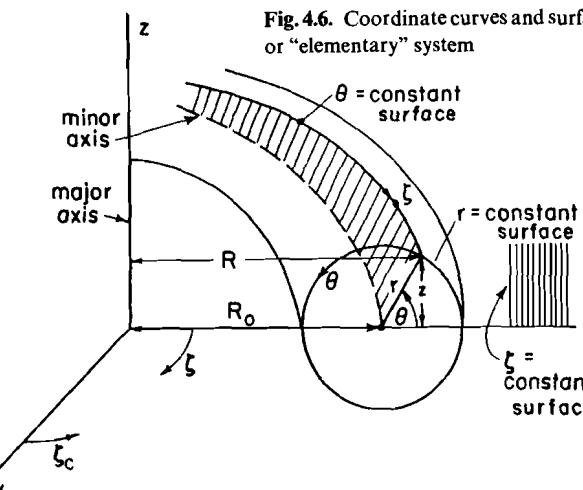


Fig. 4.6. Coordinate curves and surfaces of the "cylindrical-toroidal" or "elementary" system

the y -axis. The "toroidalization" is done by "replacing" R and z by r and θ . Because conventionally, θ and ζ are taken as the second and third coordinates, respectively, whereas ζ_c is the second in the cylindrical system, the angle ζ is measured in the positive sense from the y -axis towards the x -axis (i.e., $\zeta = \pi/2 - \zeta_c$). Only then has the trihedron of the unit vectors r , θ , ζ a right-handed orientation. Thus the coordinates are $u^1 = r$, $u^2 = \theta$, $u^3 = \zeta$.

The coordinate surfaces $r = \text{constant}$ are nested perfect tori with *major radius* R_o . The circle with radius R_o is called the *minor axis*, while the axis in the center of the doughnut hole is the *major axis* (in Fig. 4.6, this is the z -axis). r is called the *minor radius* and measures the distance from the minor axis. The coordinate surfaces $\theta = \text{constant}$ touch the minor axis and are partial cones. The third family of coordinate surfaces, $\zeta = \text{constant}$, are planes touching the major axis. An r -coordinate curve is formed by the intersection of a $\theta = \text{constant}$ "cone" and a $\zeta = \text{constant}$ plane. It is a *straight line*, with origin on the minor axis. The θ -coordinate curve is a *circle* obtained from the intersection of a torus $r = \text{constant}$ and a plane $\zeta = \text{constant}$. The third coordinate curve ζ , is a *circle* traced out by the intersection of the torus $r = \text{constant}$ and the "cone" $\theta = \text{constant}$. When moving around the minor axis (the short way around), we move in the *poloidal* direction. The θ -coordinate curve encircles the minor axis, hence the designation that θ is the *poloidal angle*. The plane in which the poloidal-angle coordinate curve lies is called the *poloidal plane* ($\zeta = \text{constant}$ is a *poloidal plane*). Analogously, going around the major axis corresponds to moving in the *toroidal* direction. ζ is the *toroidal angle*. The relationship between x, y, z and r, θ, ζ is as follows:

$$\begin{aligned} x &= (R_o + r \cos \theta) \sin \zeta \\ y &= (R_o + r \cos \theta) \cos \zeta \\ z &= r \sin \theta . \end{aligned} \quad (4.6.1)$$

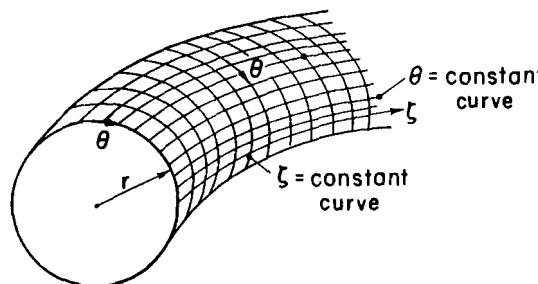


Fig. 4.7. Position location on a toroidal surface via a θ and ζ grid

The usual procedure for specifying the position of a point in space is to first choose a coordinate surface, $r = \text{constant}$, and then to lay out a grid of θ and ζ curves on this torus. The intersection of two curves, one from each grid, fixes the position of the point in question (Fig. 4.7). Note that the curves of the grid satisfy the properties that any two members of the same family never intersect each other, and that every curve of one family intersects every curve of the other family at precisely one point. This observation here, of course, is plausible, since the two sorts of coordinate curves are circles intersecting each other at right angles. The angles θ and ζ are multiple-valued functions with regard to the position on the torus. θ and ζ increase by 2π after encircling the minor axis and the major axis, respectively. Every point on the toroidal surface is thus labeled by $\theta + 2\pi m$ and $\zeta + 2\pi n$ for all integers m and n . Below, we shall denote the poloidal and toroidal angles in the “elementary” toroidal system by θ_e and ζ_e .

b) Generalized Cylindrical-Toroidal Coordinates; Flux Coordinates

In theoretical fusion-plasma physics, it is customary to choose as the first family of coordinate surfaces the *magnetic surfaces*, regardless of their shape. We label these magnetic surfaces by $\varrho = \text{constant}$, thus $u^1 \equiv \varrho$. Below, we shall identify this general label ϱ with more specific measures such as the magnetic flux enclosed by the surface, its volume or the pressure.

The coordinate surfaces for the angles θ and ζ are generalized as well. The $u^2 = \text{constant}$ surfaces are now based on the *magnetic axis* and, in general, are partial-cone like surfaces. (Note that the magnetic axis does not normally coincide with the geometrical minor axis and can be non-circular or even non-planar.) These surfaces can be thought of as being produced by bending, stretching or squeezing the partial cones of the “elementary” toroidal coordinate system (of Fig. 4.6), provided that each surface crosses the toroidal magnetic surface only once. Figure 4.8 shows an example of the cross-section in the yz Cartesian plane of a simple tokamak with non-concentric circular toroidal magnetic surfaces, where the $u^2 = \text{constant}$ surfaces are chosen such that they intersect the tori at right angles. Because the coordinate surfaces $u^2 = \text{constant}$ have topologically the same structure as the $\theta_e = \text{constant}$ cones in the “elementary” toroidal system, it is convenient to denote u^2 by the symbol θ . It is clear that θ is a poloidal

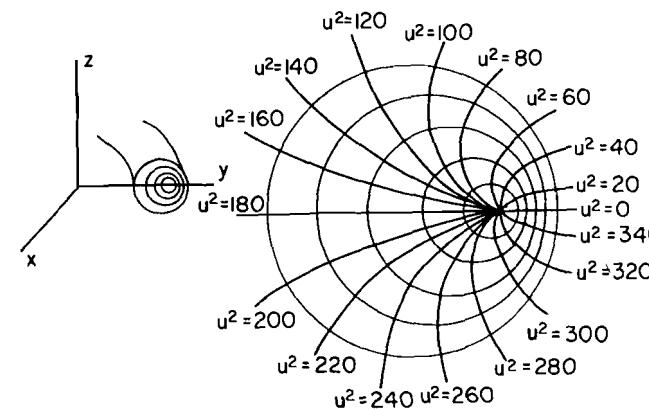


Fig. 4.8. $u^2 = \text{constant}$ surfaces which are perpendicular to the Shafranov-shifted flux-surfaces in a simple tokamak

angle-like variable or a generalized poloidal angle. In short, we shall call θ the “poloidal” angle, although it is understood that it is a generalization of the usual concept of “angle”. Likewise, the $u^3 = \text{constant}$ coordinate surfaces are, in general, a deformed version of the $\zeta_e = \text{constant}$ planes. The $u^3 = \text{constant}$ surfaces touch the major axis and fan out to cut through the magnetic surface. Because of the analogy with the variable ζ_e , we denote u^3 by ζ and call this angle-like variable the “toroidal” angle. We also say that a surface $\zeta = \text{constant}$ makes a “poloidal” cross-section of the torus. From now on, the words, “toroidal” and “poloidal” are related to the curvilinear coordinates θ and ζ , which do not necessarily coincide with the “elementary” angles θ_e and ζ_e . The context should make it clear what is meant.

Coordinate curves for θ and ζ are constructed as usual, via intersections of coordinate surfaces. With the coordinate surfaces constructed as explained above, we obtain *closed curves* for the coordinate curves as shown in Fig. 4.9. We require that the functions $\theta(x, y, z)$ and $\zeta(x, y, z)$, which are transformations from x, y, z to θ and ζ , have continuous first derivatives, implying that the tangents to the curves and the gradients to the surfaces are continuous vector fields. $\theta(x, y, z)$ and $\zeta(x, y, z)$ are multiple-valued functions since they change by 2π after each complete circuit. Each point $P(x, y, z)$ can have θ coordinates that can differ by multiples of 2π and similarly for ζ . θ and ζ must be single valued, however, on any other path on the surface which does not encircle the major or minor axes. On every magnetic surface, the θ and ζ coordinate curves lay out a grid as shown in Fig. 4.10. It should be noted that it is sufficient for *place* positioning to isolate a $\varrho = \text{constant}$ surface and lay out a grid of θ and ζ curves. However, this is not enough for a proper description of *vectors*. With the grid of contours, the covariant-(tangent-) basis vectors are determined at every point, but the contravariant-(reciprocal-) basis vectors ∇u^i can only be found if the $u^2 = \text{constant}$ and $u^3 = \text{constant}$ surfaces are known.

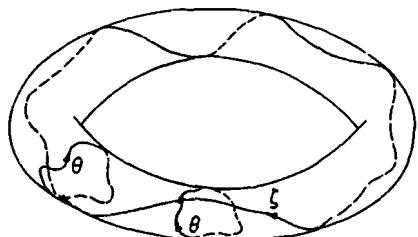


Fig. 4.9. Coordinate curves for the generalized poloidal and toroidal “angles”

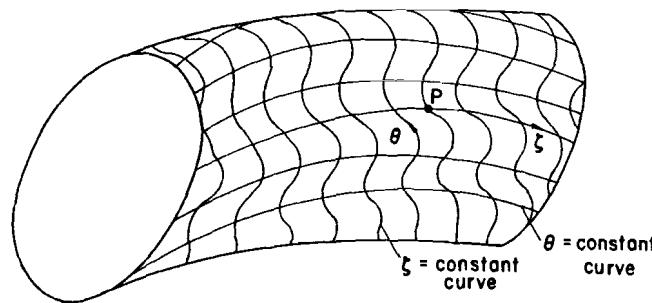


Fig. 4.10. Position location on a toroidal surface based on a grid constructed for generalized “angles” θ and ζ

We can always construct a suitable set of “new” angle functions (or variables) $\tilde{\theta} \equiv \tilde{\theta}(R)$ and $\tilde{\zeta} \equiv \tilde{\zeta}(R)$ by adding functions which are periodic with a period of 2π to the original θ and ζ :

$$\tilde{\theta} = \theta + f(\rho, \theta, \zeta) \quad (4.6.2a)$$

$$\tilde{\zeta} = \zeta + g(\rho, \theta, \zeta). \quad (4.6.2b)$$

The periodicity of f and g in both θ and ζ is equivalent to stating that f and g are single-valued functions of θ and ζ . Because of this single valuedness, the (θ, ζ) coordinate curves are deformed in a such a way that the new coordinates $(\tilde{\theta}, \tilde{\zeta})$ have proper curves as well, making them acceptable coordinates.

The reason for working in these coordinate systems is straightforward. To perform analytic calculations, it is convenient to choose a “natural” coordinate system, such that in that coordinate system the geometry of the problem under consideration becomes simple. Then the complex geometrical features that exist in a “neutral” (e.g., Cartesian) coordinate frame are taken care of in the transformation equations between the coordinate systems. In theoretical plasma-physics computations, it is very helpful to choose the $\theta(x, y, z)$ and $\zeta(x, y, z)$ functions so that the magnetic-field lines (which oscillate in the “elementary” toroidal coordinate system) appear as *straight lines*. This means that, along a field line, $d\theta/d\zeta$ is a constant (more precisely, is a flux function), for that choice of θ and ζ functions. A particular choice for the θ, ζ coordinates makes both the \mathbf{B} -field lines

and the \mathbf{J} current-density lines straight. It is customary to call coordinates in which the magnetic field is straight *magnetic coordinates* or *flux coordinates*. When we deal with angles, these coordinates will be labeled with an index f , θ_f and ζ_f . (The flux-surface label ρ is exempted from this index, since we want to make field lines that lie in $\rho^1 = \rho = \text{constant}$ surfaces straight in the remaining ρ^2 and ρ^3 coordinates.)

To summarize, given that the magnetic field \mathbf{B} is known everywhere, and consequently, so is \mathbf{J} , we try to find suitable poloidal and toroidal coordinates that make \mathbf{B} and/or \mathbf{J} straight in (θ_f, ζ_f) . One can start with any acceptable grid of θ and ζ curves as a function of which the magnetic field is known (for example, the “elementary” angles, θ_e and ζ_e), and deform the coordinate system until it possesses the desired properties.

c) An Illustrative Example

In Fig. 4.11a, we show the constant flux-surface contours of a circular tokamak, where the surfaces are non-concentrically nested. Also shown are the cross-sections of constant θ_e surfaces. The coordinate system used here is very similar to the “elementary” toroidal system, except that it is based on the magnetic axis, rather than the minor axis of the torus. Because of this close relationship we designate the angle by θ_e .

Consider one of the larger surfaces (shown dashed in Fig. 4.11a) and its (θ_e, ζ_e) grid (shown in Fig. 4.11b). The magnetic field is assumed to traverse this surface with constant pitch everywhere. Let us assume that the rotational transform τ equals 3 on this particular surface (τ will be defined rigorously below). This means that the magnetic field makes 3 poloidal transits for every toroidal transit. If we mark the points where the \mathbf{B} line crosses the $\theta_e = \text{constant}$ curves in Fig. 4.11b and transfer them to a (θ_f, ζ_f) coordinate system, we observe an oscillatory behavior of the field line as seen in Fig. 4.11c. To make the \mathbf{B} lines straight in a new (θ_f, ζ_f) coordinate system, it is obvious that the arc lengths on the constant flux-surface contours cut off by the $\theta_f = \text{constant}$ contours must be equal. This can be accomplished by constructing poloidal angle surfaces as shown in Fig. 4.11d. (This procedure is equivalent to isolating the surface under consideration and constructing angles as “usual” based on the center of the circle, not the magnetic axis. Still another method is to isolate the surface and disregard the structure of the coordinate surfaces, but simply construct an equidistant (θ_f, ζ_f) grid on the surface. This secondary procedure does not allow the determination of the gradient $\nabla\theta_f$, however.) The net result of all this is that the \mathbf{B} lines are straight in (θ_f, ζ_f) , but that our new θ_f has an oscillatory behavior as a function of the old θ_e (Fig. 4.11e).

d) (ρ, β, l) Coordinates in Toroidal Systems

An alternative (but complicated) set of flux coordinates in toroidal devices is the so-called (ρ, β, l) coordinate system. Again, ρ is a magnetic-surface label, β labels

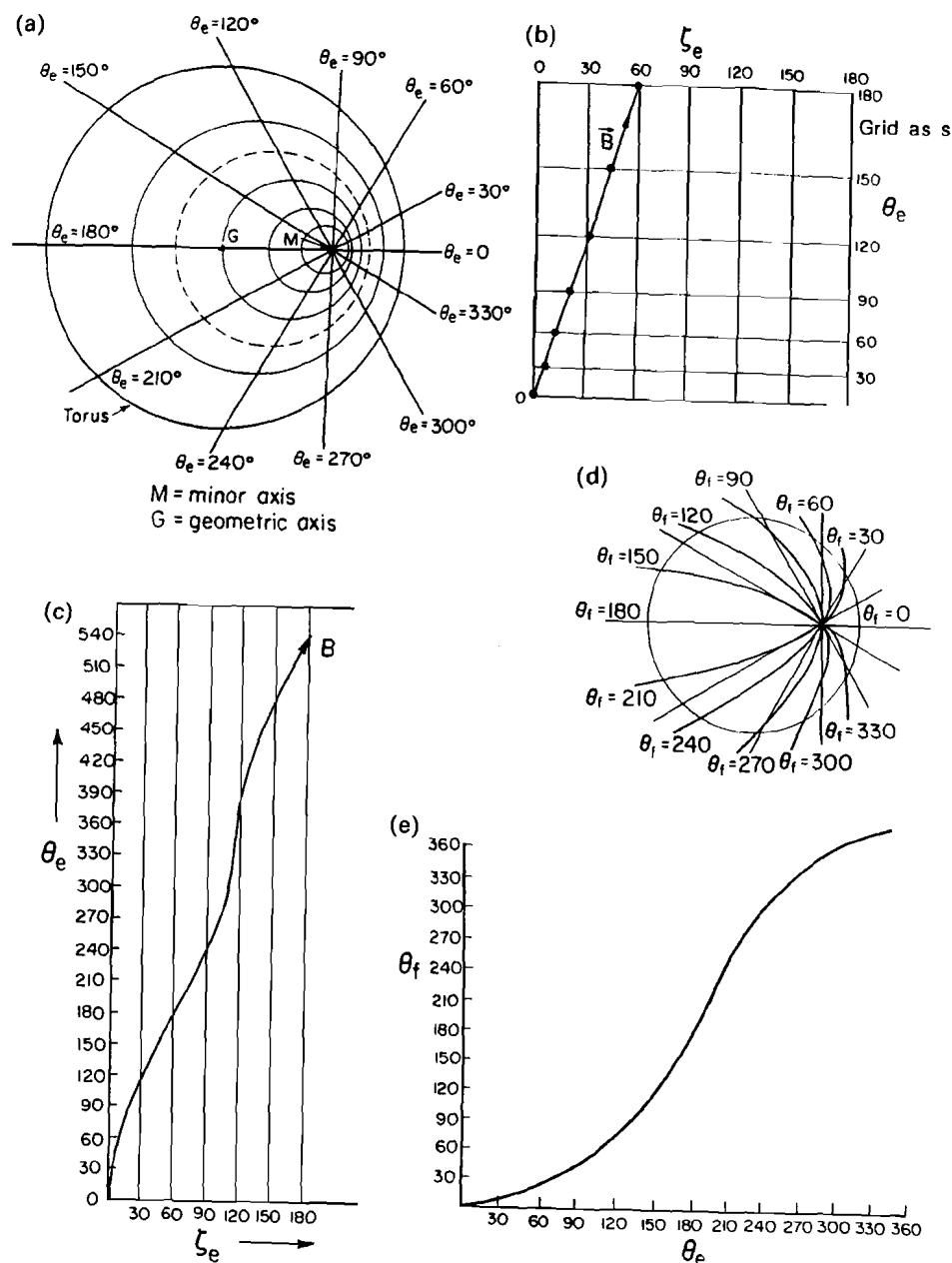


Fig. 4.11. (a) The usual poloidal angle surfaces (planes) $\theta_e = \text{constant}$, fanning out from the magnetic axis of a Shafranov-shifted tokamak plasma. (b) The (θ_e, ξ_e) grid as seen on a folded-open toroidal surface. The magnetic-field line is seen as a straight line. (c) A magnetic-field line “wiggles” in a (ξ_e, θ_e) coordinate system. (d) The constant coordinate surfaces of the flux-coordinate angle θ_f cut each toroidal surface equidistantly. (e) The flux-coordinate angle θ_f , in which the B -lines appear as straight lines, wiggles with respect to the old (usual) coordinate θ_e .

the field lines in a surface, and l is the length along a field line. A $\beta = \text{constant}$ surface is a “ribbon” starting at the magnetic axis and passing through the trajectory of a field line on a given surface, that twists around with the same pitch as the field line on that surface (see Fig. 4.12). The issue of $l = \text{constant}$ surfaces is not straightforward. First, one could choose an arbitrary origin, $l = 0$ for all field lines (this could be a common poloidal-plane cut), and then constant l surfaces must be constructed. The length l and thus these l surfaces are influenced by the pitch of the field lines. If the field lines twist around more on the outer surfaces, the l surfaces will look like paraboloids. These will not be regular paraboloids, in general, however, since the rate of twisting of a field line is, in general, not constant on a surface. An important consequence of such $l = \text{constant}$ surface structure is that in most cases, ∇l is not tangential (i.e., not parallel) to B .

On the other hand, an irrational surface can be thought of as being constructed from a single field line. Thus, on that flux surface, there would be only one point which we could assign a value $l = 0$. So what then are the $l = \text{constant}$ surfaces?

Similarly, what are the $\beta = \text{constant}$ surfaces given this last point of view? Since the magnetic lines lie on $\beta = \text{constant}$ ribbons, and one single field line covers the whole magnetic surface, it might appear that β is zero everywhere on the surface. Although in principle the entire surface can be parameterized by l -values, with $\beta = 0$, this is not very useful in practice. Rational surfaces also present problems. If a field line closes after, say, n toroidal transits, there are n (probably overlapping) β ranges ($\beta = 0 \rightarrow \beta = \beta_{\text{final}}$) with jumps between them.

A possible way to make sense out of this coordinate system would be to introduce cuts in the torus and limit l to one toroidal transit (which is effectively what we did when we constructed the $l = \text{constant}$ “paraboloids”). A complication arises, however, because of the multiple valuedness of both β and l in that

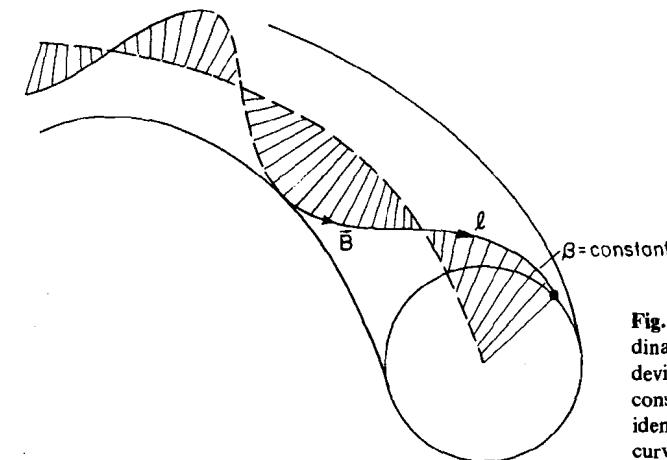


Fig. 4.12. A $\beta = \text{constant}$ coordinate surface in a toroidal device. The intersection of $\beta = \text{constant}$ and $q = \text{constant}$ is identified with the l -coordinate curve

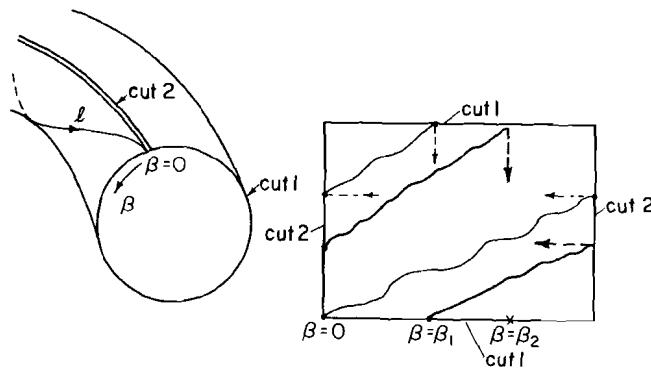


Fig. 4.13. Field-line labeling by means of the β Clebsch coordinate. The figure on the right is the topologically equivalent folded-open surface of the toroidal surface on the left. See text

picture. One considers every field line in this cut as an independent entity (independent of the fact that all irrational lines are generated by one single field line); β still labels the field lines. In Fig. 4.13, we show the cuts and the topologically equivalent folded-open surface. The opposite sides of the rectangle are topologically identical. If we pick a field line at the crossing of the cuts and label it by $\beta = 0$, it will arrive at a point where $\beta = \beta_1$. However, β is still 0 because β is constant along the field line. For single-valuedness, we must thus require that $\beta(l=L, \varrho) = \beta(l=0, \varrho)$. But if we start with a field line labeled by $\beta = \beta_1$, we will arrive at $\beta = \beta_2$, which must be identified with β_1 , which, in turn, was equal to 0. This means that every point on the flux surface can be labeled by *every* value of β , ranging from 0 to 2π , if we take $\beta = 2\pi$ the last label at the other end of the cut. These problems arise because the l coordinate curve does not close upon itself after one toroidal or poloidal transit, in contrast to the poloidal and toroidal angles θ and ζ , whose coordinate curves do close after their respective transits.

The (ϱ, β, l) coordinates, therefore, may be useful for formal reasoning, but lead to severe problems in practice. The $\mathbf{B} \cdot \nabla$ operator reduces to $B(\partial/\partial l)$ in this coordinate system, which may be useful in calculations involving local properties of the surface.

4.6.2 Open-Ended Systems

Unlike toroidal topology, the (ϱ, β, l) system is very well suited to open-ended systems. It is precisely the connection of one end of the device onto the other, which caused problems for the single valuedness of β and l in toroidal systems, that is avoided here. ϱ labels the chosen "cylindrical" flux surfaces, β is a suitable angle-like variable, and l measures the length of a field line as before. A (ϱ, β, l) system is known as a Clebsch system. It will be discussed in detail below.

4.7 Magnetic-Surface Labeling

Before we can develop flux-coordinate systems in detail, we must find adequate measures that represent magnetic surfaces. This section provides these quantities. We shall again differentiate between toroidal and open-ended systems. For toroidal systems, we first consider the magnetic fluxes enclosed by and outside of the magnetic surfaces. We distinguish between toroidal magnetic flux, poloidal ribbon flux and poloidal disk flux. Then other magnetic-surface labels such as volume and pressure are mentioned. Finally, the magnetic axis is defined. For open-ended systems, axial and azimuthal magnetic fluxes and scalar pressure are considered.

4.7.1 Toroidal Systems

A convenient way to label the magnetic surfaces is by means of the magnetic flux. Because the magnetic-field lines are by definition everywhere tangent to a magnetic surface, the flux inside or outside a magnetic surface is conserved. For this reason, we can also call magnetic surfaces constant flux surfaces or, in short, *flux surfaces*.

There are several magnetic fluxes which label a flux surface in a toroidal device uniquely. Assume that an arbitrary $(\theta(x, y, z), \zeta(x, y, z))$ coordinate system has been established. First, there is the *toroidal flux* (Fig. 4.14a):

$$\Psi_{\text{tor}} \equiv \iint_{S_{\text{tor}}} \mathbf{B} \cdot d\mathbf{S} . \quad (4.7.1)$$

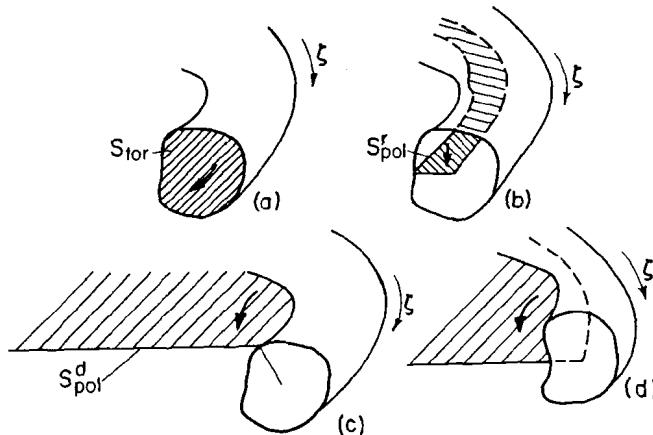


Fig. 4.14a-d. Cross sections through which the magnetic fluxes characterizing a magnetic-flux surface are defined. (a) The field lines intersecting S_{tor} determine the toroidal flux within a flux surface. (b) S_{pol} is a ribbon bounded by the magnetic axis and the flux surface. It "contains" the poloidal (ribbon) flux which resides inside the flux surface, Ψ_{pol} . (c) S_{pol}^d is a disk touching the magnetic surface. It picks up the poloidal (disk) flux outside the flux surface, Ψ_{pol}^d . (d) A special case of a disk is shown. It touches the magnetic surface where $\theta = \pi$, i.e., it lies in the equatorial plane

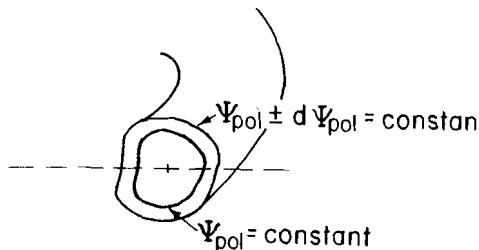


Fig. 4.15. The flux-surface volume between two infinitesimally close flux surfaces. + sign for Ψ_{pol}^r , - sign for Ψ_{pol}^d

S_{tor} is the surface produced by cutting the flux surface with a poloidal surface $\zeta = \text{constant}$. Next, the *poloidal fluxes* inside and outside a flux surface may be used. The poloidal flux *inside* the flux surface is taken, e.g., through a ribbon at constant θ between the magnetic axis and the surface, and is given by (Fig. 4.14b):

$$\Psi_{\text{pol}}^r \equiv \iint_{S_{\text{pol}}^r} \mathbf{B} \cdot d\mathbf{S}. \quad (4.7.2)$$

The superscript r refers to the ribbon. S_{pol}^r is the part of a $\theta = \text{constant}$ surface that lies within the flux surface under consideration. The flux *outside* the flux surface is obtained by integrating \mathbf{B} over a disk-like surface that is tangent to the flux surface everywhere (Fig. 4.14c), and is bounded by a ζ coordinate curve on the flux surface:

$$\Psi_{\text{pol}}^d \equiv \iint_{S_{\text{pol}}^d} \mathbf{B} \cdot d\mathbf{S} \quad (4.7.3)$$

The poloidal-disk flux includes the tokamak transformer flux, and no reference to the magnetic axis is required. The flux obtained from (4.7.3) is fixed as long as the disk is tangent to the flux surface.

To compare the two poloidal fluxes, it is convenient to choose a disk that is part of the approximately circular surface bounded by the magnetic axis (see Fig. 4.14d). Comparing Figs. 4.14b, d, it is clear that the fluxes are complementary with respect to the total flux in the system. The flux from the magnetic axis out to infinity has the same value as the flux through $S_{\text{pol}}^r + S_{\text{pol}}^d$. This remains true even if there is an additional vertical field component to keep the magnetic axis (the point where the net vertical field component is zero) from expanding. In every steady-state device which has fixed magnetic surfaces, such as a stellarator, the spatial changes in poloidal flux satisfy the relationship (see Figs. 4.14 and 4.15):

$$d\Psi_{\text{pol}}^d = -d\Psi_{\text{pol}}^r. \quad (4.7.4)$$

$d\Psi_{\text{pol}}$ is the differential flux between the two flux surfaces $\Psi_{\text{pol}} = \text{constant}$ and $\Psi_{\text{pol}} \pm d\Psi_{\text{pol}} = \text{constant}$. Because

$$d\Psi_{\text{pol}} \equiv \nabla\Psi_{\text{pol}} \cdot d\mathbf{R} \quad (4.7.5)$$

for every arbitrary differential vector $d\mathbf{R}$, (4.7.4) is equivalent to:

$$\nabla\Psi_{\text{pol}}^d = -\nabla\Psi_{\text{pol}}^r. \quad (4.7.6)$$

Of course, both sides of (4.7.5) and (4.7.6) must be evaluated at the same position.

In tokamaks, the situation is more complicated because the equilibrium is dynamic. Suppose we want to label the flux surfaces by means of the poloidal-disk flux, $\Psi_{\text{pol}}^d = \text{constant}$. In Chap. 11, we show that the increasing transformer flux forces the flux surfaces, labeled by $\Psi_{\text{pol}}^d = \text{constant}$, to shrink with a velocity \mathbf{v} such that $\mathbf{v} \cdot \nabla\Psi_{\text{pol}}^d = -\partial\Psi_{\text{pol}}^d/\partial t$. This velocity turns out to be equal to the inward pinch $\mathbf{E} \times \mathbf{B}$ velocity observed in a *fully-ionized ideal-MHD* plasma ($\sigma \rightarrow \infty$); since the infinite conductivity requires E_{\parallel} to be zero, the (induced) electric field has only a perpendicular (to \mathbf{B}) component E_{\perp} . Because both the plasma and the flux surfaces $\Psi_{\text{pol}}^d = \text{constant}$ move with the same speed, this is in effect a statement of the *frozen-flux theorem* in "ideal" plasmas. From the standpoint of the poloidal ribbon flux, the same conclusion can be drawn: because of the frozen flux theorem, the surfaces labeled by $\Psi_{\text{pol}}^r = \text{constant}$ move at the same speed as the plasma. The conclusion is that *it does not matter which label we use for the poloidal flux*.

Since we are primarily concerned with spatial variations of the flux functions, we can ignore the time evolution because it does not change the magnetic topology relative to the plasma. This is similar to taking a photograph and assuming that it reflects the time independent equilibrium state. Once this time evolution is disregarded, we can claim that (4.7.4) and (4.7.6) are still valid. (We stress that it is the spatial part of $d\Psi_{\text{pol}}$ in (4.7.4) that is of concern. We should say $\delta\Psi_{\text{pol}}^d = -\delta\Psi_{\text{pol}}^r$, where $\delta\Psi_{\text{pol}} = \nabla\Psi_{\text{pol}} \cdot d\mathbf{R}$, since, when the time t is retained as a variable, $d\Psi_{\text{pol}} = \partial\Psi_{\text{pol}}/\partial t dt + \nabla\Psi_{\text{pol}} \cdot d\mathbf{R}$. However, with the time dependence suppressed, we have that $d\Psi_{\text{pol}} = \delta\Psi_{\text{pol}}$.)

In a non-ideal plasma, the story is somewhat different. The finite resistivity allows magnetic-field diffusion to take place, while it also gives rise to transport processes such as classical and neoclassical particle diffusion fluxes across the flux surfaces. This leads to a self-consistent evolution of the magnetic structure and the plasma. We shall not deal with the time evolution of flux surfaces in this monograph, but refer the reader to the literature. (Plasma-equilibrium evolution on the resistive time scale in the context of the closure of moment equations in transport theory is discussed by Hinton and Hazeltine (1976), Chap. VII, and by Hirshman and Sigmar (1981), Chap. IX. For more detailed information, the (tokamak) theory has been developed in Grad and Hogan (1970), Grad (1974), Grad, et al. (1975), Pao (1978), Pereversev, et al. (1978), Hirshman and Jardin (1979), Blum and Le Foll (1984) and Blum (1989) among others.)

To develop the flux-coordinate formalism, we need more specific expressions for the toroidal and poloidal fluxes. In particular, we wish to derive a form-invariant expression with respect to the choice of θ and ζ functions (see, e.g., Kruskal and Kulsrud 1958). Consider the integral

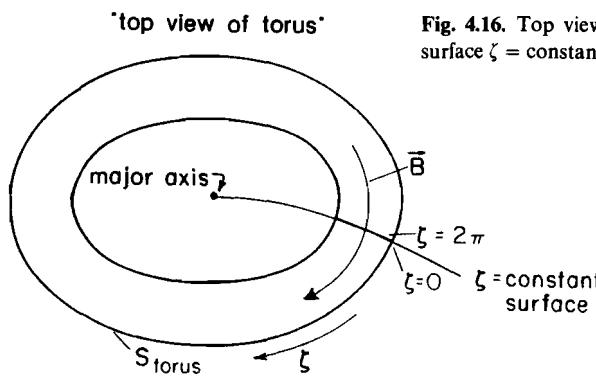


Fig. 4.16. Top view of a toroidal surface, cut by a surface $\zeta = \text{constant}$

$$\iiint_V d^3R \mathbf{B} \cdot \nabla \zeta = \iiint_V d^3R \mathbf{V} \cdot (\mathbf{B} \zeta) . \quad (4.7.7)$$

Here, d^3R is an elementary volume element and V is the volume enclosed by the flux surface. The second form follows from $\nabla \cdot \mathbf{B} = 0$. In Fig. 4.16 we show a top view of the flux surface together with a $\zeta = \text{constant}$ (poloidal) surface. In particular, we choose the $\zeta = 0 + 2\pi k$ surface. This surface cuts the torus open, allowing us to consider it as a volume enclosed by the surfaces S_{torus} , $S_{\zeta=0}$ and $S_{\zeta=2\pi}$. If we now apply Gauss' theorem to the integral of (4.7.7) over the region of the cut-open torus, we obtain

$$\iint_S \zeta \mathbf{B} \cdot dS = \iint_{S_{\zeta=2\pi}} \zeta \mathbf{B} \cdot dS + \iint_{S_{\zeta=0}} \zeta \mathbf{B} \cdot dS + \iint_{S_{\text{torus}}} \zeta \mathbf{B} \cdot dS . \quad (4.7.8)$$

The first integral on the RHS vanishes by definition: \mathbf{B} is tangent everywhere to the magnetic surface, $dS \perp \mathbf{B}$. ζ is constant on the surfaces $S_{\zeta=2\pi}$ and $S_{\zeta=0}$, meaning that ζ can be replaced by 2π and 0 in the second and third terms, respectively. Then (4.7.7) and (4.7.8) give

$$\iiint_V d^3R \mathbf{B} \cdot \nabla \zeta = 2\pi \iint_{S_{\zeta=2\pi}} \mathbf{B} \cdot dS . \quad (4.7.9)$$

$S_{\zeta=2\pi}$ (or $S_{\zeta=0}$) is what we previously called S_{tor} . A combination of (4.7.1) and (4.7.9) leads to the form-invariant expression for Ψ_{tor} :

$$\Psi_{\text{tor}} = \frac{1}{2\pi} \iiint_V \mathbf{B} \cdot \nabla \zeta d^3R . \quad (4.7.10)$$

Analogously, the poloidal ribbon flux Ψ_{pol}^r can be written as

$$\Psi_{\text{pol}}^r = \frac{1}{2\pi} \iiint_V \mathbf{B} \cdot \nabla \theta d^3R . \quad (4.7.11)$$

There exists a variety of flux-surface labels, other than the fluxes. Any function or variable that is uniform over a magnetic surface is referred to as a *flux function* or a *flux-surface quantity*. Such functions are independent of θ and ζ and depend only on the flux-surface label; hence, we can use these functions as surface labels. Any flux function can be used to label the flux surfaces. An obvious surface quantity is the *volume* enclosed by the magnetic surface:

$$V = \iiint_V d^3R = \iiint_V dx dy dz = \iiint_V \sqrt{g} d\varrho d\theta d\zeta . \quad (4.7.12)$$

Here, g is the determinant of the metric tensor, and \sqrt{g} is the Jacobian, $\sqrt{g} = [\nabla \varrho \cdot (\nabla \theta \times \nabla \zeta)]^{-1}$.

Another surface quantity (used in the plasma literature) is the scalar pressure p . From $\nabla p = \mathbf{J} \times \mathbf{B}$ which implies that $\mathbf{B} \cdot \nabla p = 0$, it follows that \mathbf{B} lies in constant pressure surfaces. Since \mathbf{B} also lies in flux surfaces, it follows that $p \equiv p(\Psi)$ only, although this may break down at points where $dp/d\Psi = 0$.

Two of the flux labels that we have referred to have the disadvantage of decreasing as we go radially outward; the pressure and the poloidal disk flux. To keep a right-handed system (and a positive Jacobian), we would have to change the orientation of, say, the toroidal angle ζ . Furthermore, this implies that positive integration along coordinate curves would give rise to a flux-surface volume and inner fluxes that are zero at the plasma boundary and maximum (in absolute value) at the magnetic axis, in contrast to common conventions. Although it is possible to keep track of these adaptations, it is quite confusing and may overload expressions with absolute value bars and/or (\pm) choices, to obtain the proper results for a general flux label ϱ . Therefore, we use only *radially increasing flux labels*. Then no angles need be changed and the formulae remain simple. It requires, however, that we use the *negative* of the flux functions that decrease radially. Thus, we can have $\varrho \equiv \Psi_{\text{tor}}$, $\varrho \equiv \Psi_{\text{pol}}^r$, $\varrho \equiv V$, $\varrho \equiv -p$, $\varrho \equiv -\Psi_{\text{pol}}^d$.

The label Ψ_{pol} is used most frequently in the tokamak literature, while Ψ_{tor} is more popular for stellarators. This is in part because the tokamak literature uses the safety factor q , whereas for stellarators the rotational transform, $t = 1/q$, is used as a measure of the rate of twisting of a field line. The volume was chosen by Hamada (1962) when he developed his coordinate system in which both \mathbf{B} and \mathbf{J} lines are straight.

The defining equation of a flux surface with flux-surface label ϱ , is:

$$\mathbf{B} \cdot \nabla \varrho \equiv 0 . \quad (4.7.13)$$

It simply states that \mathbf{B} lies in the surface $\varrho = \text{constant}$ everywhere. Its form is more familiar when $\varrho \equiv \Psi$.

$$\mathbf{B} \cdot \nabla \Psi \equiv 0 . \quad (4.7.14)$$

The degenerate magnetic surface, namely the magnetic axis, is properly defined by:

$$\nabla\varrho \equiv 0 \quad \text{or} \quad \nabla\Psi \equiv 0 . \quad (4.7.15)$$

This equation states that there is no singularity in the gradient of the flux (label) for ∇ approaching 0. Its physical meaning is easily understood. In the limit of $V \rightarrow 0$, we have that $\Psi_{\text{tor}} \propto r^2$, where r is an average minor radius. Then $\nabla\Psi_{\text{tor}} = (d\Psi_{\text{tor}}/dr)\hat{r}$ allows us to conclude that $\nabla\Psi_{\text{tor}}$, evaluated at $r = 0$ is zero. (When taking Ψ_{tor} , the dependence of the area on r^2 must be invoked to prove (4.7.15); when one wishes to prove (4.7.15) for Ψ_{pol} , one must make use of the fact that $B_\theta = 0$ at the magnetic axis.) For every other surface label $\varrho(\Psi)$, we must make use of the property that:

$$\nabla\varrho(\Psi) = \frac{d\varrho}{d\Psi} \nabla\Psi \quad (4.7.16)$$

with the condition that $d\varrho/d\Psi$ is non-singular at the axis.

4.7.2 Open-Ended Systems

In mirrors, similar flux labels as in toroidal systems are employed. The flux volume does not require further comment. Concerning magnetic fluxes, we are now limited to fluxes contained within the flux surface. By analogy to Ψ_{tor} and Ψ_{pol} in a torus, we have the axial flux Ψ_{ax} and the azimuthal flux Ψ_{az} , respectively. A scalar pressure p seems to be less convenient because the pressure is often anisotropic and not constant along a magnetic-field line. Depending on the choice of flux label, \hat{l} of the (ϱ, β, l) coordinate system points in the \mathbf{B} direction or its reverse: $\hat{l} = \pm \hat{\mathbf{B}}$.

4.8 The Rotational Transform in Toroidal Systems

The total poloidal angle $\Delta\theta$ traversed by a field line after one toroidal transit ($\Delta\zeta = 2\pi$), is called the rotational transformation angle ι (Greek small letter iota). The average rotational transformation angle is defined as (Morozov and Solov'ev 1966):

$$\iota = \lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \iota_k}{n} . \quad (4.8.1)$$

Here ι_k represents the rotation angle resulting from the k -th toroidal transit. n is the product of the number of toroidal transits traced out by a single field line times the number of independent field lines on that flux surface (for an irrational surface, the latter is obviously one). The angle ι of (4.8.1) is an average over the flux-surface and is by definition a flux-surface quantity.

It is more customary to work with the quantity called the *rotational transform*:

$$\tau = \frac{\iota}{2\pi} . \quad (4.8.2)$$

τ (iota-bar) equals the average number of poloidal transits for one toroidal transit. (In the paper by Solov'ev and Shafranov (1970), the symbol μ is used instead of τ . Their Eqs. (1.7) and (1.8) should read m/n instead of n/m). A rational surface has τ equal to a rational number; on an irrational surface, τ is irrational.

As mentioned above, in the tokamak literature, one prefers to work with the "toroidal winding number," which is the reciprocal of τ :

$$q = \frac{1}{\tau} . \quad (4.8.3)$$

Because of its importance in stability calculations, q is called the *safety factor*.

There is an equivalent and more useful expression for the (average) rotational transform τ :

$$\tau = \frac{d\Psi_{\text{pol}}^r}{d\Psi_{\text{tor}}} = \frac{\dot{\Psi}_{\text{pol}}^r}{\dot{\Psi}_{\text{tor}}} , \quad (4.8.4a)$$

or

$$\tau = - \frac{d\Psi_{\text{pol}}^d}{d\Psi_{\text{tor}}} = - \frac{\dot{\Psi}_{\text{pol}}^d}{\dot{\Psi}_{\text{tor}}} . \quad (4.8.4b)$$

The sign of τ is chosen such that it is positive if the toroidal field and the poloidal field are in the positive ζ and θ directions, respectively (then $\mathbf{B} \cdot \nabla\zeta > 0$, and $\mathbf{B} \cdot \nabla\theta > 0$). The dot represents the derivative with respect to the chosen flux-surface label ϱ :

$$\dot{f}(\varrho) \equiv \frac{df}{d\varrho} . \quad (4.8.5)$$

We should note that for time dependent problems, τ is defined by

$$\tau = \frac{\delta\Psi_{\text{pol}}^r}{\delta\Psi_{\text{tor}}} = \frac{\nabla\Psi_{\text{pol}}^r \cdot d\mathbf{R}}{\nabla\Psi_{\text{tor}} \cdot d\mathbf{R}} . \quad (4.8.6)$$

If we consider a functional relationship, $\Psi_{\text{pol}}^r \equiv \Psi_{\text{pol}}^r(\Psi_{\text{tor}}(\mathbf{R}, t), t)$, where \mathbf{R} is an arbitrary position vector, the chain rule of calculus leads to

$$\nabla\Psi_{\text{pol}}^r = \left. \frac{\partial\Psi_{\text{pol}}^r}{\partial\Psi_{\text{tor}}} \right|_t \nabla\Psi_{\text{tor}} . \quad (4.8.7)$$

Then τ , as defined in (4.8.6) becomes $\tau = \partial\dot{\Psi}_{\text{pol}}^r/\partial\Psi_{\text{tor}}$, (Hinton and Hazeltine (1976), Chap. 7).

The validity of (4.8.4) cannot easily be proved yet, so we shall give an outline which will be justified below. (For simple examples which show the equivalence of our definitions for τ , see Bateman (1978), p. 6 and Morozov and Solov'ev (1966), p. 18.) The crucial point is that the expressions for Ψ_{tor} and $\dot{\Psi}_{\text{pol}}^r$, given in (4.7.10) and (4.7.11), respectively, are valid for any toroidal and poloidal angle ζ and θ , and in particular for a (θ, ζ) coordinate system in which the field lines are straight. When \mathbf{B} is straight in (θ, ζ) , it follows from the definition of τ ($d\theta = \tau d\zeta = 2\pi$) that

$$\frac{d\theta}{d\zeta} = \tau(\varrho) \quad \text{or} \quad \frac{d\zeta}{d\theta} = q(\varrho). \quad (4.8.8)$$

As indicated, τ and q are flux functions (because of its straightness, \mathbf{B} always makes the same angles with the coordinate curves for θ and ζ). The equation of a field line, (4.1.7), requires that we always have

$$\frac{d\theta}{d\zeta} = \frac{B^\theta}{B^\zeta} = \frac{\mathbf{B} \cdot \nabla\theta}{\mathbf{B} \cdot \nabla\zeta}. \quad (4.8.9)$$

In an upcoming section we shall show that, in a straight- \mathbf{B} (θ, ζ) coordinate system:

$$\frac{B^\theta}{B^\zeta} = \frac{\dot{\Psi}_{\text{pol}}^r}{\dot{\Psi}_{\text{tor}}}. \quad (4.8.10)$$

A combination of (4.8.8), (4.8.9) and (4.8.10) gives the required result (4.8.4).

A final comment on the rotational transform. The “well-known” expression for the safety factor q

$$q = \frac{rB_T}{RB_P} \quad (4.8.11)$$

is actually a lowest order approximation. Equation (4.8.11) is only rigorously correct in an axisymmetric tokamak with circular cross-section flux surfaces where r is measured from the center of the circular flux surface. R is the distance from the major axis to a point on the surface, $B_T \equiv (\mathbf{B} \cdot \hat{\zeta}_e)$, and $B_P \equiv (\mathbf{B} \cdot \hat{\theta}_e)$, where θ_e and ζ_e are the “elementary” angles based on the center of the surface.

In almost all toroidal devices, the rotational transform changes from flux surface to flux surface. To measure the amount of change, one uses the concept of *shear* defined simply as

$$\sigma(\varrho) \equiv \frac{d\tau(\varrho)}{d\varrho} = \dot{\tau}(\varrho). \quad (4.8.12)$$

4.9 The Flux-Surface Average

The flux surface average of a function $\Phi(\mathbf{R})$ is defined by the *volume average* over an infinitesimally small shell with volume ΔV , where ΔV lies between two neighboring flux surfaces with volumes V and $V + \Delta V$. It is denoted by $\langle \Phi(\mathbf{R}) \rangle$, and is equal to:

$$\langle \Phi(\mathbf{R}) \rangle \equiv \lim_{\Delta V \rightarrow 0} \frac{\iiint_{\Delta V} \Phi(\mathbf{R}) d^3 R}{\iiint_{\Delta V} d^3 R} = \lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \iiint_{\Delta V} \Phi d^3 R. \quad (4.9.1)$$

There are several ways to write $\langle \Phi \rangle$ in a more practical form. The elemental volume element $d^3 R$ appearing under the integral sign in (4.9.1) is by definition equal to

$$d^3 R = |\mathbf{dR}(\varrho) \cdot \mathbf{dS}(\varrho)|. \quad (4.9.2)$$

The flux surface is assumed to be labeled by $\varrho = \text{constant}$; $\mathbf{dR}(\varrho)$ is the infinitesimal change in the position vector along the ϱ -coordinate curve, while $\mathbf{dS}(\varrho)$ is the elemental area vector of a $\varrho = \text{constant}$ surface. If we orient \mathbf{dS} so that with $\mathbf{dR}(\varrho)$ in the positive ϱ -direction, the dot product of (4.9.2) leads to a positive volume element, we can omit the absolute value bars. $\mathbf{dS}(\varrho)$ is perpendicular to the surface, and then we have

$$dS(\varrho) = \frac{\nabla\varrho}{|\nabla\varrho|} dS. \quad (4.9.3)$$

Thus (4.9.2) reads:

$$d^3 R = \mathbf{dR}(\varrho) \cdot \frac{\nabla\varrho}{|\nabla\varrho|} dS \quad (4.9.4)$$

or with $d\varrho \equiv \nabla\varrho \cdot \mathbf{dR}(\varrho)$:

$$d^3 R = d\varrho \frac{dS}{|\nabla\varrho|}. \quad (4.9.5)$$

Substitution of this expression for $d^3 R$ in the integral of (4.9.1) leads to a *surface integral* as follows:

$$\langle \Phi \rangle = \lim_{\substack{\Delta V \rightarrow 0 \\ \Delta\varrho \rightarrow 0}} \frac{\Delta\varrho}{\Delta V} \iint_{S(\varrho)} \Phi \frac{dS}{|\nabla\varrho|}, \quad (4.9.6)$$

or

$$\langle \Phi \rangle = \frac{d\varrho}{dV} \iint_{S(\varrho)} \Phi \frac{dS}{|\nabla\varrho|}. \quad (4.9.7)$$

dV stands for an elemental volume that is infinitesimal in only one direction, here the ϱ -direction. It is the differential volume between two shells after integration over θ and ζ . In contrast, d^3R is infinitesimal in all three coordinates as the notation suggests. (In fact, we should have used the symbol d^2R instead of dS to be consistent. However, since we deal with only one kind of area element, we kept the symbol as simple as possible, hence dS .)

This form of the flux-surface average, particularly with the pressure as flux-surface label has been used often in the earlier fusion literature. When the flux-surface label is chosen to be V , $d\varrho/dV = 1$, and

$$\langle \Phi \rangle = \iint_{S(V)} \Phi \frac{dS}{|\nabla V|}, \quad \varrho \equiv V. \quad (4.9.8)$$

4.9.1 Toroidal Systems

In arbitrary toroidal coordinates (ϱ, θ, ζ) with Jacobian $\sqrt{g} = (\nabla \varrho \cdot \nabla \theta \times \nabla \zeta)^{-1}$, the surface integral leads to an even more transparent expression. The elemental surface element dS equals $\sqrt{g} |\nabla \varrho| d\theta d\zeta$ according to Chap. 2. Equation (4.9.7) then yields

$$\langle \Phi \rangle = \frac{d\varrho}{dV} \int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta \sqrt{g} \Phi. \quad (4.9.9)$$

This equation provides an expression for $dV/d\varrho$ in terms of an integral over $d\theta$ and $d\zeta$ if we set $\Phi \equiv 1$

$$\frac{dV}{d\varrho} = \int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta \sqrt{g}. \quad (4.9.10)$$

Of course, this also follows from

$$\frac{dV}{d\varrho} = \lim_{\Delta\varrho \rightarrow 0} \frac{\Delta V}{\Delta\varrho} \quad \text{and} \quad \Delta V = \Delta\varrho \iint d\theta d\zeta \sqrt{g}.$$

The flux-surface average of a quantity Φ , expressed as an integral over the poloidal and toroidal angles, becomes:

$$\langle \Phi \rangle = \frac{\int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta \sqrt{g} \Phi}{\int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta \sqrt{g}}. \quad (4.9.11)$$

This reduces to a very simple expression if the flux-surface label is the flux-surface volume, $\varrho = V$:

$$\langle \Phi \rangle = \int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta \sqrt{g} \Phi, \quad \varrho \equiv V. \quad (4.9.12)$$

This clearly also follows from (4.9.9).

In an axisymmetric device, Φ and \sqrt{g} are independent of the toroidal angle ζ ; the integration over ζ provides a factor of 2π in the numerator and the denominator of (4.9.11). For such devices, (4.9.11) simplifies to:

$$\langle \Phi \rangle = \frac{\int_0^{2\pi} d\theta \sqrt{g} \Phi}{\int_0^{2\pi} d\theta \sqrt{g}}. \quad (4.9.13)$$

Thus in an axisymmetric device, θ integration is sufficient to sample conditions on the whole surface.

In axisymmetric machines, the flux-surface average is sometimes referred to as a $\oint dl/B$ average, where the integral is supposed to be performed along closed field lines. In particular, we have

$$\langle \Phi \rangle = \frac{\oint \left(\frac{dl}{B} \right) \Phi}{\oint \frac{dl}{B}}. \quad (4.9.14)$$

To show that (4.9.14) is equivalent to (4.9.13), we must anticipate an important “property” of the magnetic field which will be justified below. The above equations are all valid for arbitrary angles θ and ζ , including those in which the magnetic field lines are straight, whose existence will be shown below. In such a system, we recall from (4.8.8) the simple relationship $d\zeta/d\theta = q(\varrho)$. After integration, this leads to:

$$q\theta - \zeta = c(\varrho) = q\theta_0 \quad (4.9.15)$$

where $c(\varrho)$ is only a function of ϱ . θ_0 is the value of the poloidal angle when $\zeta = 0$ (Fig. 4.17).

On a particular flux surface, $\varrho = \text{constant}$, (4.9.15) represents the equation of a straight line in (θ, ζ) . However, if we let ϱ vary, in general $q(\varrho)$ varies as well, and (4.9.15) becomes the equation of a “ruled” surface¹ traced out by the “straight” B lines (such as the hatched surface in Fig. 4.17). Thus, we have the situation where the field lines lie in the $q\theta - \zeta = q\theta_0$ surfaces. As a consequence, B is perpendicular to $\nabla(q\theta - \zeta)$. From (4.7.13), we recall that B is also perpendicular to $\nabla\varrho$. A vector perpendicular to these two gradients is their vector product. Hence, we can conclude that B must lie along the cross product of $\nabla\varrho$ and

¹A ruled surface is one that is swept out by non-parallel straight lines. The quotation marks on “ruled” are there to indicate that in our case, the surface is not made up of straight lines in the literal sense, but by (field) lines, whose equation is that of a straight line.

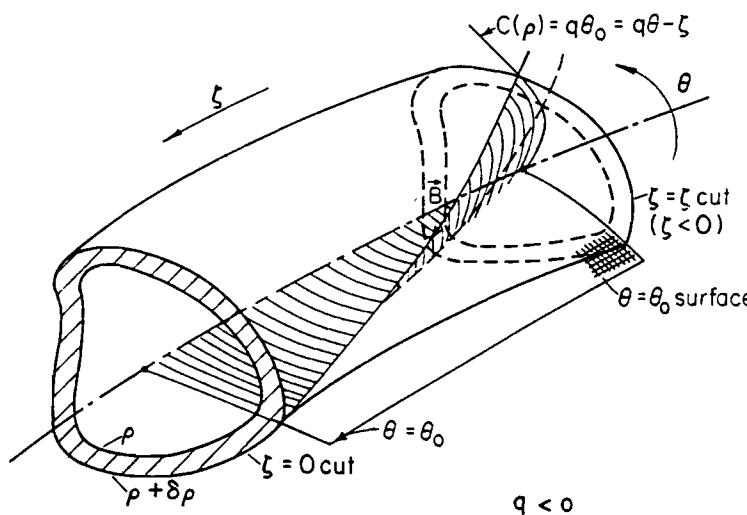


Fig. 4.17. Three-dimensional picture of an angular section of a toroidal surface between $\zeta = 0$ and $\zeta = \zeta_0$. The surfaces $\zeta = \text{constant}$, $\theta = \text{constant}$ and $\zeta - q\theta = \text{constant}$ are shown

$\nabla(q\theta - \zeta)$:

$$\mathbf{B} \propto \nabla\varrho \times \nabla(q\theta - \zeta). \quad (4.9.16)$$

The proportionality constant will be shown below to be a function of ϱ only. The dot product of $\nabla\theta$ with the \mathbf{B} expression of (4.9.16) leads to

$$\begin{aligned} \mathbf{B} \cdot \nabla\theta &\propto [(\nabla\varrho \times \nabla(q\theta)) + (\nabla\zeta \times \nabla\varrho)] \cdot \nabla\theta \\ &\propto \{[\nabla\varrho \times (q\theta\nabla\varrho + q\nabla\theta)] \cdot \nabla\theta + \nabla\varrho \cdot \nabla\theta \times \nabla\zeta\}. \end{aligned} \quad (4.9.17)$$

The first term vanishes because $\nabla\varrho \times (q\theta)\nabla\varrho \equiv 0$, and $q(\nabla\varrho \times \nabla\theta) \cdot \nabla\theta \equiv 0$. The surviving term can be identified with the reciprocal of the Jacobian:

$$\mathbf{B} \cdot \nabla\theta \propto \frac{1}{\sqrt{g}}. \quad (4.9.18)$$

We now go back to the $\oint dl/B$ expression in (4.9.14). The equation of a field line given in (4.1.7) provides us with the equality $dl/B = d\theta/\mathbf{B} \cdot \nabla\theta$. If the loop in (4.9.14) closes after N poloidal transits, the expression for $\langle \Phi \rangle$ can be written as

$$\langle \Phi \rangle = \frac{N \int_0^{2\pi} (d\theta/\mathbf{B} \cdot \nabla\theta) \Phi}{N \int_0^{2\pi} d\theta/\mathbf{B} \cdot \nabla\theta} = \frac{\int_0^{2\pi} d\theta \sqrt{g} \Phi}{\int_0^{2\pi} d\theta \sqrt{g}}. \quad (4.9.19)$$

The proportionality constant of (4.9.18) cancels in this last expression, since it appears in both numerator and denominator and does not depend on θ .

So, in an *axisymmetric* device, a closed-loop integral $\oint dl/B$ is a true flux-surface average. Since it is equivalent to an integration over θ , according to (4.9.13), it is not even necessary to close the loop; covering one period in θ suffices. Therefore, a $\int \Phi dl/B \sim \int_0^{2\pi} d\theta \sqrt{g} \Phi$ flux-surface average can be evaluated on *any* flux surface. If one insists on performing a *closed-loop* integral, however, it is clear that it can only be done on rational surfaces. In that case, one can invoke the fact that every irrational surface can be approximated arbitrarily closely by a rational one, so that *closed loop* flux-surface averaging can still be used for "essentially" any flux surface.

The situation is different in *non-axisymmetric* systems, however. A *closed-loop* integral performed on a particular rational surface leads to different results depending on the field line we start from. Hence it cannot be a flux-surface average. If we insist on using a dl/B integration to find the flux-surface average, we have two options. Either we take an average of several closed-loop integrals, each "belonging to" different closed field lines; or, we perform a $\int (dl/B)\Phi$ integration on an irrational surface and follow the non-closed field line long enough until, for all practical purposes, we have adequately covered the whole surface. Because rational and irrational surfaces can approximate each other infinitesimally closely, the methods are equivalent.

4.9.2 Open-Ended Systems

In mirror devices, a $\int dl/B$ average is insufficient because a field line does not cover the whole flux surface; it merely runs from L_l to L_r (l for left, and r for right). In addition to the l integration, it is necessary to integrate over the angular coordinate as well. The latter is commonly denoted by β and is usually a Clebsch coordinate (see below). In a (ϱ, β, l) system, we have $dS = \sqrt{g_C} |\nabla\varrho| dl d\beta$, where $\sqrt{g_C} = (\nabla\varrho \cdot \nabla\beta \times \nabla l)^{-1} = (l \cdot \nabla\varrho \times \nabla\beta)^{-1} = 1/B$. The C subscript on g stands for Clebsch, and the $1/B$ result will be shown below in the section on Clebsch coordinates. Starting from (4.9.7), we have that for mirror devices:

$$\langle \Phi \rangle = \frac{d\varrho}{dV} \iint_{S(\varrho)} \Phi \sqrt{g_C} dl d\beta = \frac{d\varrho}{dV} \int_0^{2\pi} d\beta \int_{L_1}^{L_r} \frac{dl}{B} \Phi \quad (4.9.20)$$

with

$$dV = d\varrho \iint_{S(\varrho)} \sqrt{g_C} dl d\beta = d\varrho \int_0^{2\pi} d\beta \int_{L_1}^{L_r} \frac{dl}{B}. \quad (4.9.21)$$

We then find that in mirror devices:

$$\langle \Phi \rangle = \frac{\int_0^{2\pi} d\beta \int_{L_1}^{L_r} \Phi dl/B}{\int_0^{2\pi} d\beta \int_{L_1}^{L_r} dl/B} = \frac{1}{dV/d\varrho} \int_0^{2\pi} d\beta \int_{L_1}^{L_r} \Phi \frac{dl}{B}. \quad (4.9.22)$$

In general, L_l and L_r are dependent on both ϱ and β .

4.9.3 Properties of the Flux-Surface Average

The flux surface average of the divergence of a vector \mathbf{A} can be computed by applying Gauss' law to an infinitesimally thin shell of flux volume. In *toroidal* devices, this is done after making a cut at a certain ζ -constant surface (similarly to Fig. 4.16)

$$\langle \nabla \cdot \mathbf{A} \rangle = \lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \iiint_{\Delta V} (\nabla \cdot \mathbf{A}) d^3 R = \lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \oint \mathbf{A} \cdot d\mathbf{S}. \quad (4.9.23)$$

The closed surface integral consists of three pieces. Firstly, there is the surface $S_{\Delta V}$ made up of the two toroidal flux surfaces enclosing the volume ΔV , $S_{\Delta V} = S(V) + S(V + \Delta V)$. Secondly, there are two surfaces in the poloidal "plane" $\zeta = \text{constant}$. The contribution of the last two vanishes because their normals are opposite in direction such that the flux of \mathbf{A} that "flows out" of one surface "flows in" through the other.

When we computed expressions for the flux in (4.7.8), the contributions were reversed. There was no contribution through the toroidal surface in (4.7.8) because the vector in question was \mathbf{B} and there the contribution of the cuts survived because the integrand was multiple valued, containing a ζ . Here, in (4.9.23), on the other hand, the vector \mathbf{A} must be single valued, since it represents a physical quantity. Thus, the integrand in (4.9.23) has the same value at $\zeta = \zeta_0$ and $\zeta = \zeta_0 + 2\pi$, as a consequence of which the "cut"-contributions vanish.

If we also replace $d\mathbf{S}$ in (4.9.23), using (4.9.3), but with $\varrho = V$, we obtain

$$\langle \nabla \cdot \mathbf{A} \rangle = \lim_{\Delta V \rightarrow 0} (\pm) \frac{1}{\Delta V} \iint_{S_{\Delta V}} \frac{\mathbf{A} \cdot \nabla V}{|\nabla V|} dS. \quad (4.9.24)$$

The integral over $S_{\Delta V}$ in this expression stands for the sum of two integrals, one over $S(V)$ and one over $S(V + \Delta V)$. For the outer surface, we must choose the $+$ sign, since ∇V points outward; for the inner surface, we must take the $-$ sign ($d\mathbf{S}$ in Gauss' theorem is outward with respect to the closed surface to which it applies, i.e., $d\mathbf{S}$ is outward for $S(V + \Delta V)$ but inward for $S(V)$). It actually gives a difference rather than a sum in this case. The limit of this small difference represents the derivative of an integral over $S(V)$ with respect to V

$$\langle \nabla \cdot \mathbf{A} \rangle = \frac{d}{dV} \iint_{S(V)} \frac{\mathbf{A} \cdot \nabla V}{|\nabla V|} dS. \quad (4.9.25)$$

Comparing this expression with (4.9.8), we find

$$\langle \nabla \cdot \mathbf{A} \rangle = \frac{d}{dV} \langle \mathbf{A} \cdot \nabla V \rangle. \quad (4.9.26)$$

In terms of an arbitrary flux label ϱ , this becomes

$$\langle \nabla \cdot \mathbf{A} \rangle = \frac{d\varrho}{dV} \frac{d}{d\varrho} \left\{ \frac{dV}{d\varrho} \langle \mathbf{A} \cdot \nabla \varrho \rangle \right\}, \quad (4.9.27)$$

where we made use of the fact that

$$\frac{d}{dV} = \left(\frac{d\varrho}{dV} \right) \frac{d}{d\varrho} \quad \text{and} \quad \nabla V = \frac{dV}{d\varrho} \nabla \varrho.$$

We also made use of the obvious fact that a flux-surface quantity can be taken out of the $\langle \rangle$ bracket.

The proof for *open-ended systems* proceeds analogously, with an important distinction. In contrast to toroidal systems, the "end contributions" to the surface integral in (4.9.23) do not vanish, as least in the general case.

The surface integrals at the end of the infinitesimal shell volume are most easily computed in the (ϱ, β, l) coordinate system. If we bound the shell volume by $l = \text{constant}$ surfaces, then (2.5.49) is the suitable expression for $d\mathbf{S}$, in which $(g_c)^{1/2} = 1/B$. This gives for (4.9.23):

$$\begin{aligned} \langle \nabla \cdot \mathbf{A} \rangle &= \frac{d\varrho}{dV} \frac{d}{d\varrho} \left\{ \frac{dV}{d\varrho} \langle \mathbf{A} \cdot \nabla \varrho \rangle \right\} \\ &+ \lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \int_{\varrho}^{\varrho + d\varrho} d\varrho \int_0^{2\pi} \frac{d\beta}{B} \mathbf{A} \cdot \frac{\nabla l}{|\nabla l|} \Big|_{\text{at } l=L_r} \\ &- \lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \int_{\varrho}^{\varrho + d\varrho} d\varrho \int_0^{2\pi} \frac{d\beta}{B} \mathbf{A} \cdot \frac{\nabla l}{|\nabla l|} \Big|_{\text{at } l=L_i} \end{aligned} \quad (4.9.28)$$

Written somewhat more compactly, we have

$$\langle \nabla \cdot \mathbf{A} \rangle = \frac{1}{V} \frac{d}{d\varrho} \{ \dot{V} \langle \mathbf{A} \cdot \nabla \varrho \rangle \} + \frac{1}{V} \int_0^{2\pi} \frac{d\beta}{B} \mathbf{A} \cdot \frac{\nabla l}{|\nabla l|} \Big|_{l=L_r} \quad (4.9.29)$$

Here $\dot{V} \equiv dV/d\varrho$. Upon using the definition of the flux-surface average in open-ended devices, (4.9.22), we obtain

$$\langle \nabla \cdot \mathbf{A} \rangle = \frac{1}{V} \frac{d}{d\varrho} \{ \dot{V} \langle \mathbf{A} \cdot \nabla \varrho \rangle \} + \left\langle B \frac{\partial}{\partial l} \left(\frac{\mathbf{A} \cdot \nabla l}{|\nabla l|} \right) \right\rangle, \quad (4.9.30)$$

a somewhat more general form than that given by Catto and Myra (1986) (P.J. Catto, private communication).

In general, ∇l is different from $\hat{\mathbf{B}}$. However, if the geometry is such that the $l = \text{constant}$ surfaces (at the ends of the device) can be constructed to be perpen-

icular to the field lines, then $\nabla l = \hat{\mathbf{B}}$ and one gets:

$$\langle \nabla \cdot \mathbf{A} \rangle = \frac{1}{V} \frac{d}{d\varrho} \{ \dot{V} \langle \mathbf{A} \cdot \nabla \varrho \rangle \} + \left\langle \mathbf{B} \frac{\partial}{\partial l} \left(\frac{\mathbf{A} \cdot \hat{\mathbf{B}}}{B} \right) \right\rangle . \quad (4.9.31)$$

The equality of ∇l and $\hat{\mathbf{B}}$, in order for (4.9.31) to hold, is perhaps not necessary. Since the second term in (4.9.30) and (4.9.31) concerns averaging over the “angle” β , it might be that the averages are identical in certain specific cases, even though $\hat{\mathbf{B}}$ differs from ∇l on the surfaces on which the integrals are performed.

Furthermore, for a particular vector \mathbf{A} and/or geometry, this second term might be shown to vanish. If the flux of the vector field \mathbf{A} over the infinitesimal annular surface at $l = L_1$ exactly cancels that at $l = L_r$, the end contributions disappear and (4.9.27) applies. (For more on the relationship between strict quadrupole symmetry and end losses in tandem mirrors, the reader should consult the papers by Myra and Catto (1982a) and by Myra, Catto and Hazeltine (1982b).)

By far the most important property of the flux-surface average is that it annihilates the $(\mathbf{B} \cdot \nabla)$ operator:

$$\langle \mathbf{B} \cdot \nabla \Phi \rangle \equiv 0 . \quad (4.9.32)$$

This property may be used to define the $\langle \cdot \rangle$ operation. Here, we prove it with the help of the expression derived above for $\langle \text{div} \rangle$, (4.9.27). It follows from $\nabla \cdot \mathbf{B} = 0$ and (4.9.27) that

$$\langle \mathbf{B} \cdot \nabla \Phi \rangle = \langle \nabla \cdot (\mathbf{B} \Phi) \rangle \propto \langle \Phi \mathbf{B} \cdot \nabla \varrho \rangle , \quad (4.9.33)$$

which vanishes identically because $\mathbf{B} \cdot \nabla \varrho = 0$ is the equation of a flux surface.

More intuitively, in a toroidal device,

$$\langle \mathbf{B} \cdot \nabla \Phi \rangle = \left\langle \mathbf{B} \frac{\partial \Phi}{\partial l} \right\rangle \sim \int \frac{dl}{B} \mathbf{B} \frac{\partial \Phi}{\partial l} = \int dl \frac{\partial \Phi}{\partial l} = 0 . \quad (4.9.34)$$

For an axisymmetric device, $\int dl$ stands for $\oint dl$ on the nearest rational surface so the null result is obvious. For a non-axisymmetric device, we assume that Φ is continuous over the surface, we pick an irrational surface and go around as many times as required to sample the surface well enough for practical purposes and we reach a point that is sufficiently close to the starting point. The result can be made as small as desired, with zero being the limit.

For mirrors, $\langle \mathbf{B} \cdot \nabla \Phi \rangle$ vanishes only if the end contributions cancel. We have:

$$\begin{aligned} \langle \mathbf{B} \cdot \nabla \Phi \rangle &= \left\langle \mathbf{B} \frac{\partial \Phi}{\partial l} \right\rangle \sim \int_0^{2\pi} d\beta \int_{L_1}^{L_r} dl \frac{\partial \Phi}{\partial l} \\ &\sim \int_0^{2\pi} \Phi(\beta, L_r) d\beta - \int_0^{2\pi} \Phi(\beta, L_1) d\beta . \end{aligned} \quad (4.9.35)$$

Equality implies that the average of Φ over β must be the same at $l = L_1$ and $l = L_r$.

Note that although $\langle \mathbf{B} \cdot \nabla \Phi \rangle \equiv 0$ (in toroidal devices), we have, in general that $\langle \hat{\mathbf{B}} \cdot \nabla \Phi \rangle \neq 0$, where $\hat{\mathbf{B}} = \mathbf{B}/B$ is the unit vector along \mathbf{B} . Thus $\langle (\mathbf{B} \cdot \nabla \Phi)/B \rangle \neq 0$. Also, in general, the following expression is non zero:

$$\langle \gamma(\mathbf{R}) \mathbf{B} \cdot \nabla \Phi(\mathbf{R}) \rangle = \langle \gamma \mathbf{B} \cdot \nabla \Phi \rangle \neq 0 . \quad (4.9.36)$$

$\gamma(\mathbf{R})$ is a scalar field analogous to $\Phi(\mathbf{R})$. However, (4.9.36) reduces identically to zero if $\gamma = \Phi^{-1}$

$$\left\langle \frac{\mathbf{B} \cdot \nabla \Phi}{\Phi} \right\rangle = \langle \mathbf{B} \cdot \nabla \ln \Phi \rangle = \langle \mathbf{B} \cdot \nabla \Omega \rangle \equiv 0 , \quad (4.9.37)$$

where $\Omega = \ln \Phi$.

A final expression concerning the flux-surface average is the identity (for toroidal devices)

$$\langle \nabla \varrho \cdot \nabla \times \mathbf{W} \rangle \equiv 0 , \quad (4.9.38)$$

in which ϱ is a flux-surface label. This equation is a consequence of the identity $\nabla \cdot (\mathbf{W} \times \nabla \varrho) = \nabla \varrho \cdot \nabla \times \mathbf{W} - \mathbf{W} \cdot \nabla \times \nabla \varrho = \nabla \varrho \cdot \nabla \times \mathbf{W}$, and the $\langle \text{div} \rangle$ result of (4.9.27)

4.10 The Magnetic Differential Equation

An equation of the form

$$\mathbf{B} \cdot \nabla f = S \quad (4.10.1)$$

is the prototype magnetic differential equation (Kruskal and Kulsrud 1958; Newcomb 1959; see also, Hazeltine and Meiss 1985, and Solov'ev and Shafranov 1970). S is a source term, and f is the unknown single-valued function that we must solve for. Both f and S are scalar fields: $f \equiv f(\mathbf{R})$ and $S \equiv S(\mathbf{R})$.

We limit our discussion to toroidal devices because this type of equation poses fewer problems in linear devices.

A first condition on S in order that f be single valued follows from an integration over l :

$$\mathbf{B} \cdot \nabla f = \mathbf{B} \frac{\partial f}{\partial l} = S , \quad (4.10.2)$$

which implies

$$f = f_0 + \int_0^l \frac{dl}{B} S . \quad (4.10.3)$$

Here, f_0 is the value of f at $l = 0$, an arbitrary origin. If we happen to perform the integral in (4.10.3) on a rational surface, we get a closed integral, and the single valuedness of f requires that

$$\oint \frac{dl}{B} S = 0 \quad (4.10.4)$$

on rational surfaces. Equation (4.10.4) must be satisfied for every closed field line on the rational surface.

A second solvability condition follows from taking the flux-surface average of both sides of the magnetic differential equation and applying the identity (4.9.32)

$$\langle \mathbf{B} \cdot \nabla f \rangle = \langle S \rangle \equiv 0 . \quad (4.10.5)$$

A careful examination shows that these two conditions are not independent (if one assumes, which we do throughout this monograph, that the rational surfaces are also a dense set in non-axisymmetric devices, as they are for the “imaginary” lowest-order field structure with “good” surfaces, that we assume throughout this monograph). Equation (4.10.5) follows from (4.10.4). If the requirement given in (4.10.4) is satisfied for *every* field line on the surface, then it must be concluded that the flux-surface average of S is automatically zero on that surface. This is obvious if we write the flux-surface average on the rational surface as

$$\langle S \rangle = \frac{\oint \frac{dl}{B} S + \oint \frac{dl}{B} S + \oint \frac{dl}{B} S + \cdots + \oint \frac{dl}{B} S + \cdots}{\oint \frac{dl}{B} + \oint \frac{dl}{B} + \oint \frac{dl}{B} + \cdots + \oint \frac{dl}{B} + \cdots} = 0 , \quad (4.10.6)$$

where the summation runs ad infinitum over all possible field lines in the surface. Since we then have that $\langle S \rangle$ is zero on every rational surface, continuity and the fact that the real numbers are everywhere dense mean that $\langle S \rangle$ vanishes everywhere. The reverse statement is not valid. The fact that the average of S vanishes does not, in general, guarantee that all the pieces that make up that average vanish. In other words, the fact that $\sum \oint S dl / B$ vanishes does not imply that every $\oint S dl / B$ vanishes. This would be true only for an axisymmetric device such as a tokamak which has $\langle S \rangle = \oint S dl / B / \oint dl / B$. Thus, (4.10.5) cannot be a sufficient condition to guarantee a solution of the differential equation.

The definition of the flux-surface average allows us to write an alternate expression for (4.10.5),

$$\lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \iiint_S d^3 R \equiv 0 . \quad (4.10.7)$$

It was in this form that the condition was written by Kruskal and Kulsrud (1958), and that it was discussed by Newcomb (1959). Although (4.10.5) and (4.10.7) follow from (4.10.4), it is a convenient condition to impose in order to eliminate certain features of S , after which the application of (4.10.4) is made easier.

It was Newcomb (1959) who showed that (4.10.4) is a sufficient condition for solving the magnetic differential equation. We shall demonstrate this by finding a solution for f after imposing the condition.

As usual in the theory of differential equations, the general solution of (4.10.1) consists of two parts, a homogeneous solution and a particular solution

$$f = f_h + f_p , \quad (4.10.8)$$

where f_h satisfies the homogeneous equation without a source term

$$\mathbf{B} \cdot \nabla f_h = 0 . \quad (4.10.9)$$

This is equivalent to

$$\frac{\partial f_h}{\partial l} = 0 , \quad (4.10.10)$$

which means that f_h must be a flux function. (This conclusion can only be drawn in toroidal machines; in mirrors, all we can say is that f_h is independent of l .) Hence, the total solution of (4.10.1) equals

$$f = f_h(\varrho) + f_p . \quad (4.10.11)$$

Any flux-surface quantity satisfies (4.10.10) and is thus eligible as a homogeneous solution for the magnetic differential equation.

To find a particular solution, the differential equation may be rewritten in terms of partial derivatives $\partial/\partial\theta$ and $\partial/\partial\zeta$. It is most convenient to do so in flux coordinates (θ_f, ζ_f) in which the magnetic field is straight. At this moment, we do not yet have the tools to actually solve the equation. However, the process is straightforward, and we can give an outline. Details can be found in the review by Solov'ev and Shafranov (1970).

The $\mathbf{B} \cdot \nabla f$ term becomes $B^i \partial f / \partial u^i$, where u^i stands for θ_f or ζ_f , since $B^e \equiv \mathbf{B} \cdot \nabla \varrho = 0$. The B^i in flux coordinates have a $\sqrt{g_f}$ in the denominator (see below) which upon bringing it to the right-hand side, leads to a source term $\sqrt{g_f} S$. Next, the functions f and $\sqrt{g_f} S$, which are both single valued, are expanded in a double Fourier series in θ_f and ζ_f . The two integrands in the two solvability conditions (4.10.4) and (4.10.5) contain the factor $\sqrt{g_f} S$ as well. (In the integral $\oint (dl/B) S$, we use the equation of a field line $dl/B = d\zeta_f / B^e \propto \sqrt{g_f}$, and in the volume integral, $d^3 R = \sqrt{g_f} d\varrho d\zeta_f d\theta_f$.) Then we use the equation for a field line on a flux surface in flux coordinates $\theta_{f0} = \theta_f - i\zeta_f$ (see (4.9.15)) to eliminate the θ_f variable and impose the two integral constraints.

These conditions have consequences for the expansion coefficients of $\sqrt{g_f} S$, which are denoted by $a_{mn}(\varrho)$. The volume-integral condition requires that

$a_{00}(\varrho)$ be zero on any flux surface. The $\oint S dl/B = 0$ condition states that $\sum_{mn} a_{mn}(\varrho)/(m\tau - n) = 0$, so that the $a_{mn}(\varrho_{\text{rational}})$ must vanish on their respective rational surfaces where $m\tau - n = 0$. When these conditions are satisfied, the $a_{mn}(\varrho_{\text{rational}})$ are factored as $(m\tau - n)\tilde{a}_{mn}(\varrho_{\text{rational}})$, and the solution is found by equating the respective harmonics of the Fourier expansion. Thus, $f_{mn} = (\)\tilde{a}_{mn}$, where $()$ represents a constant factor independent of m or n . This completes the proof because f can be reconstructed from its harmonics.

From this Fourier-expansion procedure, it is also evident that the $\langle S \rangle = 0$ condition follows from the $\oint S dl/B = 0$ requirement. The latter restriction, $\sum_{mn} a_{mn}(\varrho)/(m\tau - n) = 0$, not only forces $a_{mn}(\varrho_{\text{rational}}) = 0$, but also sets $a_{00}(\varrho) = 0$ if $m = n = 0$.

In a tokamak, f and S are independent of ζ , so that their Fourier expansion is only with respect to θ . Therefore, $\sum_m a_{m0}(\varrho)/m\tau = 0$, requires only that $a_{00}(\varrho) = 0$ (independent of the value of τ), leading to the observation that both conditions reduce to one and the same thing.

For the numerical solution of the magnetic differential equation, the reader is referred to the paper by Reiman and Greenside (1988).

Flux Coordinates

From a purely mathematical standpoint, Part II, on flux coordinates, can be considered merely as an extensive example of non-orthogonal curvilinear coordinate systems with certain particular properties. To recall, *magnetic-flux coordinates* are curvilinear coordinates, so chosen that *the equation of a magnetic-field line is that of a straight line*. Somebody familiar with the topics covered in Part I should have no difficulties in this part, at least from a computational standpoint. Conversely, it is likely to be very difficult to obtain an adequate understanding of this part if the previous material has not been mastered.

Providing a discussion of the computational tools for magnetic flux coordinate descriptions is the primary reason for the writing of this monograph. However, for plasma and fusion scientists, the physical interpretation of the subject matter is equally important i.e., “In what physical situation do we use which coordinate system?” Therefore, besides explaining most mathematical steps, we have tried to emphasize the relationships between the various coordinate systems. Thus, it is even instructive to discuss links with a few non-flux coordinate systems that are frequently used or have their own intrinsic value.

The emphasis in this flux-coordinate treatment is on *toroidal devices*. From a conceptual standpoint, understanding the relationships between the different types of flux coordinates in toroidal devices is quite intricate; in contrast, in open-ended systems there is only one fairly simple system that is used, namely the Clebsch system.

We nevertheless discuss *mirrors* whenever we feel this is helpful. The very general Clebsch system is one of the topics we will discuss in considerable detail. Since it is based on the properties of the magnetic field \mathbf{B} , independent of a particular device, it is useful not only in mirrors and toroids in a fusion context, but also for the description of extraterrestrial magnetic fields in *space physics*, *astronomy*, and *astrophysics*.

As far as toroidal devices are concerned, the discussions are kept completely general with regard to the type of device. To make the treatment less abstract, however, we interpret the general results by means of two prototypes. For machines with a symmetry direction, we consider the axisymmetric tokamak. For general three-dimensional devices, we relate the results to the stellarator family of devices.

In Chap. 5, we discuss Clebsch-related coordinates. There we define stream functions, treat the Clebsch coordinate system itself, point out its relation to the contra- and covariant component formalism, and discuss a closely-related system that we give the name “Boozer-Grad”. Toroidal flux coordinates are considered in Chap. 6. Here the stage is set for straight- \mathbf{B} -line coordinates, non-flux coordinates in tokamaks, straight- \mathbf{J} -line coordinates, covariant- \mathbf{B} components and the relationship to the Boozer-Grad form, Boozer’s toroidal flux coordinates, the ideal-MHD equilibrium conditions for toroidally confined plasmas, and Hamada coordinates. In Chap. 7, the focus is on the conversion of Clebsch-type coordinates to toroidal flux coordinates. We do this for the Clebsch and Boozer-Grad systems. The following Chap. 8 is on the establishment of the flux-

coordinate transformation. It summarizes how flux coordinates are found and allows us to say a few words about equilibrium calculations, including useful references. Finally, in Chap. 9, we treat a more general but related topic, namely, canonical coordinates. These coordinates superficially resemble the flux-coordinate representation and are intimately related to the existence of flux surfaces. Flux coordinates are reconsidered in a slightly different form to make their evaluation readily comparable with that of the canonical coordinates. The canonical coordinates and the magnetic field-line Hamiltonian are presented, after which the practical evaluation of the Hamiltonian and the canonical coordinates is explained.

5. The Clebsch-Type Coordinate Systems

5.1 Stream Functions

The idea of a stream function (or flow function) is borrowed from hydrodynamic theory (see, e.g., Milne-Thomson (1968)). A stream function is any function that is constant along a stream line. In plasma physics, we are interested in stream functions for the magnetic field. Hence, in our context, a stream function is *any function that is constant along a field line*. In three-dimensional space, the equation $\Phi(u^1, u^2, u^3) \equiv \Phi(\mathbf{R}) = \text{constant}$ represents a surface. A magnetic-field line can thus be identified with a curve lying at the intersection of two surfaces belonging to two different stream functions.

An interesting property of stream functions $\alpha(\mathbf{R})$ and $\beta(\mathbf{R})$ to be defined below in (5.2.1), is that the differential magnetic flux through a surface (other than the $\alpha = \text{constant}$ and $\beta = \text{constant}$ surfaces) is proportional $d\alpha d\beta$, or

$$\Psi \propto \int \int d\alpha d\beta . \quad (5.1.1)$$

We will show below that this property is consistent with the definition of stream functions.

5.2 Generic Clebsch Coordinates

The Clebsch coordinates also have their origin in fluid mechanics, (e.g., Milne-Thomson 1968; Bedaux 1980). They are related to what is known in hydrodynamics as the Clebsch transformations (Lamb 1945). They were introduced to magnetic-fusion theory by Kruskal and Kulsrud (1958) and by Grad and Rubin (1958). These Clebsch coordinates, commonly denoted by α and β are stream functions of the magnetic field and allow us to write \mathbf{B} as the cross product of their gradients

$$\mathbf{B} = \nabla\alpha \times \nabla\beta . \quad (5.2.1)$$

An intuitive way of obtaining (5.2.1) follows from writing \mathbf{B} in terms of a vector potential \mathbf{A}

$$\mathbf{B} = \nabla \times \mathbf{A} . \quad (5.2.2)$$

Equation (5.2.2) is a consequence of $\nabla \cdot \mathbf{B} = 0$. If we now pick two functions α and β such that

$$\mathbf{A} = \alpha \nabla\beta , \quad (5.2.3)$$

we obtain for (5.2.2)

$$\mathbf{B} = \nabla \times (\alpha \nabla\beta) = (\nabla\alpha \times \nabla\beta) + \alpha (\nabla \times \nabla\beta) . \quad (5.2.4)$$

Since the last term vanishes identically because of a vector identity, we recover (5.2.1).

A constructive proof of (5.2.1) shows how to find functions α and β (Grad and Rubin 1958; see also, Bateman 1978; Kulsrud 1983; Bernstein 1971). Consider a (transverse) surface S that is intersected by \mathbf{B} -field lines (Fig. 5.1a) – thus S is nowhere tangent to \mathbf{B} . In this surface S , we choose a family of lines that we label by $\tilde{\alpha} = \text{constant}$. Any of these lines, take for instance $\tilde{\alpha} = \tilde{\alpha}_0$ (Fig. 5.1b), can be viewed as the intersection of the surface S with a surface traced out

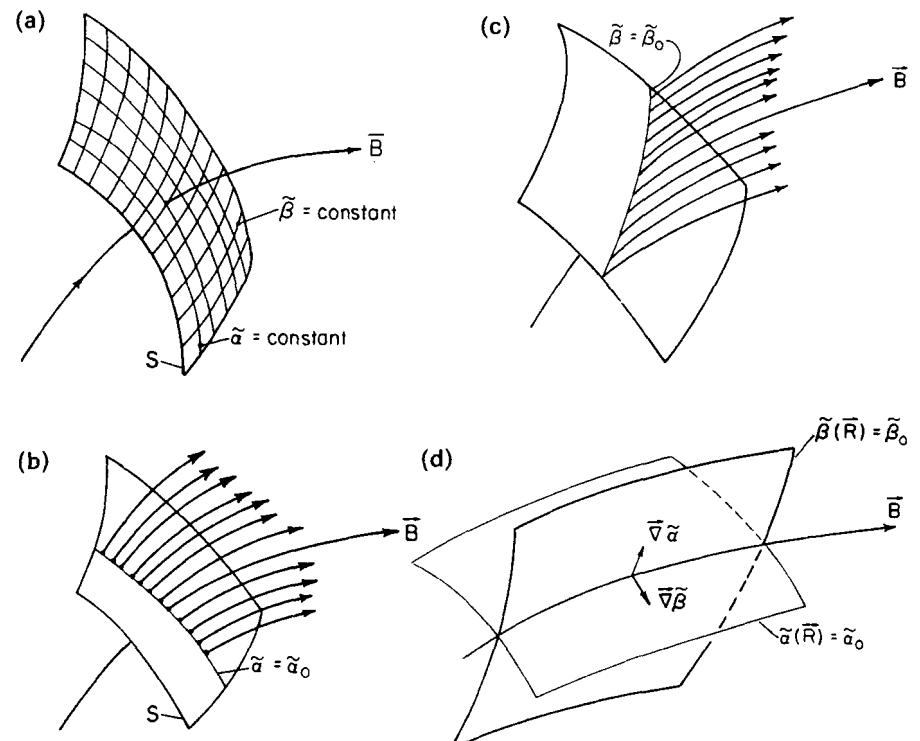


Fig. 5.1a-d. Construction of an (α, β) Clebsch coordinate system. (a) $\tilde{\alpha}, \tilde{\beta}$ grid laid out on a surface S intersecting a field line \mathbf{B} . (b) $\tilde{\alpha} = \tilde{\alpha}_0$ surface, being constructed from the magnetic-field lines intersecting the curve $\tilde{\alpha} = \tilde{\alpha}_0$ on S . (c) $\tilde{\beta} = \tilde{\beta}_0$ surface, constructed from the field lines intersecting the curve $\tilde{\beta} = \tilde{\beta}_0$ on S . (d) Representation of the $\tilde{\alpha} = \tilde{\alpha}_0$ and $\tilde{\beta} = \tilde{\beta}_0$ surfaces based on a magnetic-field line. The gradients $\nabla\tilde{\alpha}$ and $\nabla\tilde{\beta}$ to the respective surfaces are also shown

by all the magnetic-field lines passing through $\tilde{\alpha} = \tilde{\alpha}_0$. The latter surface is labeled by carrying the value $\tilde{\alpha} = \tilde{\alpha}_0$ off S on the field lines. Hence, the equation of the surface is $\tilde{\alpha} = \tilde{\alpha}_0$. Analogously, we construct another family of lines $\tilde{\beta} = \text{constant}$ in S (Fig. 5.1a) such that the $\tilde{\alpha}$ and $\tilde{\beta}$ lines are nowhere tangent. In other words, the $\tilde{\alpha}$ and $\tilde{\beta}$ lines must form a proper grid in S . Again, as shown in Fig. 5.1c, a particular line $\tilde{\beta} = \tilde{\beta}_0$ is carried off on the field lines, resulting in a surface $\tilde{\beta} = \tilde{\beta}_0$. When we repeat this same procedure for all the field lines, we can thus construct functions $\tilde{\alpha} = \tilde{\alpha}(\mathbf{R})$ and $\tilde{\beta} = \tilde{\beta}(\mathbf{R})$, such that the geometric loci of points of $\tilde{\alpha} = \text{constant}$ and $\tilde{\beta} = \text{constant}$ are precisely the types of surfaces discussed above. By construction, the \mathbf{B} lines lie in $\tilde{\alpha} = \text{constant}$ and $\tilde{\beta} = \text{constant}$ surfaces. In that respect, these surfaces can be called “magnetic surfaces”. However, in fusion plasma physics, it is customary to reserve this name for toroidal or cylindrical flux surfaces.

A particular magnetic-field line is defined by the intersection of $\tilde{\alpha} = \text{constant}$ and $\tilde{\beta} = \text{constant}$ surfaces (as is shown in Fig. 5.1d). We can thus interpret the pair of values $\tilde{\alpha}_0$ and $\tilde{\beta}_0$ as the coordinates of the particular field line. A point in space can then be determined by specifying $\tilde{\alpha}$, $\tilde{\beta}$ and a third variable which is related to the length along a field line.

By construction,

$$\mathbf{B} \perp \nabla \tilde{\alpha} \quad \text{or} \quad \mathbf{B} \cdot \nabla \tilde{\alpha} = 0 \quad (5.2.5a)$$

$$\mathbf{B} \perp \nabla \tilde{\beta} \quad \text{or} \quad \mathbf{B} \cdot \nabla \tilde{\beta} = 0 . \quad (5.2.5b)$$

This states that $\tilde{\alpha}$ and $\tilde{\beta}$ are magnetic stream functions. Another vector that is everywhere orthogonal to both $\nabla \tilde{\alpha}$ and $\nabla \tilde{\beta}$ is their vector product. We must thus have that \mathbf{B} is parallel to $\nabla \tilde{\alpha} \times \nabla \tilde{\beta}$ or

$$\mathbf{B} = C \nabla \tilde{\alpha} \times \nabla \tilde{\beta} , \quad (5.2.6)$$

where C is a proportionality factor. In general, C can vary from point to point. It could thus be a function of $\tilde{\alpha}$, $\tilde{\beta}$ and the third variable which we take for simplicity equal to the length l along a field line. However, from $\nabla \cdot \mathbf{B} = 0$, it follows that

$$0 = \nabla \cdot [C \nabla \tilde{\alpha} \times \nabla \tilde{\beta}] = \nabla C \cdot (\nabla \tilde{\alpha} \times \nabla \tilde{\beta}) + C \nabla \cdot (\nabla \tilde{\alpha} \times \nabla \tilde{\beta}) , \quad (5.2.7)$$

where the last term vanishes because it involves expressions of the form $\nabla \times \nabla f$ which vanish identically. Thus, we have that

$$\nabla C \cdot (\nabla \tilde{\alpha} \times \nabla \tilde{\beta}) = 0 . \quad (5.2.8)$$

With the help of the expression for \mathbf{B} in (5.2.6), we obtain

$$\begin{aligned} \frac{\nabla C}{C} \cdot \mathbf{B} = 0 \quad \text{or} \quad \mathbf{B} \cdot \nabla C = 0 \\ \text{or} \quad \mathbf{B} \frac{\partial C}{\partial l} = 0 . \end{aligned} \quad (5.2.9)$$

This states that C is constant along the field lines, and thus is independent of l . Hence, $C = C(\tilde{\alpha}, \tilde{\beta})$ alone.

We now redefine $\tilde{\alpha}$ and $\tilde{\beta}$ such that the factor C is absorbed in the new α and β :

$$\alpha = \tilde{\alpha} \quad (5.2.10a)$$

$$\beta = \int_0^{\tilde{\beta}} d\tilde{\beta} C(\tilde{\alpha}, \tilde{\beta}) \quad \text{or} \quad \frac{\partial \beta}{\partial \tilde{\beta}} = C(\tilde{\alpha}, \tilde{\beta}) . \quad (5.2.10b)$$

It follows from these definitions that

$$\nabla \alpha = \nabla \tilde{\alpha} \quad (5.2.11a)$$

and

$$\nabla \beta = \frac{\partial \beta}{\partial \tilde{\alpha}} \nabla \tilde{\alpha} + \frac{\partial \beta}{\partial \tilde{\beta}} \nabla \tilde{\beta} . \quad (5.2.11b)$$

With the replacement (5.2.10b) for $\partial \beta / \partial \tilde{\beta}$, we can thus write for the \mathbf{B} expression of (5.2.6):

$$\mathbf{B} = C \nabla \tilde{\alpha} \times \nabla \tilde{\beta} = \nabla \tilde{\alpha} \times C \nabla \tilde{\beta} = \nabla \tilde{\alpha} \times \left[\nabla \beta - \frac{\partial \beta}{\partial \tilde{\alpha}} \nabla \tilde{\alpha} \right] = \nabla \alpha \times \nabla \beta . \quad (5.2.12)$$

Note that our discussion does not require $\alpha(\mathbf{R})$ and $\beta(\mathbf{R})$ to be single valued functions. However, \mathbf{B} , which is a physical quantity, must be single valued. In magnetic-confinement systems, α is usually identified with the flux-surface label q , while β is a magnetic field-line label. $\beta = \text{constant}$ surfaces are shown in Figs. 4.12, 17 for toroidal devices. In mirrors, they are quite similar. The problems encountered with the multiple valuedness of β in tori have been discussed above. In mirrors, β represents an angle-like variable and its use is straightforward.

The functions α and β are not unique for any given magnetic field. For example, we can add to α any arbitrary function of β , or the other way around (but not both at the same time):

$$\nabla \alpha \times \nabla(\beta + g(\alpha)) = (\nabla \alpha \times \nabla \beta) + \left(\nabla \alpha \times \frac{\partial g}{\partial \alpha} \nabla \alpha \right) = \nabla \alpha \times \nabla \beta . \quad (5.2.13)$$

5.3 Relationship to the Contra- and Covariant Formalism

When we work with Clebsch coordinates α and β and a variable along a field line, we imply that space has been parameterized by these three variables. In many cases, this last variable is the length l of the field line (from some particular origin). The coordinate surfaces for the coordinates (α, β, l) are thus $\alpha(\mathbf{R}) = \text{constant}$, $\beta(\mathbf{R}) = \text{constant}$ and $l(\mathbf{R}) = \text{constant}$. The coordinate lines are intersections

of coordinate surfaces as usual. The special feature of the Clebsch representation is that the particular vector we want to describe, namely, \mathbf{B} , is tangent to one of the coordinate curves: the l curve (since it is by definition the intersection of $\alpha = \text{constant}$ and $\beta = \text{constant}$ surfaces).

We shall assume that \hat{l} is in the direction of \mathbf{B} , although, in fact, $\mathbf{B} = B\hat{\mathbf{B}} = \pm B\hat{l}$.

The covariant-basis vectors in the (α, β, l) coordinate system are denoted by e_α , e_β , and e_l ; the contravariant-basis vectors by e^α , e^β , e^l . The latter are defined by the gradients of the coordinate surfaces

$$e^\alpha \equiv \nabla\alpha, \quad e^\beta \equiv \nabla\beta, \quad e^l \equiv \nabla l. \quad (5.3.1)$$

The covariant-basis vectors are tangent to the coordinate curves. Thus, e_l is tangent to the field lines. Moreover, since l measures the length along that coordinate curve, its magnitude is unity:

$$e_l = \frac{\mathbf{B}}{B} \equiv \hat{\mathbf{B}} \equiv \hat{l}. \quad (5.3.2)$$

(Often, our $\hat{\mathbf{B}}$ is written as $\hat{\mathbf{b}}$; however, we avoid $\hat{\mathbf{b}}$ here because we used it to denote the binormal of the Frenet-Serret trihedron.) We shall use the symbols of (5.3.2) interchangeably. The other covariant-basis vectors do not have such a simple form but can be written in terms of the contravariant-basis vectors by means of the Jacobian J :

$$J = \frac{1}{\nabla\alpha \cdot \nabla\beta \times \nabla l} = \frac{1}{\nabla l \cdot \nabla\alpha \times \nabla\beta}. \quad (5.3.3)$$

We get

$$e_\alpha = Je^\beta \times e^l = J\nabla\beta \times \nabla l, \quad (5.3.4a)$$

$$e_\beta = Je^l \times e^\alpha = J\nabla l \times \nabla\alpha. \quad (5.3.4b)$$

Nothing prohibits us from writing the unit vector along \mathbf{B} in a similar form:

$$e_l \equiv \hat{\mathbf{B}} \equiv \hat{l} = J\nabla\alpha \times \nabla\beta. \quad (5.3.4c)$$

The magnetic field can be written in terms of covariant- or contravariant-basis vectors:

$$\mathbf{B} = B^\alpha e_\alpha + B^\beta e_\beta + B^l e_l, \quad \text{with} \quad B^i = \mathbf{B} \cdot e^i, \quad (5.3.5a)$$

or

$$\mathbf{B} = B_\alpha e^\alpha + B_\beta e^\beta + B_l e^l, \quad \text{with} \quad B_i = \mathbf{B} \cdot e_i. \quad (5.3.5b)$$

Equation (5.3.5a), the contravariant form of \mathbf{B} (since it is written with contravariant components), simplifies to a plausible result. By construction $B^\alpha \equiv$

$\mathbf{B} \cdot \nabla\alpha = 0$, and $B^\beta \equiv \mathbf{B} \cdot \nabla\beta = 0$, such that the equation in question becomes with the help of (5.3.2):

$$\mathbf{B} = B^l e_l = B^l \hat{l} = B^l \hat{\mathbf{B}} = \frac{B^l}{B} \mathbf{B}. \quad (5.3.6)$$

From this we conclude that $B^l \equiv B$ and thus that

$$\mathbf{B} = B\hat{\mathbf{B}} \quad (5.3.7)$$

is the *contravariant representation* of \mathbf{B} (in α, β, l coordinates).

Consider the expression for $\hat{\mathbf{B}}$ given in (5.3.4c), but with the Jacobian written explicitly

$$\hat{\mathbf{B}} = \frac{\nabla\alpha \times \nabla\beta}{\nabla l \cdot \nabla\alpha \times \nabla\beta}. \quad (5.3.8)$$

If we take the dot product of either side with ∇l , we obtain:

$$\hat{\mathbf{B}} \cdot \nabla l = 1. \quad (5.3.9)$$

Equation (5.3.9) also follows trivially from the observation that $\hat{\mathbf{B}} \cdot \nabla \equiv \partial/\partial l$. It should be noted that (5.3.9) is (in general) the best we can do for a relationship between $\hat{\mathbf{B}}$ and ∇l . In general, and unlike e_l , $e^l \equiv \nabla l$ is *not* equal to $\hat{\mathbf{B}}$; they point in different directions (contrary to what has been assumed frequently in the literature, as e.g., by Lee et al. 1986).

With this expression and the Clebsch representation, $\mathbf{B} = \nabla\alpha \times \nabla\beta$, we find the value for the Jacobian of (5.3.3)

$$J = \frac{1}{\nabla l \cdot \nabla\alpha \times \nabla\beta} = \frac{1}{\nabla l \cdot \mathbf{B}} = \frac{1}{B \nabla l \cdot \hat{\mathbf{B}}} = \frac{1}{B}. \quad (5.3.10)$$

This result was also obvious from (5.3.4c), and allows us to identify the contravariant form of (5.3.7) with the Clebsch form:

$$\mathbf{B} = B\hat{\mathbf{B}} = BJ\nabla\alpha \times \nabla\beta = \nabla\alpha \times \nabla\beta. \quad (5.3.11)$$

Thus, the “complete” *contravariant representation* of \mathbf{B} is

$$\mathbf{B} = B\hat{\mathbf{B}} = \nabla\alpha \times \nabla\beta. \quad (5.3.12)$$

The covariant representation does not provide a simple form because ∇l is in general not parallel to \mathbf{B} . From (5.3.4) and (5.3.10), we can rewrite the *covariant form* of (5.3.5b) as follows:

$$\begin{aligned} \mathbf{B} &= \mathbf{B} \cdot \frac{\nabla\beta \times \nabla l}{B} \nabla\alpha + \mathbf{B} \cdot \frac{\nabla l \times \nabla\alpha}{B} \nabla\beta + \mathbf{B} \cdot \frac{\nabla\alpha \times \nabla\beta}{B} \nabla l \\ &= (\hat{\mathbf{B}} \cdot \nabla\beta \times \nabla l) \nabla\alpha + (\hat{\mathbf{B}} \cdot \nabla l \times \nabla\alpha) \nabla\beta + B \nabla l. \end{aligned} \quad (5.3.13)$$

In the special case where ∇l happens to be equal to \vec{B} , this reduces simply to $B = B\vec{B}$. In that case, the co- and contravariant representations coincide.

We shall now determine the proportionality constant in (5.1.1). Consider a surface $l = \text{constant}$. An infinitesimal-area element on this surface is equal to

$$dS(l) = J \nabla l d\alpha d\beta = \frac{\nabla l}{B} d\alpha d\beta , \quad (5.3.14)$$

where we have set the Jacobian $J = 1/B$. The magnetic flux through an element $dS(l)$ equals

$$d\Psi = \vec{B} \cdot dS(l) = \frac{\vec{B} \cdot \nabla l}{B} d\alpha d\beta = d\alpha d\beta . \quad (5.3.15a)$$

This is shown in Fig. 5.2. Every surface S that is intersected by \vec{B} has a similar (α, β) grid. B -lines can never cross α or β surfaces, so they cannot leave a certain volume tube bounded by the surfaces $\alpha = \text{constant}$, $\alpha + d\alpha = \text{constant}$, $\beta = \text{constant}$ and $\beta + d\beta = \text{constant}$. Hence, our conclusion is that the magnetic flux in that tube must remain the same. Thus, the flux through any surface S intersecting $\vec{B} = \nabla\alpha \times \nabla\beta$ equals

$$\Psi = \iint d\alpha d\beta . \quad (5.3.15b)$$

In the case where we dealt with the stream functions $\tilde{\alpha}$ and $\tilde{\beta}$, \vec{B} was written as $C \nabla\tilde{\alpha} \times \nabla\tilde{\beta}$. If we parameterize a field line by $\tilde{\alpha}$, $\tilde{\beta}$, and l , the Jacobian in these coordinates is

$$J = \frac{1}{\nabla l \cdot (\nabla\tilde{\alpha} \times \nabla\tilde{\beta})} = \frac{C}{\nabla l \cdot \vec{B}} = \frac{C}{B} . \quad (5.3.16)$$

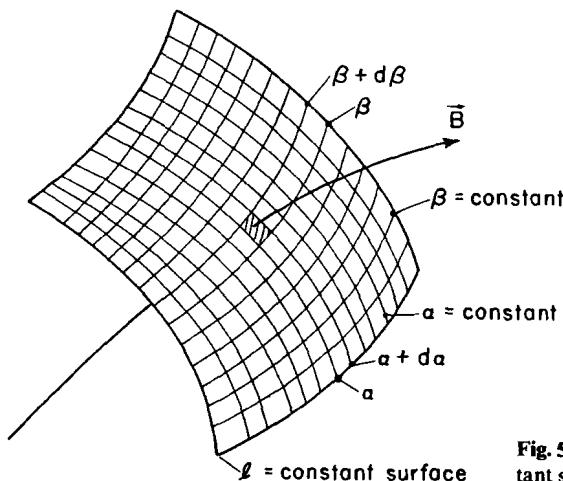


Fig. 5.2. (α, β) grid laid out on an $l = \text{constant}$ surface

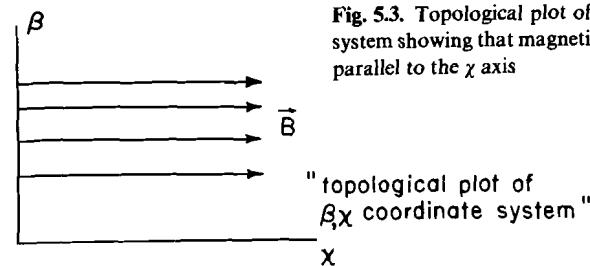


Fig. 5.3. Topological plot of the (β, χ) Boozer-Grad coordinate system showing that magnetic-field lines appear as straight lines, parallel to the χ axis

Consequently, the combination of $d\Psi = \vec{B} \cdot dS(l)$, with $dS(l)$ now equal to $C \nabla l d\tilde{\alpha} d\tilde{\beta}/B$ leads to

$$d\Psi = \frac{C}{B} \vec{B} \cdot \nabla l d\tilde{\alpha} d\tilde{\beta} = C d\tilde{\alpha} d\tilde{\beta} \quad (5.3.17a)$$

or

$$\Psi = \iint C d\tilde{\alpha} d\tilde{\beta} . \quad (5.3.17b)$$

This explains what the proportionality sign in (5.1.1) means. Thus, α , β , $\tilde{\alpha}$ and $\tilde{\beta}$ are all stream functions.

One of the properties of the Clebsch coordinate system is that the field lines are straight lines in the coordinates other than α (the one labeling the flux surface), i.e., the field-line label β and the third coordinate which is used to indicate the "distance" along a field line (which may be but does not have to be the length l). We call this third variable χ . The differential equation for a field line in (β, χ) coordinates is

$$\frac{d\beta}{d\chi} = \frac{B^\beta}{B^\chi} \quad (5.3.18)$$

where $B^\beta = \vec{B} \cdot \nabla\beta$, $B^\chi = \vec{B} \cdot \nabla\chi$. With $\vec{B} = \nabla\alpha \times \nabla\beta$, this equation becomes

$$\frac{d\beta}{d\chi} = 0 \quad \text{or} \quad \beta = \text{constant.} \quad (5.3.19)$$

This means that the field line represents a straight line parallel to the χ axis (see Fig. 5.3 – note that this is a "topological" plot, since, in reality, the β and χ axes do not intersect at right angles).

5.4 Boozer-Grad Coordinates

Boozer-Grad coordinates are characterized by a particular choice of the third coordinate in the Clebsch representation. Recently, this coordinate system has been employed frequently by Boozer (1980, 1981) (after discovering it inde-

pendently) but it was previously used by Grad (1971) (see also, Catto and Hazeltine 1981; White et al. 1983; and Betancourt and Garabedian 1985). We shall discuss the system for toroidal devices that possess flux surfaces, each of which is traced out by one single field line (at least for the irrational flux surfaces). A discussion for mirrors would be similar, except for a few minor points. (One of which is that $\partial\Phi/\partial l = 0$ does not imply that Φ is a flux-surface function in mirrors, but only that Φ is independent of l .)

In this coordinate system, we require that the contravariant representation for \mathbf{B} is the usual Clebsch form

$$\mathbf{B} = \nabla\varrho \times \nabla v . \quad (5.4.1)$$

We have made a change of notation: $\alpha \equiv \varrho$ and $\beta \equiv v$. Although α and β are normally used as the flux-surface label and field-line label, respectively, several authors have used other notations. Boozer (1980) uses ψ as a flux-surface label and α as a field-line label. He uses β for the Clebsch coordinate of the current – see below. Betancourt and Garabedian (1985) use s as a flux label and ψ as a field-line label. We opted for a different set of variables ϱ and v , since ϱ normally represents a radial variable, and we used it before as a flux-surface label. v has been used with the same meaning as here by Solov'ev and Shafranov (1970).

In general, $\nabla\varrho \cdot \nabla v \neq 0$ and $e_\varrho \cdot e_v \neq 0$. The former implies that the ∇v direction is not tangent to a flux surface (as is often erroneously claimed), although it is perpendicular to \mathbf{B} . The covariant-basis vector e_v is by definition tangent to a flux surface but is not necessarily orthogonal to \mathbf{B} (see, e.g., Figs. 4.2 and 5.4).

The idea of the Boozer-Grad coordinate system is to find a covariant representation which reduces to the gradient of a scalar function (the magnetic scalar potential) in a current free region. Rather than taking the arc length l along \mathbf{B} as the third coordinate, a function $\tilde{\chi}$ is used such that the third term in the covariant expression, similar to (5.3.13), reads $\nabla\tilde{\chi}$ instead of $B\nabla l$. A second reason for working with $\tilde{\chi}$ rather than with l is that there is a simple relationship between $\tilde{\chi}$ and the angle variables θ and ζ , whereas, no such simple relationship exists between l and (θ, ζ) . This issue will be discussed in detail later.

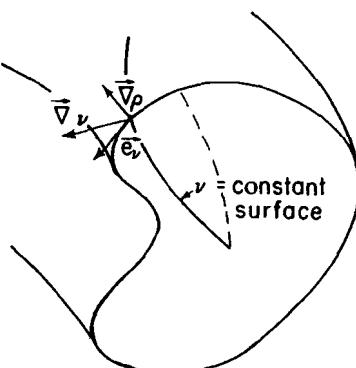


Fig. 5.4. $\varrho = \text{constant}$ and $v = \text{constant}$ coordinate surfaces. The contravariant basis vector ∇v is perpendicular to the $v = \text{constant}$ surface, whereas the covariant basis vector e_v is parallel to the $\varrho = \text{constant}$ surface. ∇v is in general not tangent to the flux surface $\varrho = \text{constant}$

As in the previous section, we assume that $\hat{\mathbf{l}} = +\hat{\mathbf{B}}$ for simplicity of notation.

Writing $\nabla\tilde{\chi}$ as the third term in the covariant form does *not* imply that $B\nabla l$ equals $\nabla\tilde{\chi}$. The main reason is that our new choice implies a new Jacobian, as well as new covariant-basis vectors. The covariant components B_ϱ and B_v of the (ϱ, v, l) system of (5.3.13), have now changed into B'_ϱ and B'_v , respectively, and consequently, so has the vector $(B_\varrho \nabla\varrho + B_v \nabla v)$, which now becomes $(B'_\varrho \nabla\varrho + B'_v \nabla v) -$ which has in general, a different direction. Since both

$$B_\varrho \nabla\varrho + B_v \nabla v + B \nabla l \quad \text{and} \quad B'_\varrho \nabla\varrho + B'_v \nabla v + \nabla\tilde{\chi}$$

must add up to the same vector \mathbf{B} , it is clear that in general $\nabla\tilde{\chi}$ cannot be parallel to ∇l . (A rigorous proof of this will be given presently.)

For the covariant representation, we make the Ansatz in the third term

$$\mathbf{B} = [\mathbf{B} \cdot (\nabla v \times \nabla\tilde{\chi}) J] \nabla\varrho + [\mathbf{B} \cdot (\nabla\tilde{\chi} \times \nabla\varrho) J] \nabla v + \nabla\tilde{\chi} \quad (5.4.2a)$$

or

$$\mathbf{B} = \tilde{\lambda} \nabla\varrho + \tilde{\xi} \nabla v + \nabla\tilde{\chi} . \quad (5.4.2b)$$

We have used the abbreviations

$$\tilde{\lambda} \equiv B'_\varrho \equiv \mathbf{B} \cdot (\nabla v \times \nabla\tilde{\chi}) J$$

and

$$\tilde{\xi} \equiv B'_v \equiv \mathbf{B} \cdot (\nabla\tilde{\chi} \times \nabla\varrho) J .$$

The new Jacobian is easily found by taking the dot product of the contravariant and covariant forms of \mathbf{B} , (5.4.1) and (5.4.2):

$$J = (\nabla\tilde{\chi} \cdot \nabla\varrho \times \nabla v)^{-1} = \frac{1}{(B)^2} . \quad (5.4.3)$$

This result is also clear from our requirement, in (5.4.2a), that the third term is $\nabla\tilde{\chi} \equiv [\mathbf{B} \cdot (\nabla\varrho \times \nabla v) J] \nabla\tilde{\chi}$.

This Jacobian allows us to obtain some more insight concerning the function $\tilde{\chi}$. If we take $\hat{\mathbf{B}} \cdot \nabla\tilde{\chi}$, (5.4.3) gives

$$\hat{\mathbf{B}} \cdot \nabla\tilde{\chi} = \frac{\mathbf{B} \cdot \nabla\tilde{\chi}}{B} = \frac{\nabla\varrho \times \nabla v}{B} \cdot \nabla\tilde{\chi} = \frac{(B)^2}{B} \quad (5.4.4)$$

or thus

$$\hat{\mathbf{B}} \cdot \nabla\tilde{\chi} = \left. \frac{\partial \tilde{\chi}}{\partial l} \right|_{\varrho, v} = B . \quad (5.4.5)$$

This is an important result. The differential distance along a field line equals $d\tilde{\chi}/B$:

$$dl|_{\text{along } \mathbf{B}} = \left. \frac{d\tilde{\chi}}{B} \right|_{\text{along } \mathbf{B}} . \quad (5.4.6)$$

The tangent $\tilde{\chi}$ -basis vector $e_{\tilde{\chi}}$ is by definition equal to

$$e_{\tilde{\chi}} = \frac{\nabla \varrho \times \nabla v}{\nabla \tilde{\chi} \cdot (\nabla \varrho \times \nabla v)} = \frac{B \hat{B}}{(B)^2} = \frac{\hat{B}}{B} \quad \text{or} \quad e_{\tilde{\chi}} = \frac{\hat{l}}{B} . \quad (5.4.7)$$

From (5.4.7), it also follows that

$$e_{\tilde{\chi}} \cdot \nabla l = \frac{\hat{l}}{B} \cdot \nabla l = \frac{1}{B} . \quad (5.4.8)$$

To show that $B \nabla l \neq \nabla \tilde{\chi}$ for general magnetic-field structures, we apply the chain rule to $\nabla \tilde{\chi}$ in ϱ, v, l coordinates

$$\nabla \tilde{\chi} = \frac{\partial \tilde{\chi}}{\partial \varrho} \nabla \varrho + \frac{\partial \tilde{\chi}}{\partial v} \nabla v + \frac{\partial \tilde{\chi}}{\partial l} \nabla l . \quad (5.4.9)$$

Because of (5.4.5), this gives:

$$\nabla \tilde{\chi} = B \nabla l + \left. \frac{\partial \tilde{\chi}}{\partial \varrho} \right|_{v,l} \nabla \varrho + \left. \frac{\partial \tilde{\chi}}{\partial v} \right|_{\varrho,l} \nabla v . \quad (5.4.10)$$

If $B \nabla l$ were equal to $\nabla \tilde{\chi}$, then we would conclude that either $\nabla \varrho = -[(\partial \tilde{\chi}/\partial v)/(\partial \tilde{\chi}/\partial \varrho)] \nabla v$, or that $\partial \tilde{\chi}/\partial \varrho$ and $\partial \tilde{\chi}/\partial v$ are zero. The former is impossible by the definition of the functions ϱ and v . The latter is not true by the argument which will now be developed. The fact that $\partial \tilde{\chi}/\partial l = B$, (5.4.5), implies that

$$\tilde{\chi}(\varrho, v, l) - \tilde{\chi}_0(\varrho, v) = \int_0^l B(\varrho, v, l) dl . \quad (5.4.11)$$

$\tilde{\chi}_0$ is the value of $\tilde{\chi}$ at $l = 0$. In general, $B = |\mathbf{B}|$ depends on all three coordinates, (ϱ, v, l) . Then the following statement must exhibit a contradiction:

$$\left. \frac{\partial \tilde{\chi}}{\partial \varrho} \right|_{v,l} \stackrel{?}{=} 0 \stackrel{?}{=} \frac{\partial \tilde{\chi}_0}{\partial \varrho} + \int_0^l \frac{\partial B}{\partial \varrho}(\varrho, v, l) dl \quad (5.4.12)$$

Indeed, $\partial \tilde{\chi}_0/\partial \varrho$ is independent of l , while the integral is dependent on l unless the partial derivative $\partial B/\partial \varrho$ vanishes for all l , which is not the case for general B . An analogous argument can be developed for $\partial \tilde{\chi}/\partial v$.

Next, we want to redefine $\tilde{\chi}$, $\tilde{\lambda}$, and $\tilde{\xi}$ such that (5.4.2b) simplifies to a form like

$$\mathbf{B} = \lambda \nabla \varrho + \nabla \chi . \quad (5.4.13)$$

In other words, we want the term involving $\tilde{\xi}$ to vanish. However, this forces us to change $\tilde{\chi}$ to χ , and $\tilde{\lambda}$ to λ as well. In order to do so, we make use of the ideal MHD force balance

$$\nabla p = \mathbf{J} \times \mathbf{B} . \quad (5.4.14)$$

(\mathbf{J} here is the current and not the Jacobian.) From $\mathbf{B} \cdot \nabla p = 0$, we conclude that since, in a toroidal system, \mathbf{B} traces out flux surfaces $\varrho = \text{constant}$, p is a function of ϱ only, $p \equiv p(\varrho)$. Furthermore, the condition $\mathbf{J} \cdot \nabla p = 0$ is equivalent to $\mathbf{J} \cdot \nabla \varrho dp/d\varrho = 0$ or

$$\mathbf{J} \cdot \nabla \varrho = 0 \quad (5.4.15)$$

(except possibly at points where $dp/d\varrho = 0$). This says that the current-density lines lie in constant ϱ surfaces.

Ampère's law permits us to find an expression for \mathbf{J} . From $\mathbf{J} = (1/\mu_0) \nabla \times \mathbf{B}$, we obtain with the help of (5.4.2b)

$$\mathbf{J} = \frac{1}{\mu_0} \nabla \times [\tilde{\lambda} \nabla \varrho + \tilde{\xi} \nabla v + \nabla \tilde{\chi}] = \frac{1}{\mu_0} [(\nabla \tilde{\lambda} \times \nabla \varrho) + (\nabla \tilde{\xi} \times \nabla v)] . \quad (5.4.16)$$

The condition $\mathbf{J} \cdot \nabla \varrho = 0$ implies that

$$\nabla \varrho \cdot (\nabla \tilde{\xi} \times \nabla v) = 0$$

$$(\nabla \varrho \times \nabla v) \cdot \nabla \tilde{\xi} = 0$$

so

$$\mathbf{B} \cdot \nabla \tilde{\xi} = 0 \quad (5.4.17)$$

or

$$\frac{\partial \tilde{\xi}}{\partial l} = 0 . \quad (5.4.18)$$

In a toroidal device with flux surfaces traced out by single field lines, this states that $\tilde{\xi}$ is a flux-surface quantity

$$\tilde{\xi} \equiv \tilde{\xi}(\varrho) . \quad (5.4.19)$$

If we look back at our original covariant form for \mathbf{B} , $\mathbf{B} = \tilde{\lambda} \nabla \varrho + \tilde{\xi} \nabla v + \nabla \tilde{\chi}$, we observe that the term $\tilde{\xi}(\varrho) \nabla v$ pushes the field line in the ∇v direction (which would potentially cause a secular behavior) (see Fig. 5.4). A field line does not move in the ∇v direction, however. Because $\mathbf{B} \cdot \nabla v = 0$, field lines lie in constant v surfaces. From (5.4.10) and the discussion below (5.4.12), we see that $\nabla \tilde{\chi}$ does have components in the $\nabla \varrho$ and ∇v directions. The terms $\nabla \tilde{\chi}$ and $\tilde{\lambda} \nabla \varrho$ will cancel the "secular" $\tilde{\xi} \nabla v$ term. By looking for a new χ , we shall try to absorb the component of $\nabla \tilde{\chi}$ which lies in the ∇v direction into the new $\nabla \chi$. (Equation (5.4.2b) does not imply that \mathbf{B} has a component along $\nabla \varrho$ either. Since $\mathbf{B} \cdot \nabla \varrho = 0$, the two other terms $\tilde{\xi} \nabla v$ and $\nabla \tilde{\chi}$ must have components in the $\nabla \varrho$ direction to make the net component of \mathbf{B} along $\nabla \varrho$ vanish.)

Because $\tilde{\xi}$ is a flux function, its gradient is along the $\nabla \varrho$ direction: $\nabla \tilde{\xi} = \dot{\tilde{\xi}} \nabla \varrho$, where the dot represents the derivative with respect to ϱ , $\dot{\tilde{\xi}} \equiv d\tilde{\xi}/d\varrho$. The expression for the current density in (5.4.16) can be written successively as follows:

$$\begin{aligned}
\mu_0 \mathbf{J} &= (\nabla \tilde{\lambda} \times \nabla \varrho) + (\dot{\tilde{\lambda}} \nabla \varrho \times \nabla v) \\
&= \nabla \varrho \times [\dot{\tilde{\lambda}} \nabla v - \nabla \tilde{\lambda}] \\
&= \nabla \varrho \times [\nabla(\dot{\tilde{\lambda}} v) - \nabla \tilde{\lambda}] \\
&= \nabla \varrho \times \nabla(\dot{\tilde{\lambda}} v - \tilde{\lambda}) .
\end{aligned} \tag{5.4.20}$$

Introducing a new function λ , such that the current has a simple Clebsch representation, gives

$$\mu_0 \mathbf{J} = \nabla \lambda \times \nabla \varrho \tag{5.4.21}$$

with

$$\lambda \equiv (\tilde{\lambda} - \dot{\tilde{\lambda}} v) . \tag{5.4.22}$$

Thus, λ is a Clebsch (stream) function for the current density \mathbf{J} .

We now substitute the $\tilde{\lambda}$ from (5.4.22) in the covariant \mathbf{B} for (5.4.2b), $\mathbf{B} = \tilde{\lambda} \nabla \varrho + \dot{\tilde{\lambda}} \nabla v + \nabla \tilde{\lambda}$, and obtain

$$\begin{aligned}
\mathbf{B} &= (\lambda + \dot{\tilde{\lambda}} v) \nabla \varrho + \dot{\tilde{\lambda}} \nabla v + \nabla \tilde{\lambda} \\
&= \lambda \nabla \varrho + (\dot{\tilde{\lambda}} v \nabla \varrho + \dot{\tilde{\lambda}} \nabla v) + \nabla \tilde{\lambda} \\
&= \lambda \nabla \varrho + \nabla(\dot{\tilde{\lambda}} v) + \nabla \tilde{\lambda} \\
&= \lambda \nabla \varrho + \nabla(\dot{\tilde{\lambda}} v + \tilde{\lambda}) .
\end{aligned} \tag{5.4.23}$$

When we define the new function χ as

$$\chi \equiv \tilde{\lambda} + \dot{\tilde{\lambda}} v , \tag{5.4.24}$$

the magnetic field becomes in *covariant form*:

$$\mathbf{B} = \lambda \nabla \varrho + \nabla \chi . \tag{5.4.25}$$

Since, in general, $\nabla v \cdot \nabla \varrho \neq 0 \neq \nabla \chi \cdot \nabla \varrho$, both $\nabla \varrho$ and $\nabla \chi$ have components along ∇v . However, they offset each other so that no secular term remains.

In summary, to arrive at the simple covariant form given in (5.4.25), we have made two substitutions:

$$\begin{aligned}
\chi &= \tilde{\lambda} + \dot{\tilde{\lambda}} v \\
\lambda &= \tilde{\lambda} - \dot{\tilde{\lambda}} v .
\end{aligned}$$

Neither $\tilde{\lambda}$ nor v varies with l . Consequently,

$$\hat{\mathbf{B}} \cdot \nabla \chi = \frac{\partial \chi}{\partial l} = \frac{\partial \tilde{\lambda}}{\partial l} = B .$$

Because of this result, the relationships expressed in (5.4.3), (5.4.5), and (5.4.7) are still valid for χ as well as $\tilde{\lambda}$:

$$\hat{\mathbf{B}} \cdot \nabla \chi = B \tag{5.4.26}$$

$$J^{-1} = \nabla \chi \cdot (\nabla \varrho \times \nabla v) = (B)^2 \tag{5.4.27}$$

$$e_\chi = \frac{\hat{l}}{B} = \frac{\hat{\mathbf{B}}}{B} . \tag{5.4.28}$$

Having established the identity of the new function χ , we can rewrite the “complete” contravariant form of \mathbf{B} . Multiplying the Clebsch representation in (5.4.1) by J in the numerator and the denominator, and using (5.4.27), we obtain

$$\mathbf{B} = (B)^2 e_\chi = \nabla \varrho \times \nabla v . \tag{5.4.29}$$

These are the two equivalent contravariant representations of \mathbf{B} .

It is not easy to visualize the $\chi = \text{constant}$ surfaces. In Chap. 6, we shall show that the $v = \text{constant}$ surfaces are ruled surfaces in the toroidal coordinates (θ_f, ζ_f) in which the magnetic field is “straight”, with an equation of the form $v = A(\varrho)\theta_f - D(\varrho)\zeta_f$, (6.6.14b) (see also the discussion related to Fig. 4.17, below (4.9.15)). The v surfaces are traced out by “straight” field lines. Similarly, we shall show there that the $\chi = \text{constant}$ surfaces are ruled surfaces traced out by a “straight” line, obeying an equation of the form $\chi = a(\varrho)\theta_f + d(\varrho)\zeta_f$, (6.6.14c) (which intersects a field line).

The (ϱ, v, χ) Boozer-Grad coordinate system suffers from the same shortcomings when it comes to single valuedness in toroidal devices, as those encountered with the (α, β, l) coordinates. Again, this system is useful for formal derivations, but practical calculations are better done after conversion to an angle system. This transformation will be discussed in an upcoming section.

Finally, we show that the λ Clebsch coordinate of the current density can be related to the pressure gradient and the Pfirsch-Schlüter current. Following Boozer (1980), we obtain from (5.4.21) with $\nabla \lambda = (\partial \lambda / \partial \varrho) \nabla \varrho + (\partial \lambda / \partial v) \nabla v + (\partial \lambda / \partial \chi) \nabla \chi$

$$\mu_0 \mathbf{J} = \frac{\partial \lambda}{\partial v} \nabla v \times \nabla \varrho + \frac{\partial \lambda}{\partial \chi} \nabla \chi \times \nabla \varrho \tag{5.4.30}$$

or

$$\mu_0 \mathbf{J} = -\mathbf{B} \frac{\partial \lambda}{\partial v} + \nabla \chi \times \nabla \varrho \frac{\partial \lambda}{\partial \chi} . \tag{5.4.31}$$

Since $\hat{\mathbf{B}} \cdot \nabla \varrho \times \nabla \chi = \hat{\mathbf{B}} \cdot \nabla \varrho \times (\mathbf{B} - \lambda \nabla \varrho) = 0$, the parallel component of (5.4.30) leads to

$$\frac{\partial \lambda}{\partial v} = -\mu_0 \frac{J_l}{B} . \tag{5.4.32}$$

To find an interesting relationship for λ from the perpendicular component, we substitute \mathbf{J} in $\nabla p = \mathbf{J} \times \mathbf{B}$. Recalling that $p \equiv p(\varrho)$, we get from (5.4.30)

$$\dot{p} \nabla \varrho = \frac{1}{\mu_0} \left(\nabla \chi \times \nabla \varrho \frac{\partial \lambda}{\partial \chi} \right) \times \mathbf{B} \quad (5.4.33)$$

which, with the *bac-cab* rule, becomes

$$\begin{aligned} \dot{p} \nabla \varrho &= \frac{1}{\mu_0} \frac{\partial \lambda}{\partial \chi} [\nabla \varrho (\mathbf{B} \cdot \nabla \chi) - \nabla \chi (\mathbf{B} \cdot \nabla \varrho)] \\ &= \frac{1}{\mu_0} \frac{\partial \lambda}{\partial \chi} [\nabla \varrho (\nabla \varrho \times \nabla v \cdot \nabla \chi)] \\ &= \frac{(B)^2}{\mu_0} \frac{\partial \lambda}{\partial \chi} \nabla \varrho . \end{aligned}$$

Equating the coefficients of both sides, we obtain an expression for $\partial \lambda / \partial \chi$:

$$\frac{\partial \lambda}{\partial \chi} = \frac{\mu_0}{(B)^2} \frac{dp}{d\varrho} . \quad (5.4.34)$$

The function λ is determined via the differential equations, (5.4.32) and (5.4.34), and is thus determined up to an additive constant (constant in v and χ). To take away some of the arbitrariness of λ (and χ), we impose the initial condition at $\chi = v = 0$,

$$\lambda(\varrho, v = 0, \chi = 0) = 0 . \quad (5.4.35)$$

This condition implies that $\lambda = 0$ when $\mathbf{J} = 0$, as we shall show momentarily. Indeed, $\mathbf{J} = 0$ makes J_{\parallel} and \dot{p} vanish so that we obtain from (5.4.32) and (5.4.34),

$$\lambda \equiv \lambda(\varrho) \quad \text{when} \quad \mathbf{J} = 0 . \quad (5.4.36)$$

This can also be seen from (5.4.21). The “initial” condition (5.4.36) then sets

$$\lambda = 0 \quad \text{when} \quad \mathbf{J} = 0 . \quad (5.4.37)$$

In vacuum, the covariant form of \mathbf{B} becomes

$$\mathbf{B} = \nabla \chi, \quad \mathbf{J} = 0 . \quad (5.4.38)$$

When $\mathbf{J} = 0$, the function χ is the magnetic scalar potential.

However, this still does not remove all the arbitrariness of the (ϱ, v, χ) coordinate system. As with any Clebsch coordinate, v is determined only to within an additive flux function (cf. (5.2.13)). If $v^* = v + g(\varrho)$, \mathbf{B} is unchanged: $\mathbf{B} = \nabla \varrho \times \nabla v = \nabla \varrho \times \nabla v^*$. Since v does not appear in the covariant form, this change in v does not force any other quantities to change. The magnetic field and the properties of χ are also left the same if we construct a new χ by adding a flux

function. That the properties of a new $\chi^* = \chi + f(\varrho)$ do not change has already been made clear when we substituted χ for $\tilde{\chi}$, using (5.4.24). We saw that χ retains its properties as long as the additive term is independent of l . This is certainly true for a flux function. However, changing χ to χ^* forces λ to change to a new λ^* as follows:

$$\chi^* = \chi + f(\varrho) \Rightarrow \lambda^* = \lambda - \frac{df}{d\varrho} . \quad (5.4.39)$$

6. Toroidal Flux Coordinates

6.1 Straight Field-Line Coordinates

Toroidal flux coordinates are a set of poloidal and toroidal “angles” θ_f and ζ_f chosen such that the equation of a field line is that of a straight line in those coordinates. It is common to say that the magnetic-field lines appear as straight lines, or that the magnetic field is straight in (θ_f, ζ_f) .

To construct functions $\theta_f(\mathbf{R})$ and $\zeta_f(\mathbf{R})$ it is necessary that the *magnetic field* \mathbf{B} is known completely. In this section, we assume the vector \mathbf{B} to be known everywhere as a function of arbitrary angle coordinates θ and ζ . (These angles could, for example, be the “elementary” angles θ_e and ζ_e .) In a later section, we shall see how to transform from Cartesian components of \mathbf{B} to the desired flux coordinates (θ_f, ζ_f) .

The first coordinate of the system will be the flux-surface label ϱ . Surfaces of constant ϱ are the flux surfaces. ϱ stands for any flux-surface function and can be “replaced” anywhere by a quantity such as the poloidal or toroidal flux enclosed by (or outside of) the flux surface Ψ_{tor} , Ψ_{pol} , the volume of a flux surface V , the pressure on that flux surface p , the rotational transform or its inverse, t , q, \dots (as long as $dp/d\varrho \neq 0$, $dt/d\varrho \neq 0$, etc. at places other than the magnetic axis). Recall that we only accept radially increasing ϱ ’s. Thus, $\varrho \equiv -p$; $\varrho \equiv -\Psi_{\text{pol}}^d$; $\varrho \equiv t$ and $\varrho \equiv -q$ in most stellarators; $\varrho \equiv q$ and $\varrho \equiv -t$ in tokamaks.

The angles θ and ζ on the surface are defined by *multiple-valued functions* $\theta(\mathbf{R})$ and $\zeta(\mathbf{R})$, in the sense that a point on the flux surface has (θ, ζ) coordinates equal to $\theta + 2\pi m$ and $\zeta + 2\pi n$, where m and n are arbitrary integers. Their gradients must be single valued, however (i.e., periodic functions of θ and ζ).

The magnetic field \mathbf{B} can be written in terms of its contravariant components in the (ϱ, θ, ζ) coordinate system as usual:

$$\mathbf{B} = B^\varrho \mathbf{e}_\varrho + B^\theta \mathbf{e}_\theta + B^\zeta \mathbf{e}_\zeta , \quad (6.1.1)$$

where $B^i = \mathbf{B} \cdot \nabla u^i$. From the definition of a flux surface,

$$\mathbf{B} \cdot \nabla \varrho = 0 , \quad (6.1.2)$$

we conclude immediately that

$$B^\varrho = 0 . \quad (6.1.3)$$

Thus \mathbf{B} equals:

$$\mathbf{B} = B^\theta \mathbf{e}_\theta + B^\zeta \mathbf{e}_\zeta = B^\theta \sqrt{g} (\nabla \zeta \times \nabla \varrho) + B^\zeta \sqrt{g} (\nabla \varrho \times \nabla \theta) . \quad (6.1.4)$$

Here we used $\mathbf{e}_i = (\nabla u^j \times \nabla u^k) \sqrt{g}$, where $\sqrt{g} = [\nabla \varrho \cdot (\nabla \theta \times \nabla \zeta)]^{-1}$ is the Jacobian. Recall that $g = (J)^2$, with g the determinant of the metric coefficient matrix. Instead of using J for the Jacobian, we shall henceforth use the symbol \sqrt{g} . (In the flux-coordinate context, the symbol J is often used for a current-density flux, which is a “current”, not a current-density. However, we shall not use the symbol J at all to avoid confusion with the Jacobian or the magnitude of the current-density.) Note that g , in general, is a function of (ϱ, θ, ζ) : $g \equiv g(\varrho, \theta, \zeta)$; the Jacobian is a known single valued (i.e., periodic) function in (θ, ζ) .

The two components B^θ and B^ζ are not independent of each other as can be seen from $\nabla \cdot \mathbf{B} = 0$,

$$\frac{\partial}{\partial \theta} (\sqrt{g} B^\theta) + \frac{\partial}{\partial \zeta} (\sqrt{g} B^\zeta) = 0 . \quad (6.1.5)$$

This relationship suggests that B^θ and B^ζ are derived from a single function $v \equiv v(\varrho, \theta, \zeta)$ as follows:

$$B^\theta = -\frac{1}{\sqrt{g}} \frac{\partial v}{\partial \zeta} , \quad B^\zeta = \frac{1}{\sqrt{g}} \frac{\partial v}{\partial \theta} . \quad (6.1.6a, b)$$

These B^θ and B^ζ satisfy (6.1.5), as can be checked by substitution. We can now rewrite (6.1.4):

$$\mathbf{B} = \nabla \varrho \times \left(\frac{\partial v}{\partial \zeta} \nabla \zeta + \frac{\partial v}{\partial \theta} \nabla \theta \right) . \quad (6.1.7)$$

Adding a term $(\partial v / \partial \varrho) \nabla \varrho$ between the brackets leaves \mathbf{B} unchanged, since $\nabla \varrho \times \nabla \varrho = 0$. Then we can use $\nabla v = (\partial v / \partial \varrho) \nabla \varrho + (\partial v / \partial \theta) \nabla \theta + (\partial v / \partial \zeta) \nabla \zeta$, allowing us to write the simple Clebsch form for \mathbf{B} :

$$\mathbf{B} = \nabla \varrho \times \nabla v . \quad (6.1.8)$$

This equation states that the \mathbf{B} lines lie on constant v surfaces. The same conclusion can be drawn from the equation of a field line on a $\varrho = \text{constant}$ surface. If we substitute (6.1.6) in $d\theta/B^\theta = d\zeta/B^\zeta$, we have that $dv = 0$, since $v \equiv v(\varrho, \theta, \zeta)$ and $d\varrho = 0$ on a flux surface. Hence, the equation of a field line lying in a flux surface $\varrho = \varrho_0$ is

$$v(\varrho_0, \theta, \zeta) = \text{constant} . \quad (6.1.9)$$

From the expression in (6.1.7), we can infer what form the function $v(\varrho, \theta, \zeta)$ must have. \mathbf{B} is a physical quantity and is thus single valued, i.e., its components must be periodic in (θ, ζ) . We need not require the gradient of v to be single valued (periodic), but $\nabla \varrho \times \nabla v$ must be. Because of the cross product with $\nabla \varrho$, the ϱ

dependence of v is immaterial as far as $\nabla\varrho \times \nabla v$ is concerned. If v is a periodic function in θ and ζ , then ∇v is periodic and so evidently is $\nabla\varrho \times \nabla v$. Hence with regard to the requirement that $\nabla\varrho \times \nabla v$ be single valued, a function $v(\varrho, \theta, \zeta)$, with an arbitrary ϱ dependence but which is periodic in (θ, ζ) is permitted. Any non-periodic function which is not a simple polynomial in θ or ζ must be rejected, since $\nabla\varrho \times \nabla v$ would then have non-periodic terms containing θ or ζ , implying multiple valuedness (the reader may wish to try functions like $\exp(\theta)$, $\ln(\zeta)$, $1/(1 + \zeta^2)$, ...). A function v having the form of a polynomial in, say θ ,

$$a(\varrho) + b(\varrho)\theta + c(\varrho)\theta^2 + \dots \quad (6.1.10)$$

is acceptable provided we keep only the first two terms

$$a(\varrho) + b(\varrho)\theta . \quad (6.1.11)$$

Otherwise, ∇v would contain terms like $(\partial(c\theta^2)/\partial\theta)\nabla\theta = 2c\theta\nabla\theta$, which is not acceptable because of the multiple valuedness in θ which remains after the cross product with $\nabla\varrho$. Note that with (6.1.11), we would have that $\nabla v = (\partial a/\partial\varrho)\nabla\varrho + (\partial b/\partial\varrho)\theta\nabla\varrho + b\nabla\theta$, which contains a secular term involving θ as well. However, this is permissible, since it contains a $\nabla\varrho$, and $\nabla\varrho \times (\)$ eliminates that term.

Similar arguments hold for polynomials in ζ . Our conclusion is that the most general form of $v(\varrho, \theta, \zeta)$, consistent with the periodicity of $\nabla\varrho \times \nabla v$ is

$$v(\varrho, \theta, \zeta) = A(\varrho) + B(\varrho)\theta + C(\varrho)\zeta + \tilde{v}(\varrho, \theta, \zeta) . \quad (6.1.12)$$

Here, \tilde{v} represents a periodic function in θ and ζ . $A(\varrho)$ can be absorbed into \tilde{v} so that we write:

$$v(\varrho, \theta, \zeta) = B(\varrho)\theta + C(\varrho)\zeta + \tilde{v}(\varrho, \theta, \zeta) . \quad (6.1.13)$$

One could have come to the same conclusion by requiring that the components B^θ and B^ζ in (6.1.6) are periodic. The form of (6.1.13) can rigorously be obtained if one expands both $\sqrt{g}B^\theta$ and $\sqrt{g}B^\zeta$ in Fourier series in θ and ζ . Then one could separate the terms $n = 0$ and $m = 0$; $n = 0$ and $m \neq 0$; $n \neq 0$ and $m = 0$; and m and n different from zero and manipulate these forms according to the prescriptions expressed by the differential equations, (6.1.6). To actually solve for v , one could integrate (6.1.6a) with respect to ζ such that the result contains an integration constant which is only a function of θ . Then that form of v is differentiated with respect to θ and is set equal to $\sqrt{g}B^\zeta$ according to (6.1.6b). This would provide an equation for the unknown function of θ , which could be solved for, after integration with respect to θ (up to an additive constant).

B^θ , B^ζ and \sqrt{g} are known functions of ϱ ; so the ϱ dependence of v is fixed. The ϱ -dependent functions $B(\varrho)$ and $C(\varrho)$ are found from (6.1.6) as follows. Recall the form-invariant expression for Ψ_{tor} (4.7.10)

$$\Psi_{\text{tor}} = \frac{1}{2\pi} \iiint_v \mathbf{B} \cdot \nabla\zeta d^3R . \quad (6.1.14)$$

The volume element d^3R stands for $\sqrt{g}d\varrho d\theta d\zeta$. The derivative of Ψ_{tor} with respect to ϱ , is equal to:

$$\dot{\Psi}_{\text{tor}} \equiv \frac{d\Psi_{\text{tor}}}{d\varrho} = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta \sqrt{g} \mathbf{B} \cdot \nabla\zeta . \quad (6.1.15)$$

In the literature, this quantity, $\dot{\Psi}_{\text{tor}}$ is sometimes referred to as a magnetic flux density (see, e.g., Hirshman and Sigmar 1981). We do not use this name here, since a flux density in physics is commonly identified with the magnetic induction (field) \mathbf{B} . We can substitute $\sqrt{g}\mathbf{B} \cdot \nabla\zeta = \sqrt{g}B^\zeta$ from (6.1.6b) in order to obtain

$$\dot{\Psi}_{\text{tor}} = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta \frac{\partial v}{\partial \theta} . \quad (6.1.16)$$

From the general form of v in (6.1.13), we find that:

$$\frac{\partial v}{\partial \theta} = B(\varrho) + \frac{\partial \tilde{v}}{\partial \theta} . \quad (6.1.17)$$

The integral over θ of the function $\partial \tilde{v}/\partial \theta$ vanishes because \tilde{v} is periodic in θ . The other integration is trivial:

$$\dot{\Psi}_{\text{tor}} = 2\pi B(\varrho) . \quad (6.1.18)$$

Analogous manipulations for the flux Ψ_{pol}^r lead to a similar result:

$$\begin{aligned} \dot{\Psi}_{\text{pol}}^r &\equiv \frac{d\Psi_{\text{pol}}^r}{d\varrho} = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta \sqrt{g} \mathbf{B} \cdot \nabla\theta \\ &= -\frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta \frac{\partial v}{\partial \zeta} = -2\pi C(\varrho) . \end{aligned} \quad (6.1.19)$$

As a result of (6.1.18) and (6.1.19), the function v becomes:

$$v(\varrho, \theta, \zeta) = \frac{1}{2\pi} (\dot{\Psi}_{\text{tor}}\theta - \dot{\Psi}_{\text{pol}}^r\zeta) + \tilde{v}(\varrho, \theta, \zeta) . \quad (6.1.20)$$

If \tilde{v} in (6.1.20) happens to be zero or a flux-surface function, the θ and ζ coordinates we started from are already flux coordinates. Indeed, because the field lines lie on constant v surfaces (6.1.9) and (6.1.20) would give for the equation of a field line

$$\dot{\Psi}_{\text{tor}}\theta - \dot{\Psi}_{\text{pol}}^r\zeta = \text{constant} . \quad (6.1.21)$$

This is the equation of a straight line in (θ, ζ) coordinates. Thus, $\zeta_f = \zeta$ and $\theta_f = \theta$.

If \tilde{v} is a function of θ and ζ , we perform a change of variables in order to eliminate \tilde{v} :

$$\theta_f = \theta + 2\pi \frac{\tilde{v}}{\dot{\Psi}_{\text{tor}}} \quad \text{and} \quad \zeta_f = \zeta \quad (6.1.22a)$$

or

$$\theta_f = \theta \quad \text{and} \quad \zeta_f = \zeta - 2\pi \frac{\tilde{v}}{\dot{\Psi}_{\text{pol}}} . \quad (6.1.22b)$$

Note that these transformations are consistent with (4.6.2). The function v in terms of θ_f and ζ_f becomes:

$$v(\varrho, \theta_f, \zeta_f) = \frac{1}{2\pi} \dot{\Psi}_{\text{tor}} \theta_f - \frac{1}{2\pi} \dot{\Psi}_{\text{pol}} \zeta_f . \quad (6.1.23)$$

Again, $v = \text{constant}$ is the equation of a straight line

$$\dot{\Psi}_{\text{tor}} \theta_f - \dot{\Psi}_{\text{pol}} \zeta_f = \text{constant} . \quad (6.1.24)$$

The contravariant components of \mathbf{B} are still defined by (6.1.6) but with θ replaced by θ_f and ζ by ζ_f :

$$\mathbf{B}^\varrho = 0 \quad (6.1.25a)$$

$$B^{\theta_f} = \frac{1}{2\pi\sqrt{g_f}} \dot{\Psi}_{\text{pol}} \quad (6.1.25b)$$

$$B^{\zeta_f} = \frac{1}{2\pi\sqrt{g_f}} \dot{\Psi}_{\text{tor}} . \quad (6.1.25c)$$

From these equations for B^{θ_f} and B^{ζ_f} , we observe that

$$\frac{B^{\theta_f}}{B^{\zeta_f}} = \frac{\dot{\Psi}_{\text{pol}}}{\dot{\Psi}_{\text{tor}}} = \text{flux function} , \quad (6.1.26)$$

thus proving a result that we anticipated before, in (4.8.10), during the “derivation” of the expression for the rotational transform. The fact that B^{θ_f}/B^{ζ_f} is a flux function is crucial for the straightness of \mathbf{B} . From the equation for a field line,

$$\frac{d\theta_f}{d\zeta_f} = \frac{B^{\theta_f}}{B^{\zeta_f}} = \frac{\dot{\Psi}_{\text{pol}}}{\dot{\Psi}_{\text{tor}}} = \tau(\varrho) , \quad (6.1.27)$$

we see that the slope of the line is indeed constant on the flux surface, and equal to the rotational transform τ .

In summary, a transformation of one of the coordinates is sufficient to make the field lines straight. In tokamaks, it is customary to retain the symmetry angle (which measures rotation about the major axis) $\zeta = \zeta_f$. Consequently, θ then must be changed.

We should remark that the angles θ_f and ζ_f are not the only ones in which the magnetic field appears straight. Indeed, for example, if we further transform

the θ_f and ζ_f coordinates into new coordinates θ_F and ζ_F , as follows:

$$\theta_F = \theta_f + \dot{\Psi}_{\text{pol}} G(\varrho, \theta_f, \zeta_f) \quad (6.1.28a)$$

$$\zeta_F = \zeta_f + \dot{\Psi}_{\text{tor}} G(\varrho, \theta_f, \zeta_f) , \quad (6.1.28b)$$

where G is an arbitrary periodic function, we see that upon substitution of these equations into (6.1.23) for v , we obtain

$$v(\varrho, \theta_F, \zeta_F) = \frac{1}{2\pi} \dot{\Psi}_{\text{tor}} \theta_F - \frac{1}{2\pi} \dot{\Psi}_{\text{pol}} \zeta_F , \quad (6.1.29)$$

showing that θ_F and ζ_F are flux coordinates as well. One can use this degree of freedom to deform the coordinates further to make other quantities look simpler. In Hamada coordinates, the system is deformed until the current-density lines become straight in addition to the magnetic-field lines. Boozer uses this freedom to free the magnetic scalar potential (that would exist in vacuum) from a periodic part. In axisymmetric devices, it is customary to choose θ_f such that $\nabla \theta_f \perp \nabla \zeta_o$, where ζ_o is the ignorable coordinate. These three choices have been discussed qualitatively above, but will be considered in detail below.

Using the flux coordinates θ_f and ζ_f , the magnetic field in Clebsch form is equal to:

$$\mathbf{B} = \nabla \varrho \times \nabla \left(\frac{\dot{\Psi}_{\text{tor}}}{2\pi} \theta_f - \frac{\dot{\Psi}_{\text{pol}}}{2\pi} \zeta_f \right) . \quad (6.1.30)$$

In terms of the poloidal flux through a disk, Ψ_{pol}^d , the expression for \mathbf{B} is

$$\mathbf{B} = \nabla \varrho \times \nabla \left(\frac{\dot{\Psi}_{\text{tor}}}{2\pi} \theta_f + \frac{\Psi_{\text{pol}}^d}{2\pi} \zeta_f \right) . \quad (6.1.31)$$

Here we have used the fact that $d\Psi_{\text{pol}}^d = -d\Psi_{\text{pol}}$ (see (4.7.4)).

If the flux surface volume V is chosen as a flux-surface label, (6.1.30) becomes

$$\mathbf{B} = \nabla V \times \nabla \left(\frac{1}{2\pi} \frac{d\Psi_{\text{tor}}}{dV} \theta_f - \frac{1}{2\pi} \frac{d\Psi_{\text{pol}}^d}{dV} \zeta_f \right) . \quad (6.1.32)$$

In tokamaks, it is customary to use the poloidal flux as a flux label, and to use the safety factor q . If one takes $\varrho \equiv \Psi_{\text{pol}}^d/2\pi \equiv \Psi$, one obtains from (6.1.30)

$$\mathbf{B} = \nabla \Psi \times \nabla (q\theta_f - \zeta_f) . \quad (6.1.33)$$

For the poloidal disk flux as a flux label, the expression is exactly the same as in (6.1.33), but with $\varrho \equiv \Psi \equiv -\Psi_{\text{pol}}^d/2\pi$.

In stellarators, the choice of flux label ϱ is commonly Ψ_{tor} , such that with $\varrho \equiv \psi \equiv \Psi_{\text{tor}}/2\pi$, we obtain

$$\mathbf{B} = \nabla \psi \times \nabla (\theta_f - \zeta_f) . \quad (6.1.34)$$

The general Clebsch form of (6.1.30) can be written as a sum of two terms

$$\mathbf{B} = \frac{\dot{\Psi}_{\text{pol}}}{2\pi} (\nabla\zeta_f \times \nabla\varrho) + \frac{\dot{\Psi}_{\text{tor}}}{2\pi} (\nabla\varrho \times \nabla\theta_f) \equiv \mathbf{B}_P + \mathbf{B}_T . \quad (6.1.35)$$

\mathbf{B}_P is called the poloidal field, and \mathbf{B}_T the toroidal field. Note that \mathbf{B}_P does not lie along $\nabla\theta_f$ but along e_{θ_f} . Analogously, \mathbf{B}_T is not directed along $\nabla\zeta_f$ but along e_{ζ_f} . Only in axisymmetric devices is $\nabla\zeta_f \propto e_{\zeta_f} \propto \hat{\ell}_f$, so, $\mathbf{B}_T \propto \hat{\ell}_f$ if ζ_f is chosen as symmetry angle.

The covariant components of \mathbf{B} , $B_i = \mathbf{B} \cdot e_i$ can readily be found from the contravariant components B^i via the use of the metric coefficients:

$$B_i = (g_{ij})_f B^j = g_{i\theta_f} B^{\theta_f} + g_{i\zeta_f} B^{\zeta_f} . \quad (6.1.36)$$

The metric coefficients are those of the flux-coordinate system (θ_f, ζ_f) .

With the form of \mathbf{B} given in (6.1.30), the symbol $\sqrt{g_f}$ could be replaced in all of the above as follows:

$$\sqrt{g_f} = \sqrt{g_f(\varrho, \theta_f, \zeta_f)} = \frac{\dot{\Psi}_{\text{pol}}}{2\pi} (\mathbf{B} \cdot \nabla\theta_f)^{-1} , \quad (6.1.37)$$

or, for example, in the tokamak case (6.1.33):

$$\sqrt{g_f} = \sqrt{g_f(\Psi, \theta_f, \zeta_f)} = (\mathbf{B} \cdot \nabla\theta_f)^{-1} . \quad (6.1.38)$$

The straight \mathbf{B} coordinate systems are discussed in many references. They basically all follow the treatment discussed in Appendix I of the classic paper by Greene and Johnson (1962). The possibility of writing \mathbf{B} as $\nabla\varrho \times \nabla\Psi$ was suggested by Kruskal and Kulnsrud (1958), and the transformation to angles in which \mathbf{B} lines are straight, was, according to Greene and Johnson, first used by Kadomtsev (1960) and Mercier (1961) – Refs. 3 and 8 in the Greene and Johnson paper. Discussions similar to those in Greene and Johnson (1962) can be found in Bateman (1978), p. 125 (note a disturbing misprint on page 126: the designations contravariant and covariant basis vectors must be reversed); Solov'ev and Shafranov (1970); Miyamoto (1980); Hazeltine and Meiss (1985); Kieras (1982). Important also are the contributions by Shafranov (1968), and Solov'ev (1975). For an alternative approach to straight \mathbf{B} toroidal-flux coordinates, see the sequence of Boozer's papers (1980–1983).

6.2 Symmetry Flux Coordinates in a Tokamak

In an ideal axisymmetric tokamak, common sense suggests that we take the symmetry angle ζ_o as the toroidal angle. The ζ_o -coordinate surfaces are vertical planes so that $\nabla\zeta_o$ points in the symmetry direction. In axisymmetric devices the flux surfaces are such that

$$\nabla\Psi \cdot \nabla\zeta_o = 0 . \quad (6.2.1)$$

To have $\partial/\partial\zeta_o|_{\theta_f, \varrho} \equiv 0$ for every physical quantity, the $\theta_f = \text{constant}$ surfaces must be chosen so that

$$\nabla\zeta_o \cdot \nabla\theta_o = 0 . \quad (6.2.2)$$

We use a subscript “o” for angles (flux coordinates or not) that satisfy (6.2.1) and (6.2.2) to indicate their symmetry aspects. (Thus we drop the subscript “f” in favor of “o”.) These statements are equivalent to requiring that the ζ_o -coordinate curves are circles, whose tangents point in the symmetry direction:

$$e_{\zeta_o} \propto \nabla\zeta_o . \quad (6.2.3)$$

In general, we *cannot* take $\nabla\varrho \cdot \nabla\theta_o = 0$. If we insist on doing so, we have a completely orthogonal system, and this spoils the straightness of \mathbf{B} lines; because of the Shafranov shift amongst other effects, the $\theta = \text{constant}$ and $\varrho = \text{constant}$ surfaces are not orthogonal. Because of (6.2.1) and (6.2.2), we can say that we work in a partially orthogonal system.

The fact that ζ_o is an ignorable coordinate allows us to write \mathbf{B} in an alternative form. The \mathbf{B} -field representation in Clebsch form is still valid in (θ_o, ζ_o) flux coordinates

$$\mathbf{B} = \nabla\Psi \times \nabla(q\theta_o - \zeta_o) = q(\nabla\Psi \times \nabla\theta_o) + (\nabla\zeta_o \times \nabla\Psi) . \quad (6.2.4)$$

We have chosen the poloidal flux as a flux label for simplicity (see (6.1.33)). We represent the distance from the major (symmetry) axis to a particular point on the surface by R ; it represents the radial coordinate in a cylindrical coordinate system ($R, \zeta_c = \pi/2 - \zeta_o, z$). It is then clear that

$$|\nabla\zeta_o| = \frac{1}{R} . \quad (6.2.5)$$

From orthogonality, $\nabla\Psi \cdot \nabla\zeta_o = 0 = \nabla\theta_o \cdot \nabla\zeta_o$, it follows that $\nabla\Psi \times \nabla\theta_o$ is parallel to $\nabla\zeta_o$:

$$\nabla\Psi \times \nabla\theta_o = k \nabla\zeta_o . \quad (6.2.6)$$

k is the proportionality factor. The Jacobian $\sqrt{g_o}$ is defined as usual by $(\nabla\zeta_o \cdot \nabla\Psi \times \nabla\theta_o)^{-1}$, so that from (6.2.6)

$$\sqrt{g_o} = \frac{1}{k |\nabla\zeta_o|^2} . \quad (6.2.7)$$

This can be solved for k , as g_o is known,

$$k = \frac{1}{\sqrt{g_o} |\nabla\zeta_o|^2} . \quad (6.2.8)$$

The toroidal field $\mathbf{B}_T = q \nabla \Psi \times \nabla \theta_o$ can thus be written upon combining (6.2.6) and (6.2.8)

$$\mathbf{B}_T = q(\nabla \Psi \times \nabla \theta_o) = \frac{q \nabla \zeta_o}{\sqrt{g_o} |\nabla \zeta_o|^2} \quad (6.2.9)$$

$$= \frac{qR^2}{\sqrt{g_o}} \nabla \zeta_o . \quad (6.2.10)$$

It is customary to introduce a separate symbol for the factor in front of the $\nabla \zeta_o$:

$$I \equiv \frac{qR^2}{\sqrt{g_o}} . \quad (6.2.11)$$

This I can also be written as

$$I = \sqrt{g_o} (\nabla \Psi \times \nabla \theta_o) \cdot \mathbf{B} = \frac{\sqrt{g_o}}{q} \mathbf{B}_T \cdot \mathbf{B} = \frac{\sqrt{g_o}}{q} \mathbf{B}_T \cdot \mathbf{B}_T . \quad (6.2.12)$$

The magnetic field in a tokamak then becomes

$$\mathbf{B} = I \nabla \zeta_o + (\nabla \zeta_o \times \nabla \Psi) , \quad (6.2.13)$$

which is the form used most often in the tokamak literature.

An important feature of the “function” I is that it is a flux-surface function: $I \equiv I(\Psi)$ for an isotropic pressure plasma equilibrium. This can be shown by considering Ampère’s law:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \quad (6.2.14)$$

$$(\nabla \times \mathbf{B})^i = \frac{1}{\sqrt{g}} \left(\frac{\partial}{\partial u^j} B_k - \frac{\partial}{\partial u^k} B_j \right) , \quad i, j, k \text{ cyc } 1, 2, 3 . \quad (6.2.15)$$

For the radial (Ψ) component, this becomes

$$J^\Psi = \mu_0^{-1} (\nabla \times \mathbf{B})^\Psi = \frac{1}{\mu_0 \sqrt{g_o}} \left(\frac{\partial}{\partial \theta_o} B_{\zeta_o} - \frac{\partial}{\partial \zeta_o} B_{\theta_o} \right) . \quad (6.2.16)$$

Because, on the one hand, axisymmetry implies that $\partial/\partial \zeta_o \equiv 0$, and on the other hand $\mathbf{J} \cdot \nabla \Psi = 0$ requires that $J^\Psi = 0$, we have

$$\frac{\partial}{\partial \theta_o} B_{\zeta_o} = 0 . \quad (6.2.17)$$

The third covariant component of \mathbf{B} , $B_3 = \mathbf{B} \cdot \mathbf{e}_3$ turns out to be just equal to I :

$$B_{\zeta_o} \equiv \mathbf{B} \cdot \mathbf{e}_{\zeta_o} = \mathbf{B} \cdot (\nabla \Psi \times \nabla \theta_o) \sqrt{g_o} = I . \quad (6.2.18)$$

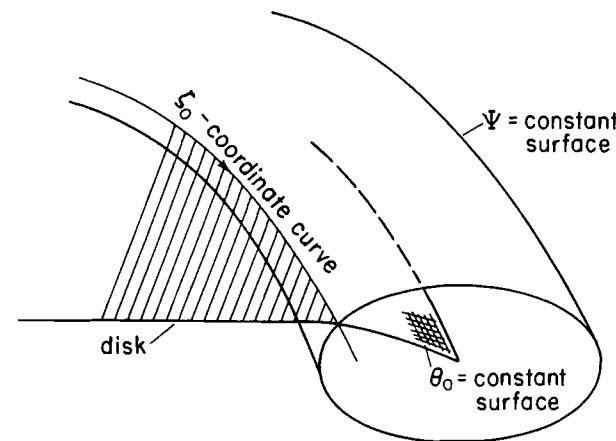


Fig. 6.1. A ζ_o -coordinate curve intersecting a constant Ψ and θ_o surface in an axisymmetric tokamak

Here we used (6.2.12). Thus,

$$\frac{\partial}{\partial \theta_o} I = 0 , \quad (6.2.19)$$

and because of axisymmetry,

$$\frac{\partial}{\partial \zeta_o} I = 0 ,$$

hence,

$$I \equiv I(\Psi) . \quad (6.2.20)$$

(When viscosity is included in the force balance, I is a function of θ as well.) The standard form for \mathbf{B} is thus

$$\mathbf{B} = I(\Psi) \nabla \zeta_o + (\nabla \zeta_o \times \nabla \Psi) . \quad (6.2.21)$$

To clarify the physical meaning of the function $I \equiv I(\Psi)$, we consider the integral form of Ampère’s law:

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 \times (\text{enclosed current}) . \quad (6.2.22)$$

We apply this loop integral along a ζ_o -coordinate curve for θ_o values on the inside of a toroidal surface as shown in Fig. 6.1. In Chap. 2, we found that the differential arc length along a coordinate curve u^i equals $dl(i) = |\mathbf{e}_i| du^i$ (without summation over i , because along the u^i curve, $du^j = du^k = 0$ for $j \neq i \neq k$). We then have that $dl_{\zeta_o} = \mathbf{e}_{\zeta_o} d\zeta_o$. From (6.2.3), it is easily established that $\mathbf{e}_{\theta_o} \cdot \mathbf{e}_{\zeta_o} = 0$. Then with $\mathbf{B} = I \nabla \zeta_o + \nabla \zeta_o \times \nabla \Psi$ and $\nabla \zeta_o \times \nabla \Psi \propto \mathbf{e}_{\theta_o}$, we have from the LHS of (6.2.22)

$$\oint (I \nabla \zeta_o + \nabla \zeta_o \times \nabla \Psi) \cdot e_{\zeta_o} d\zeta_o = \oint I \nabla \zeta_o \cdot e_{\zeta_o} d\zeta_o = 2\pi I . \quad (6.2.23)$$

The “enclosed current” is the total poloidal current outside the flux surface which flows through the disk in Fig. 6.1. We represent this enclosed current by I_{pol}^d in analogy with our notations for magnetic fluxes. In conclusion, we have

$$I \equiv \frac{\mu_0}{2\pi} I_{\text{pol}}^d . \quad (6.2.24)$$

Thus the function $I \equiv I(\Psi)$ effectively measures the total poloidal current outside the flux surface $\Psi = \text{constant}$. I_{pol}^d contains both plasma currents and the toroidal-field-coil currents.

6.3 Interlude: Non-Flux Coordinates in Tokamaks

In the previous section we saw that given a proper choice of the *flux coordinates* θ_f and ζ_o in tokamaks (denoted here by θ_o and ζ_o because they satisfy (6.2.1) and (6.2.2)), the Clebsch form of (6.1.33) can be written in the alternative form of (6.2.21):

$$\mathbf{B} = \nabla \Psi \times \nabla (q\theta_o - \zeta_o) \rightarrow \mathbf{B} = I(\Psi) \nabla \zeta_o + (\nabla \zeta_o \times \nabla \Psi) . \quad (6.3.1)$$

In this section we shall show that axisymmetry, which permits us to choose angles θ_o and ζ_o satisfying (6.2.1) and (6.2.2), always allows us to write \mathbf{B} in a form like (6.2.21), even when θ_o and ζ_o are not flux coordinates. Only when θ_o and ζ_o are flux coordinates in addition to satisfying (6.2.1) and (6.2.2), can the arrow in (6.3.1) be reversed. Below we shall discuss the requirement which must be imposed on θ_o in order for it to become a flux-coordinate angle. Furthermore, we shall show how to construct a flux-coordinate angle $\theta_{o,f} \equiv \Theta_o$ given the non-flux coordinate θ_o .

To see that a form like (6.2.21) is possible even in a non-flux coordinate system, we argue as follows. Consider the poloidal flux through a disk,

$$\Psi_{\text{pol}}^d \equiv \iint_{S_{\text{disk}}} \mathbf{B} \cdot d\mathbf{S} = \iint_{S_{\text{disk}}} d\mathbf{S} \cdot \nabla \times \mathbf{A} = \oint_{\Gamma} d\mathbf{l} \cdot \mathbf{A} . \quad (6.3.2)$$

The contour Γ along which the loop integral must be taken is the boundary curve of the disk (where the disk touches the axisymmetric toroidal magnetic surface). Because of axisymmetry, Γ is a circle with circumference $2\pi R$, and \mathbf{A} must be constant along that loop: the magnetic vector potential \mathbf{A} is a physical quantity, not merely a mathematical construct. Hence, we obtain the simple result:

$$\Psi_{\text{pol}}^d = 2\pi A_{\zeta_o} = 2\pi R \tilde{A}_{\zeta_o} . \quad (6.3.3)$$

Here $A_T \equiv A_{\zeta_o} \nabla \zeta_o = \tilde{A}_{\zeta_o} \hat{\zeta}$ so $\tilde{A}_{\zeta_o} = A_{\zeta_o}/R$. (\tilde{A}_{ζ_o} is the *physical* toroidal component, as opposed to the contravariant or covariant components.) We define the function Ψ as

$$\Psi \equiv -\frac{\Psi_{\text{pol}}^d}{2\pi} . \quad (6.3.4)$$

The minus sign is to make Ψ increase radially. Then,

$$\Psi = -A_{\zeta_o} = -R \tilde{A}_{\zeta_o} . \quad (6.3.5)$$

We can now define the poloidal magnetic field \mathbf{B}_p . It is the field that is responsible for the flux through the disk; it is thus everything except the toroidal field \mathbf{B}_T , which is along $\nabla \zeta_o$. We have

$$\begin{aligned} \mathbf{B}_p &\equiv \nabla \times \mathbf{A}_T = \nabla \times (A_{\zeta_o} \nabla \zeta_o) \\ &= A_{\zeta_o} (\nabla \times \nabla \zeta_o) + (\nabla A_{\zeta_o} \times \nabla \zeta_o) . \end{aligned} \quad (6.3.6)$$

The first term vanishes and $A_{\zeta_o} = -\Psi$, so

$$\mathbf{B}_p = -\nabla \Psi \times \nabla \zeta_o . \quad (6.3.7)$$

Ψ is thus a stream function for the poloidal field \mathbf{B}_p .

To find the field $\mathbf{B} = \mathbf{B}_p + \mathbf{B}_T$, we write \mathbf{B}_T (the field along $\nabla \zeta_o$) as

$$\mathbf{B}_T = B_{\zeta_o} \nabla \zeta_o . \quad (6.3.8)$$

If we now use the symbol I to denote $B_{\zeta_o} = R \tilde{A}_{\zeta_o}$ we obtain

$$\mathbf{B} = I \nabla \zeta_o - (\nabla \Psi \times \nabla \zeta_o) . \quad (6.3.9)$$

$B_{\zeta_o} \equiv I$ is a flux function, the proof of which (see (6.2.14–20)) applies also to the present situation. In conclusion, we find

$$\mathbf{B} = I(\Psi) \nabla \zeta_o + (\nabla \zeta_o \times \nabla \Psi) , \quad (6.3.10)$$

which is the required expression. The flux function $I = I(\Psi)$ is

$$I = R B_T . \quad (6.3.11)$$

Because of (6.2.1), expressing orthogonality between $\nabla \zeta_o$ and $\nabla \Psi$, it follows from (6.3.10) that

$$|\nabla \Psi| = R B_p . \quad (6.3.12)$$

So far we have not mentioned the angle θ_o . Let us find out which θ_o we must choose to convert (6.3.9) into a Clebsch form. The conditions in (6.2.1) and (6.2.2) permit us to write $\nabla \zeta_o$ in (6.3.9) as a cross product, $\nabla \zeta_o = c(\nabla \Psi \times \nabla \theta_o)$. To find c , we dot multiply this last expression by $\nabla \zeta_o$ and find:

$$\nabla \zeta_o = \frac{\sqrt{g}}{R^2} (\nabla \Psi \times \nabla \theta_o) . \quad (6.3.13)$$

Here $g \equiv g(\Psi, \theta_o, \zeta_o)$. Equation (6.3.9) for \mathbf{B} is now

$$\mathbf{B} = \nabla \Psi \times \left[\frac{I(\Psi) \sqrt{g}}{R^2} \nabla \theta_o - \nabla \zeta_o \right] . \quad (6.3.14)$$

If $\sqrt{g/R^2}$ is a flux-surface function (and only then), we can move the ∇ operator in front of the square brackets, leading to a Clebsch form (the $\nabla\Psi \times (\)$ will eliminate the secular $\theta_o [d(I\sqrt{g/R^2})/d\Psi] \nabla\Psi$ term). If θ_o were a flux coordinate, the Jacobian would have been such that (6.2.11) was satisfied: $I\sqrt{g/R^2} = q(\Psi)$. Thus, we conclude that it is possible to go from a form like (6.3.9) to a Clebsch form provided θ_o is chosen such that (6.2.11) is satisfied and thus

$$\sqrt{g} \propto R^2 \quad (6.3.15)$$

in order to make $I\sqrt{g/R^2}$ a flux surface quantity.

Note that even in the “elementary” toroidal system (θ_e, ζ_e) (but based on the magnetic axis), we can write \mathbf{B} as in (6.3.9) regardless of the shape of the tokamak’s flux surfaces. However, it is not possible to write \mathbf{B} in a Clebsch form in terms of θ_e .

Because $\mathbf{B} = I\nabla\zeta_o + \mathbf{V}\zeta_o \times \nabla\Psi$ in an axisymmetric device, we can easily write an explicit expression for a newly constructed flux-coordinate angle, $\theta_{o,f} \equiv \Theta_o$, in terms of the old poloidal angle θ_o .

The stream function v of (6.1.20) expressed in the non-flux coordinates $(\varrho = \Psi, \theta = \theta_o, \zeta = \zeta_o)$ is

$$v(\Psi, \theta_o, \zeta_o) = (q\theta_o - \zeta_o) + \tilde{v}(\Psi, \theta_o). \quad (6.3.16)$$

Because of axisymmetry, \tilde{v} is independent of ζ_o ; q is the safety factor defined in (4.8.3) and (4.8.4): $q \equiv d\Psi_{\text{tor}}/d\Psi_{\text{pol}}^r = -d\Psi_{\text{tor}}/d\Psi_{\text{pol}}^d$. For this $v(\Psi, \theta_o, \zeta_o)$, we can compute the B^{ζ_o} component from (6.1.6b):

$$\mathbf{B} \cdot \nabla\zeta_o = B^{\zeta_o} = \frac{1}{\sqrt{g}} \frac{\partial v}{\partial \theta_o} = \frac{1}{\sqrt{g}} \left(q + \frac{\partial \tilde{v}}{\partial \theta_o} \right). \quad (6.3.17)$$

Integration with respect to θ_o gives:

$$\tilde{v}(\Psi, \theta_o) = -q\theta_o + \int_0^{\theta_o} \sqrt{g}(\mathbf{B} \cdot \nabla\zeta_o) d\theta_o. \quad (6.3.18)$$

From \mathbf{B} in (6.3.10) we have that $\mathbf{B} \cdot \nabla\zeta_o = I/R^2$, leading to

$$\tilde{v}(\Psi, \theta_o) = -q\theta_o + \int_0^{\theta_o} \frac{I\sqrt{g}}{R^2} d\theta_o. \quad (6.3.19)$$

To find a flux-coordinate angle $\theta_{o,f} \equiv \Theta_o$, we apply the transformation in (6.1.22a) for $\varrho \equiv \Psi \equiv -\Psi_{\text{pol}}^d/2\pi$ in order to find $\theta_{o,f} = \theta_o + \tilde{v}/q$, or

$$\theta_{o,f} \equiv \Theta_o = \frac{1}{q} \int_0^{\theta_o} \sqrt{g} \frac{I}{R^2} d\theta_o. \quad (6.3.20)$$

This expression shows explicitly that θ_o is already a flux coordinate, equal to Θ_o , if (6.2.11) is satisfied.

6.4 Straight Current-Density-Line Coordinates

The current density \mathbf{J} satisfies two important relationships:

$$\mathbf{J} \cdot \nabla\varrho = 0 \quad (6.4.1)$$

$$\nabla \cdot \mathbf{J} = 0. \quad (6.4.2)$$

Similar equations are obeyed by the magnetic field \mathbf{B} . This means that we can use the same procedure for the current density \mathbf{J} as we did for \mathbf{B} for we looked for coordinates in which \mathbf{B} is straight. Our goal this time is to start from arbitrary angle coordinates θ and ζ and to find coordinates in which the current-density lines are straight (regardless of what the \mathbf{B} lines look like). We will call these coordinates *current coordinates*, and represent them by θ , and ζ_J . The treatment is analogous to that for \mathbf{B} in Sect. 6.1., so we omit details.

The current density \mathbf{J} is presumed to be known. From (6.4.2), we have:

$$\frac{\partial}{\partial \theta}(\sqrt{g} J^\theta) + \frac{\partial}{\partial \zeta}(\sqrt{g} J^\zeta) = 0, \quad (6.4.3)$$

implying that we can write J^θ and J^ζ as derivatives of a Clebsch (stream) function $\eta(\varrho, \theta, \zeta)$:

$$J^\theta = -\frac{1}{\sqrt{g}} \frac{\partial \eta}{\partial \zeta}, \quad J^\zeta = \frac{1}{\sqrt{g}} \frac{\partial \eta}{\partial \theta}. \quad (6.4.4)$$

This means that \mathbf{J} is expressible in a Clebsch form:

$$\mathbf{J} = \nabla\varrho \times \nabla\eta \quad (6.4.5)$$

such that the current-density lines on a flux surface ($\varrho = \text{constant}$) are defined by the equation

$$\eta(\varrho, \theta, \zeta) = \text{constant}. \quad (6.4.6)$$

η is of the form

$$\eta(\varrho, \theta, \zeta) = \frac{1}{2\pi} \dot{I}_{\text{tor}}(\varrho)\theta - \frac{1}{2\pi} \dot{I}_{\text{pol}}^r(\varrho)\zeta + \tilde{\eta}(\varrho, \theta, \zeta), \quad (6.4.7)$$

Here, $\dot{I}_{\text{pol}}^r(\varrho) \equiv dI_{\text{pol}}^r/d\varrho$, where $I_{\text{pol}}^r(\varrho)$ represents the poloidal current flowing *within* the flux surface. I_{pol}^r is the poloidal current-density flux (just as Ψ_{pol}^r is the poloidal magnetic-density flux):

$$I_{\text{pol}}^r = \frac{1}{2\pi} \iiint_V (\mathbf{J} \cdot \nabla\theta) d^3R = \iint_{S_{\text{pol}}} \mathbf{J} \cdot d\mathbf{S}. \quad (6.4.8)$$

One could have used the poloidal current *outside* a flux surface instead, by analogy with the disk flux Ψ_{pol}^d . In that case, the current would include the coil

current. The flux function $I_{\text{tor}}(\varrho)$ represents the total toroidal current flowing within the volume enclosed by the flux surface.

$$I_{\text{tor}}(\varrho) = \frac{1}{2\pi} \iiint_V (\mathbf{J} \cdot \nabla \zeta) d^3R = \iint_{S_{\text{tor}}} \mathbf{J} \cdot d\mathbf{S}. \quad (6.4.9)$$

$\tilde{\eta}$ is a periodic function in θ and ζ . To make \mathbf{J} -lines straight, we must get rid of $\tilde{\eta}$ or make it a flux function. A transformation similar to (6.1.22a, b) will suffice:

$$\theta_J = \theta + \frac{2\pi\tilde{\eta}}{I_{\text{tor}}} \quad \text{and} \quad \zeta_J = \zeta \quad (6.4.10a)$$

or

$$\theta_J = \theta \quad \text{and} \quad \zeta_J = \zeta - \frac{2\pi\tilde{\eta}}{I_{\text{pol}}}. \quad (6.4.10b)$$

The equation of a current-density line in (θ_J, ζ_J) coordinates is

$$\frac{\dot{I}_{\text{tor}}}{2\pi} \theta_J - \frac{\dot{I}_{\text{pol}}}{2\pi} \zeta_J = \text{constant}, \quad (6.4.11)$$

and the Clebsch form for \mathbf{J} reads

$$\mathbf{J} = \nabla \varrho \times \nabla \left(\frac{\dot{I}_{\text{tor}}}{2\pi} \theta_J - \frac{\dot{I}_{\text{pol}}}{2\pi} \zeta_J \right). \quad (6.4.12)$$

The contravariant components of \mathbf{J} are

$$J^\varrho = 0 \quad (6.4.13a)$$

$$J^{\theta_J} = \frac{\dot{I}_{\text{pol}}}{2\pi\sqrt{g_J}} \quad (6.4.13b)$$

$$J^{\zeta_J} = \frac{\dot{I}_{\text{tor}}}{2\pi\sqrt{g_J}}. \quad (6.4.13c)$$

Usually, straight \mathbf{J} -lines are not used if the magnetic field is not straight. It would be interesting, however, if we could start with flux coordinates θ_f and ζ_f in which \mathbf{B} is straight, and deform these so that in addition to \mathbf{B} , \mathbf{J} also becomes straight. If we replaced the θ and ζ by θ_f and ζ_f in the development of this section, everything would remain correct until (6.4.9). However, the transformation equations, (6.4.10), would indeed make \mathbf{J} straight, but they would spoil the straightness of \mathbf{B} . Consequently, we have to invent new transformation relations that make \mathbf{J} straight but keep \mathbf{B} straight. A transformation of the form given in (6.1.28) is adequate (i.e., make $\tilde{\eta}$ vanish or make it a flux function, while keeping the \tilde{v} of \mathbf{B} equal to zero or a flux function). We must deform both θ_f and ζ_f . The coordinates

created in this way are known as Hamada coordinates. They are characterized by a very simple Jacobian. This system will be discussed in Sect. 6.8.

6.5 Covariant \mathbf{B} Components and Their Relationship to the Boozer-Grad Form

In this section, we further investigate a flux coordinate system in which \mathbf{B} is straight regardless of the form of the \mathbf{J} lines. This means that we set $\theta = \theta_f$ and $\zeta = \zeta_f$ in the previous section. We also keep $\tilde{\eta}$ non-zero in (6.4.7).

6.5.1 The Vector Potential

The contravariant components of \mathbf{B} , expressed in flux coordinates, are according to (6.1.25):

$$B^i = \frac{1}{2\pi\sqrt{g_f}} (0, \dot{\psi}_{\text{pol}}^r, \dot{\psi}_{\text{tor}}^r). \quad (6.5.1)$$

These components were derived from a function $v(\varrho, \theta_f, \zeta_f)$ in (6.1.23)

$$v(\varrho, \theta_f, \zeta_f) = \frac{\dot{\psi}_{\text{tor}}^r}{2\pi} \theta_f - \frac{\dot{\psi}_{\text{pol}}^r}{2\pi} \zeta_f \quad (6.5.2)$$

via (6.1.6)

$$B^{\theta_f} = -\frac{1}{\sqrt{g_f}} \frac{\partial v}{\partial \zeta_f}; \quad B^{\zeta_f} = \frac{1}{\sqrt{g_f}} \frac{\partial v}{\partial \theta_f}. \quad (6.5.3)$$

We now want to give physical meaning to the function v . It is clear that it acts as a “potential” function for \mathbf{B} . The expressions for the \mathbf{B} components in (6.5.3) suggest that there might be a “curl” operation involved. The formula for the curl relates covariant and contravariant components as follows:

$$(\nabla \times \mathbf{A})^i = \frac{1}{\sqrt{g}} \left(\frac{\partial A_k}{\partial u^j} - \frac{\partial A_j}{\partial u^k} \right), \quad i, j, k \text{ cyc } 1, 2, 3. \quad (6.5.4)$$

If we identify \mathbf{A} in the above formula with a “vector potential”, which we will denote \mathbf{A}^* , so that $\mathbf{B} = \nabla \times \mathbf{A}^*$, we observe that (6.5.4) is a relationship between the contravariant components of \mathbf{B} , B^i , and the covariant components of \mathbf{A}^* , A_i^* . The meaning of the asterisk will be explained below. For now, it can be disregarded. This suggests that according to (6.5.3), \mathbf{B} is derivable from a “vector potential” \mathbf{A}^* in which v is the first and only non-vanishing component:

$$A_i^* = (-v, 0, 0) \quad (6.5.5a)$$

or

$$\mathbf{A}^* = \left(-\frac{\dot{\Psi}_{\text{tor}}}{2\pi} \theta_f + \frac{\dot{\Psi}_{\text{pol}}^r}{2\pi} \zeta_f, 0, 0 \right). \quad (6.5.5b)$$

Thus $v(\varrho, \theta_f, \zeta_f)$ is a “potential” function in the sense that it is the only component of a “vector potential” \mathbf{A}^* .

The quotation marks and the asterisk in \mathbf{A}^* indicate that although the v and \mathbf{A}^* produce the correct magnetic field \mathbf{B} , they are *not* acceptable as physical quantities themselves because of their non-periodicity, i.e., non-single valuedness in θ_f and ζ_f . This problem is not insurmountable, since it follows from vector identities that as far as \mathbf{B} is concerned, a magnetic vector potential is only defined up to an additive term which is the gradient of any “arbitrary” function. Since $\mathbf{B} = \nabla \times \mathbf{A}$, we have that $\nabla \times (\mathbf{A} + \nabla f) = \nabla \times \mathbf{A} + \nabla \times \nabla f = \nabla \times \mathbf{A} = \mathbf{B}$. Both \mathbf{A} and $\mathbf{A} + \nabla f$ are legitimate mathematical vector potentials for \mathbf{B} ; f is completely arbitrary.

Although it was thought for a long time that a magnetic vector potential is merely a mathematical tool, it has been established that a magnetic vector potential does have a physical meaning (see, e.g., Feynman, Leighton and Sands 1963). In order for \mathbf{A} to be a physical quantity, we must require that it be single valued, i.e., periodic in θ_f and ζ_f . Thus $f(\varrho, \theta_f, \zeta_f)$ is arbitrary with the only limitation that $\mathbf{A} + \nabla f$ must be periodic.

Our strategy is straightforward; we choose a particular form of f such that \mathbf{A}^* of (6.5.5) will be transformed into an acceptable vector potential $\mathbf{A} = \mathbf{A}^* + \nabla f$. To convert \mathbf{A}^* into a periodic vector potential, we must split up f into a periodic part \tilde{f} and a secular part f . The secular part must be chosen such that ∇f cancels the linear terms in θ_f and ζ_f of \mathbf{A}^* in (6.5.5). The components of \mathbf{A}^* appearing in (6.5.5) are covariant components, so that \mathbf{A}^* itself equals

$$\begin{aligned} \mathbf{A}^* &= A_{\varrho}^* \nabla \varrho + A_{\theta_f}^* \nabla \theta_f + A_{\zeta_f}^* \nabla \zeta_f \\ &= \left(-\frac{\dot{\Psi}_{\text{tor}}}{2\pi} \theta_f + \frac{\dot{\Psi}_{\text{pol}}^r}{2\pi} \zeta_f \right) \nabla \varrho . \end{aligned} \quad (6.5.6)$$

The most general form of f for our purpose is

$$f = \left[a + \frac{\Psi_{\text{tor}}(\varrho)}{2\pi} \right] \theta_f + \left[b - \frac{\Psi_{\text{pol}}^r(\varrho)}{2\pi} \right] \zeta_f , \quad (6.5.7)$$

where a and b are constants, independent of ϱ . With this form of f , and with $f = f + \tilde{f}$, ∇f becomes

$$\nabla f = \frac{\dot{\Psi}_{\text{tor}}}{2\pi} \theta_f \nabla \varrho - \frac{\dot{\Psi}_{\text{pol}}^r}{2\pi} \zeta_f \nabla \varrho + \left(a + \frac{\Psi_{\text{tor}}}{2\pi} \right) \nabla \theta_f + \left(b - \frac{\Psi_{\text{pol}}^r}{2\pi} \right) \nabla \zeta_f + \nabla \tilde{f} . \quad (6.5.8)$$

The first two terms of ∇f in (6.5.8), which are secular terms, cancel the secular first covariant component of \mathbf{A}^* in (6.5.5). The new and acceptable vector

potential \mathbf{A} can thus be written as

$$\mathbf{A} = \left(a + \frac{\Psi_{\text{tor}}}{2\pi} \right) \nabla \theta_f + \left(b - \frac{\Psi_{\text{pol}}^r}{2\pi} \right) \nabla \zeta_f + \nabla \tilde{f} \quad (6.5.9)$$

or

$$A_i = \left(0, a + \frac{\Psi_{\text{tor}}}{2\pi}, b - \frac{\Psi_{\text{pol}}^r}{2\pi} \right) + \frac{\partial \tilde{f}}{\partial u^i} , \quad (6.5.10)$$

where i and u^i stand for $\varrho, \theta_f, \zeta_f$. \tilde{f} is a completely arbitrary periodic function. The first component of \mathbf{A} is zero, whereas in \mathbf{A}^* the first component was the only nonzero component. The crucial point here is that any physically acceptable vector potential must have non-vanishing second and third components in addition to those involving \tilde{f} . The vector potential given by Solov'ev (1975) is equivalent to (6.5.10) but with the constants $a = b = 0$ and $\tilde{f} = 0$.

We could have written down the “bracketed” components of \mathbf{A} in (6.5.10) right away from a term by term identification of (6.5.4) with (6.5.3) after replacing the dot by $d/d\varrho$ or $\partial/\partial\varrho$.

6.5.2 Covariant B Components

It is now appropriate to discuss some aspects of Ampère's law

$$\mu_0 \mathbf{J} = \nabla \times \mathbf{B} . \quad (6.5.11)$$

This represents an inhomogeneous vector differential equation for the magnetic field with a source term equal to $\mu_0 \mathbf{J}$. (In component form, it represents a system of three coupled inhomogeneous partial differential equations.) According to the theory of differential equations, if \mathbf{J} is given, the equation is linear and the solution consists of two parts, \mathbf{B}_0 , the solution of the homogeneous equation (also called the “complementary solution”), and \mathbf{B}_J , a particular solution of the inhomogeneous equation, i.e., the driven equation with source term $\mu_0 \mathbf{J}$:

$$\mathbf{B} = \mathbf{B}_0 + \mathbf{B}_J . \quad (6.5.12)$$

\mathbf{B}_0 satisfies

$$\nabla \times \mathbf{B}_0 = 0 , \quad (6.5.13)$$

while \mathbf{B}_J obeys

$$\nabla \times \mathbf{B}_J = \mu_0 \mathbf{J} . \quad (6.5.14)$$

Equation (6.5.13) represents the equation for a magnetic field in vacuum. It implies that \mathbf{B}_0 can be derived from a *magnetic scalar potential* Φ via the gradient

$$\mathbf{B}_0 = \nabla \Phi . \quad (6.5.15)$$

Then the total \mathbf{B} can be written as

$$\mathbf{B} = \nabla\Phi + \mathbf{B}_J . \quad (6.5.16)$$

This problem is somewhat different from those encountered in the theory of simple differential equations because $\nabla \times \nabla h = 0$ for every arbitrary h . If we want to give Φ and \mathbf{B}_J their usual meanings, we must make sure that \mathbf{B}_J is not merely a particular solution of (6.5.14), but that it is the correct physical solution with all the integration constants included, whereas in simple differential equation theory, we need not worry about integration constants for a particular solution. An easy way to find the proper solution for \mathbf{B}_J is to find the simplest solution of (6.5.14) by observation or manipulation. We then check whether that solution is in agreement with physical common sense by means of taking "Ampère loop integrals." Then, to obtain the proper \mathbf{B}_J , we add the gradient of an appropriate scalar function.

Once \mathbf{B}_J is established correctly, we can say that the Φ solution of (6.5.13) and (6.5.15) is the magnetic scalar potential in vacuum. This potential Φ is determined by currents exterior to the region of interest; in our applications, it is related to the coil currents. However, Φ is not always single valued. Since $\mathbf{B}_0 = \nabla\Phi$ is a single-valued function, Φ can have at most linear terms in its secular part. This means that a closed loop integral can give a non-zero result if Φ has such a secular part Φ . For example, for $\Phi = \alpha(\varrho)\theta$:

$$\oint_{\text{closed } \theta \text{ coordinate curve}} \mathbf{B} \cdot d\mathbf{l} = \oint \nabla\Phi \cdot d\mathbf{l} = \oint d\Phi = \alpha(\varrho)2\pi = A\Phi . \quad (6.5.17)$$

For a single-valued (i.e., periodic) Φ , this integral would have vanished. A closed integration loop *not* encircling the $\varrho = 0$ axis—in the above example—also vanishes.

The loop integral $\oint \mathbf{B} \cdot d\mathbf{l}$ around a current-carrying conductor is constant and equal to the current enclosed. We can thus say that the magnetic field \mathbf{B} produced by such a conductor can be derived from a secular (non-periodic) potential. If there exist poloidal or toroidal integration loops that do not encircle current conductors and stay entirely inside the torus, the magnetic field potential due to these conductors is single valued inside the torus. (A *poloidal* (*toroidal*) loop is any closed curve that encircles the *magnetic* (*major*) axis of the toroid.)

As a simple example, consider Figs. 6.2, 3. In Fig. 6.2, a circular loop of wire and a straight wire are shown. Both carry steady currents as indicated. In a (r, θ, ζ) coordinate system, with $R = R_0 + r \cos \theta$, the magnetic scalar potential is $\Phi = (\mu_0/2\pi)\{I_s\zeta + I_p\theta\}$; \mathbf{B} is computed via $\mathbf{B} = \nabla\Phi = (1/r)(\partial\Phi/\partial\theta)\hat{\theta} + (1/R)(\partial\Phi/\partial\zeta)\hat{\zeta}$. In Fig. 6.3, we show a torus and several sets of coils. There is no plasma present. Every closed contour that remains inside the torus will never enclose the helical windings or the equilibrium coils. Hence, these coils "give rise to" a periodic potential. The coil in the middle (on the minor axis) can easily be encircled and consequently, it will "generate" a potential that is secular along the path of the loop.

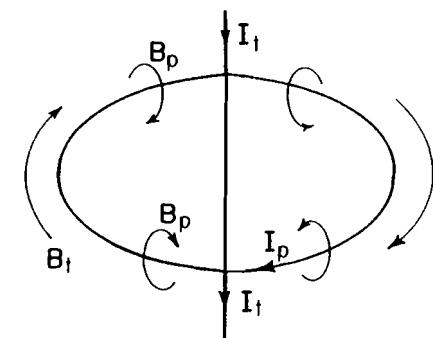


Fig. 6.2. "Toroidal" and "poloidal" magnetic-field behavior for a straight wire in the center of a circular loop. The current I_p generates B_p and B_t is due to I_t .

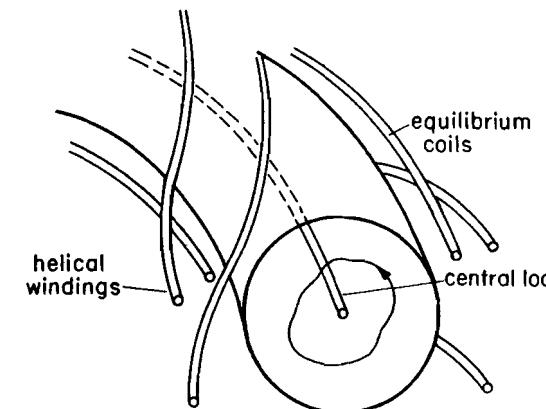


Fig. 6.3. A closed integration curve which encircles a central current carrying loop picks up a secular term in the scalar magnetic potential. None of the helical windings is encircled by an integration curve lying entirely inside the torus, and hence, these windings can only generate a single-valued (or periodic) magnetic potential, inside the torus

Scalar magnetic "potentials" with a secular part as discussed above are only mathematical potentials because they are not single valued. In the latter part of this section, we will consider only physical or single valued (i.e., periodic) potentials.

Now we find the covariant components of \mathbf{B}_J in terms of the contravariant current components. Ampere's law for \mathbf{B}_J given in (6.5.14) reads in component form:

$$J^i = \frac{1}{\mu_0} (\nabla \times \mathbf{B})^i = \frac{1}{\mu_0} \frac{1}{\sqrt{g}} \left(\frac{\partial B_k}{\partial u^j} - \frac{\partial B_j}{\partial u^k} \right) , \quad i, j, k \text{ cyc } 1, 2, 3 . \quad (6.5.18)$$

We have omitted the index J on the \mathbf{B} -components for clarity. Here we are dealing with \mathbf{B} due to currents unless stated otherwise. The contravariant current-density components were given in (6.4.4), with η given by (6.4.7); however, here we are interested in $\theta = \theta_f$ and $\zeta = \zeta_f$. With these components on the left side, (6.5.18) leads to three coupled equations:

$$J^\theta = 0 = \left(\frac{\partial B_\zeta}{\partial \theta_f} - \frac{\partial B_\theta}{\partial \zeta_f} \right) \quad (6.5.19a)$$

$$J_{\theta_f} = \frac{1}{\sqrt{g_f}} \left[\frac{1}{2\pi} \frac{\partial I_{\text{pol}}^r}{\partial \varrho} - \frac{\partial \tilde{\eta}}{\partial \zeta_f} \right] = \frac{1}{\mu_0} \frac{1}{\sqrt{g_f}} \left(\frac{\partial B_\varrho}{\partial \zeta_f} - \frac{\partial B_{\zeta_f}}{\partial \varrho} \right) \quad (6.5.19b)$$

$$J_{\zeta_f} = \frac{1}{\sqrt{g_f}} \left[\frac{1}{2\pi} \frac{\partial I_{\text{tor}}}{\partial \varrho} + \frac{\partial \tilde{\eta}}{\partial \theta_f} \right] = \frac{1}{\mu_0} \frac{1}{\sqrt{g_f}} \left(\frac{\partial B_{\theta_f}}{\partial \varrho} - \frac{\partial B_\varrho}{\partial \theta_f} \right). \quad (6.5.19c)$$

Here we have set $\dot{I}(\varrho) \equiv dI/d\varrho \equiv \partial I/\partial \varrho$, because $I \equiv I(\varrho)$. A glance at the second and third equations suggests the identification

$$B_\varrho^* = -\mu_0 \tilde{\eta}(\varrho, \theta_f, \zeta_f) \quad (6.5.20a)$$

$$B_{\theta_f}^* = \frac{\mu_0}{2\pi} I_{\text{tor}}(\varrho) \quad (6.5.20b)$$

$$B_{\zeta_f}^* = -\frac{\mu_0}{2\pi} I_{\text{pol}}^r(\varrho). \quad (6.5.20c)$$

The solutions are consistent with the first equation (6.5.19a) as can be checked by substitution.

To check whether the components in (6.5.20) make sense physically, we consider the loop integrals along the θ_f and ζ_f coordinate curves, respectively.

The geometry for the θ_f loop integral is depicted in Fig. 6.4. As at the end of Sect. 6.2., we borrow from Chap. 2 the result that a differential arc length vector along the u^i coordinate curve equals $dl(i) = e_i du^i$. There is no summation over i here, since along the u^i curve $du^j = 0 = du^k$, where $j \neq i \neq k$. With $B = \sum_j B_j e^j$, we have then the θ_f -loop integral:

$$\begin{aligned} \oint_{\Gamma_{\theta_f}} \mathbf{B} \cdot dl &= \oint_{\Gamma_{\theta_f}} \sum_j B_j e^j \cdot e_i du^i \\ &= \oint_{\Gamma_{\theta_f}} B_i du^i \quad (\text{no summation over } i) \\ &= \int_0^{2\pi} B_{\theta_f} d\theta_f. \end{aligned} \quad (6.5.21)$$

From Ampère's law:

$$\oint \mathbf{B} \cdot dl = \mu_0 \times (\text{enclosed current})$$

$$\int_0^{2\pi} B_{\theta_f} d\theta_f = \mu_0 I_{\text{tor}}(\varrho). \quad (6.5.22)$$

This relationship is in agreement with the B_{θ_f} component suggested in (6.5.20). (Boozer (1981) uses an alternative form of $dl(i)$ to arrive at the same conclusion. As shown in Chap. 2, $dl(i) = \sqrt{g} (\nabla u^i \times \nabla u^k) du^i$, with $i, j, k \in \{1, 2, 3\}$. To find (6.5.22), the definition of $(\sqrt{g})^{-1} = \nabla u^i \cdot (\nabla u^j \times \nabla u^k)$ must be used.)

For the toroidal component B_{ζ_f} , we consider a contour integral along the ζ_f curve (Fig. 6.5). Performing this integral, we see that we pick up the contribution from the poloidal current $I_{\text{pol}}^{\text{out}}$ which represents the plasma current outside the

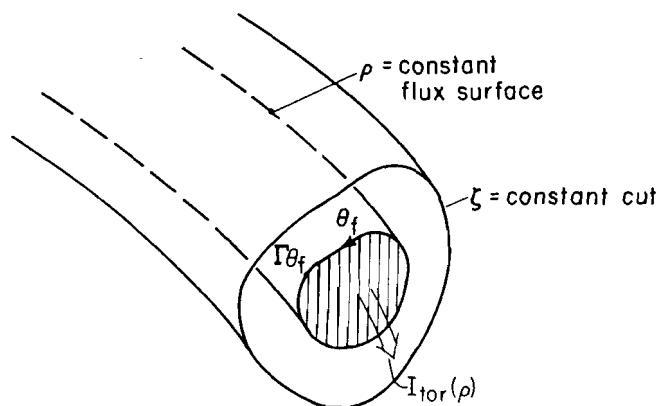


Fig. 6.4. Geometry for integration along a poloidal coordinate curve Γ_{θ_f} , lying at the intersection of a $\varrho = \text{constant}$ flux surface and a $\zeta = \text{constant}$ cut

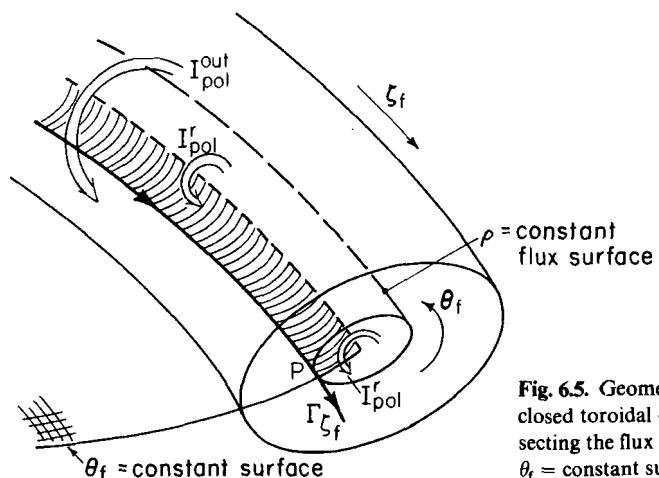


Fig. 6.5. Geometry for integration along a closed toroidal coordinate curve Γ_{ζ_f} , intersecting the flux surface $\varrho = \text{constant}$ and a $\theta_f = \text{constant}$ surface

flux surface $\varrho = \text{constant}$:

$$\int_0^{2\pi} B_{\zeta_f} d\zeta_f = \mu_0 I_{\text{pol}}^{\text{out}}(\varrho) = \mu_0 (I_{\text{pol}}^{\text{plas}} - I_{\text{pol}}^r). \quad (6.5.23)$$

$I_{\text{pol}}^{\text{plas}}$ is the total poloidal current inside the plasma. (Note that the current inside the ribbon I_{pol}^r does not contribute to the field at point P (Fig. 6.5).) Coil currents are not included in this part of the discussion because we are still solving for the components of \mathbf{B}_J , which is the magnetic field component due to only the plasma currents. Equation (6.5.23) requires that the solution for B_{ζ_f} should include an integration constant $\mu_0 I_{\text{pol}}^{\text{plas}}/2\pi$. Instead of (6.5.20), we have

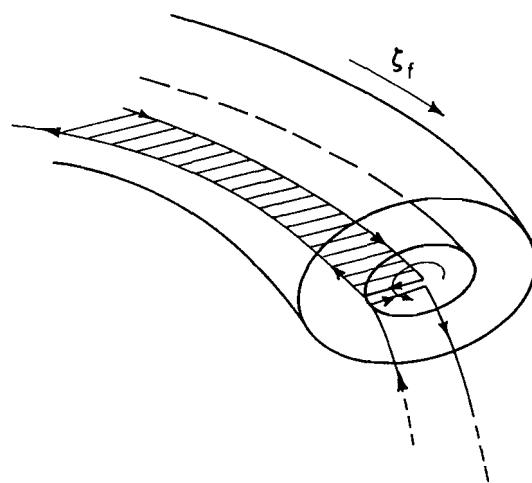


Fig. 6.6. Integration along a closed curve made up of a toroidal coordinate curve, the magnetic axis and two infinitesimally close connection pieces

$$(B_\varrho)_J = -\mu_0 \tilde{\eta}(\varrho, \theta_f, \zeta_f) \quad (6.24a)$$

$$(B_{\theta_f})_J = \frac{\mu_0}{2\pi} I_{\text{tor}}(\varrho) \quad (6.24b)$$

$$(B_{\zeta_f})_J = \frac{\mu_0}{2\pi} (I_{\text{pol}}^{\text{plas}} - I_{\text{pol}}^r(\varrho)) . \quad (6.24c)$$

Equation (6.24c) is consistent with (6.23).

An alternative way to look at this result for $(B_{\zeta_f})_J$ is to consider the loop along a ζ_f -curve that encloses the poloidal ribbon current $I_{\text{pol}}^r(\varrho)$ as shown in Fig. 6.6. Then Ampère's loop law becomes

$$\oint \mathbf{B} \cdot d\mathbf{l} = - \int_0^{2\pi} B_{\zeta_f}(\varrho) d\zeta_f + \int_0^{2\pi} B_{\zeta_f}(0) d\zeta_f = \mu_0 I_{\text{pol}}^r(\varrho) , \quad (6.25)$$

with

$$\int_0^{2\pi} B_{\zeta_f}(0) d\zeta_f = \mu_0 I_{\text{pol}}^{\text{plas}} . \quad (6.26)$$

Thus the loop integral for $B_{\zeta_f}(\varrho)$ equals (6.23).

The argument developed above is equivalent to saying that the assumed particular solution \mathbf{B}_J^* must be augmented by the gradient of a function $h = (\mu_0/2\pi) I_{\text{pol}}^{\text{plas}} \zeta_f$ such that $\mathbf{B}_J = \mathbf{B}_J^* + \nabla h$.

Next we want to find the structure of the vacuum scalar potential Φ . It is the vacuum field component \mathbf{B}_0 or scalar magnetic potential Φ that will bring the coil-current contributions into the picture. Referring to Fig. 6.3, we see that there are no closed poloidal loops that enclose helical or equilibrium coils as long as we stay inside the torus. The central conductor parallel to the minor axis

in Fig. 6.3 is absent except in a levitron or levitated octupole producing a contribution to the potential $(\mu_0/2\pi) I_{\text{coil}} \theta_f$. We assume it is absent from here on. Usually, however, the scalar potential cannot have a term linear in θ_f (cf. our discussion near (6.5.17)).

For toroidal integration loops, remaining inside the torus, it should be clear that the coils crossing the disk in the hole of the doughnut that carry a unidirectional current will contribute to a net toroidal field and $\oint \mathbf{B} \cdot d\mathbf{l} \sim \oint \nabla \Phi \cdot d\mathbf{l} \sim \oint d\Phi \neq 0$. Consequently, the magnetic scalar potential must contain a linear (secular) term in ζ_f : $\Phi \sim (\mu_0/2\pi) (\sum I_{\text{coils}}) \zeta_f$. Except for this term, the potential must be single valued. For example, the part of the scalar potential due to the helical coils in a stellarator—which carry currents in opposite directions—is periodic. Thus, we can represent the magnetic scalar potential as follows:

$$\Phi = \Phi' + \tilde{\Phi} = \frac{\mu_0}{2\pi} (\sum I_{\text{coils}}) \zeta_f + \tilde{\Phi}(\varrho, \theta_f, \zeta_f) . \quad (6.27)$$

Note that there exists a hierarchy in the contributions of the coils. The coils that wrap around in the toroidal direction (but which stay outside the torus) only give rise to a single valued potential. No secular part exists. The coils going around the short way can provide a secular part to the potential, since the toroidal loop integrals do not always vanish. Of course, the latter coils can also provide a periodic part. In a stellarator, the helical coils will also intersect the disk in the doughnut hole. However, in a stellarator the current is of opposite sign in neighboring wires so the net poloidal coil current due to these coils is zero. In a torsatron all helical coils conduct current in the same direction; these coils thus provide a net toroidal field and a secular term $\propto \zeta_f$ in the potential.

The result for the covariant \mathbf{B} components is, with $\mathbf{B} = \mathbf{B}_J + \nabla \Phi = \mathbf{B}_J + \nabla \Phi + \nabla \tilde{\Phi}$:

$$B_\varrho = -\mu_0 \tilde{\eta}(\varrho, \theta_f, \zeta_f) + \frac{\partial \tilde{\Phi}(\varrho, \theta_f, \zeta_f)}{\partial \varrho} \quad (6.28a)$$

$$B_{\theta_f} = \frac{\mu_0}{2\pi} I_{\text{tor}}(\varrho) + \frac{\partial \tilde{\Phi}(\varrho, \theta_f, \zeta_f)}{\partial \theta_f} \quad (6.28b)$$

$$B_{\zeta_f} = \frac{\mu_0}{2\pi} (I_{\text{pol}}^{\text{plas}} - I_{\text{pol}}^r(\varrho)) + \frac{\mu_0}{2\pi} (\sum I_{\text{coils}}) + \frac{\partial \tilde{\Phi}(\varrho, \theta_f, \zeta_f)}{\partial \zeta_f} . \quad (6.28c)$$

A more rigorous school of thought requires the magnetic scalar potential to be a physical quantity. This implies that Φ cannot have secular terms and must be periodic. This viewpoint requires that a scalar magnetic potential cannot be used if one considers loop integrals encircling currents (Panofsky and Phillips 1962, p. 127). In our case, it means that the coil-currents which are encircled and consequently contribute to secular terms cannot be described in terms of a physical scalar potential. As a result, (6.5.13) and (6.5.15) would not be equivalent for a *physical* potential Φ . The implication for the covariant \mathbf{B} -components is that we only deal with a periodic potential $\tilde{\Phi}$ and that we absorb $\sum I_{\text{coils}}$ into the integration constant of (6.28c) as follows:

$$I_{\text{pol}}^{\text{tot}} = I_{\text{pol}}^{\text{plas}} + \sum I_{\text{coils}} . \quad (6.5.29)$$

$I_{\text{pol}}^{\text{tot}}$ is the total poloidal current crossing a disk touching the magnetic axis. The poloidal plasma current and the coil currents crossing that disk are included. It is convenient to introduce a symbol for the poloidal current (including the coil currents) outside a flux surface:

$$I_{\text{pol}}^{\text{d}}(\varrho) = I_{\text{pol}}^{\text{tot}} - I_{\text{pol}}^{\text{r}}(\varrho) . \quad (6.5.30)$$

The index "d" stands for *disk*.

The covariant \mathbf{B} -components are thus alternatively

$$B_{\varrho} = -\mu_0 \tilde{\eta}(\varrho, \theta_f, \zeta_f) + \frac{\partial \tilde{\Phi}(\varrho, \theta_f, \zeta_f)}{\partial \varrho} \quad (6.5.31a)$$

$$B_{\theta_f} = \frac{\mu_0}{2\pi} I_{\text{tor}}(\varrho) + \frac{\partial \tilde{\Phi}(\varrho, \theta_f, \zeta_f)}{\partial \theta_f} \quad (6.5.31b)$$

$$B_{\zeta_f} = \frac{\mu_0}{2\pi} I_{\text{pol}}^{\text{d}}(\varrho) + \frac{\partial \tilde{\Phi}(\varrho, \theta_f, \zeta_f)}{\partial \zeta_f} . \quad (6.5.31c)$$

Similar expressions have been given by Shafranov (1968). Although not explicitly stated, it must be understood that our integration constants, in (6.5.28c), have been absorbed in Shafranov's I . See also Hirshman (1982).

6.5.3 Relationship to the Boozer-Grad Form

Now that we know the covariant magnetic-field components, we convert the usual covariant form to the covariant Boozer-Grad form of (5.4.25). The generic covariant representation of \mathbf{B} is

$$\mathbf{B} = B_{\varrho} \nabla \varrho + B_{\theta_f} \nabla \theta_f + B_{\zeta_f} \nabla \zeta_f \quad (6.5.32)$$

with the covariant components given in (6.5.31). After substitution of these components, we can rearrange (6.5.32) as follows:

$$\begin{aligned} \mathbf{B} &= -\mu_0 \tilde{\eta} \nabla \varrho + \frac{\partial \tilde{\Phi}}{\partial \varrho} \nabla \varrho + \frac{\mu_0}{2\pi} I_{\text{tor}} \nabla \theta_f + \frac{\partial \tilde{\Phi}}{\partial \theta_f} \nabla \theta_f + \frac{\mu_0}{2\pi} I_{\text{pol}}^{\text{d}} \nabla \zeta_f + \frac{\partial \tilde{\Phi}}{\partial \zeta_f} \nabla \zeta_f \\ &= \left(\frac{\partial \tilde{\Phi}}{\partial \varrho} \nabla \varrho + \frac{\partial \tilde{\Phi}}{\partial \theta_f} \nabla \theta_f + \frac{\partial \tilde{\Phi}}{\partial \zeta_f} \nabla \zeta_f \right) - \mu_0 \tilde{\eta} \nabla \varrho + \frac{\mu_0}{2\pi} I_{\text{tor}} \nabla \theta_f + \frac{\mu_0}{2\pi} I_{\text{pol}}^{\text{d}} \nabla \zeta_f \\ &= \nabla \tilde{\Phi} - \mu_0 \tilde{\eta} \nabla \varrho + \frac{\mu_0}{2\pi} \nabla (I_{\text{tor}} \theta_f) - \frac{\mu_0}{2\pi} \theta_f I_{\text{tor}} \nabla \varrho \\ &\quad + \frac{\mu_0}{2\pi} \nabla (I_{\text{pol}}^{\text{d}} \zeta_f) - \frac{\mu_0}{2\pi} \zeta_f I_{\text{pol}}^{\text{d}} \nabla \varrho \end{aligned} \quad (6.5.33a)$$

$$\begin{aligned} &= \nabla \tilde{\Phi} + \nabla \left(\frac{\mu_0}{2\pi} I_{\text{tor}} \theta_f \right) + \nabla \left(\frac{\mu_0}{2\pi} I_{\text{pol}}^{\text{d}} \zeta_f \right) \\ &\quad + \nabla \varrho \left(-\mu_0 \tilde{\eta} - \frac{\mu_0}{2\pi} I_{\text{tor}} \theta_f - \frac{\mu_0}{2\pi} I_{\text{pol}}^{\text{d}} \zeta_f \right) \\ &= \nabla \left(\tilde{\Phi} + \frac{\mu_0}{2\pi} I_{\text{tor}} \theta_f + \frac{\mu_0}{2\pi} I_{\text{pol}}^{\text{d}} \zeta_f \right) + \lambda \nabla \varrho . \end{aligned} \quad (6.5.33b)$$

Here we have defined the following "Boozer-Grad quantity"

$$\lambda \equiv -\mu_0 \left(\tilde{\eta} + \frac{I_{\text{tor}}}{2\pi} \theta_f + \frac{I_{\text{pol}}^{\text{d}}}{2\pi} \zeta_f \right) . \quad (6.5.34)$$

In terms of the poloidal ribbon current, $I_{\text{pol}}^{\text{r}}$, for which we have

$$I_{\text{pol}}^{\text{d}} = -I_{\text{pol}}^{\text{r}} , \quad (6.5.35)$$

we obtain

$$\lambda \equiv -\mu_0 \left(\tilde{\eta} + \frac{I_{\text{tor}}}{2\pi} \theta_f - \frac{I_{\text{pol}}^{\text{r}}}{2\pi} \zeta_f \right) \quad (6.5.36)$$

or with (6.4.7)

$$\lambda = -\mu_0 \eta . \quad (6.5.37)$$

With this λ , we obtain for the current density \mathbf{J} which was originally given as $\mathbf{J} = \nabla \varrho \times \nabla \eta$ in (6.4.5)

$$\mathbf{J} = \frac{1}{\mu_0} \nabla \lambda \times \nabla \varrho . \quad (6.5.38)$$

This is exactly the Boozer-Grad form obtained earlier in (5.4.21).

If we introduce the "Boozer-Grad quantity" χ in (6.5.33),

$$\chi = \tilde{\Phi} + \frac{\mu_0}{2\pi} (I_{\text{tor}} \theta_f + I_{\text{pol}}^{\text{d}} \zeta_f) , \quad (6.5.39)$$

the covariant form of \mathbf{B} reduces to precisely the Boozer-Grad form

$$\mathbf{B} = \nabla \chi + \lambda \nabla \varrho , \quad (6.5.40)$$

or with (6.5.37),

$$\mathbf{B} = \nabla \chi - \mu_0 \eta \nabla \varrho . \quad (6.5.41)$$

To make the link with Boozer's work somewhat clearer, we give below a table of our symbols and his (omitting c 's, 4π 's, μ_0 , ...). For completeness, we also give the symbols used by Solov'ev and Shafranov.

Ours	Boozer (1981)	Solov'ev-Shafranov (1970)
$\tilde{\eta}$	$-\beta^*$	$\tilde{\nu}$
λ	β	ν
I_{pol}^d	g	constant- I
I_{tor}	I	J
θ_f	θ	θ
ζ_f	ϕ	ζ
ν	θ_0	ϑ

6.6 Boozer's Toroidal Flux Coordinates

We have already mentioned that the flux coordinates θ_f and ζ_f are not unique. In other words, there is more than one pair of angle functions $\theta_f(\mathbf{R})$ and $\zeta_f(\mathbf{R})$ in which the magnetic field is straight. Boozer makes that particular choice of flux coordinates, which we denote by θ_B and ζ_B , such that the periodic part of the scalar magnetic potential vanishes. Thus, he deforms the (θ_f, ζ_f) coordinate curves so that $\tilde{\Phi} \equiv 0$. This gives his Jacobian a rather simple form. Following Boozer's (1981) paper, we compute the Jacobian from the contravariant and covariant representations of \mathbf{B} , (6.1.30) and (6.5.33a), respectively:

$$\mathbf{B} = \frac{\dot{\Psi}_{\text{tor}}}{2\pi} \nabla \varrho \times \nabla \theta_f - \frac{\dot{\Psi}_{\text{pol}}^r}{2\pi} \nabla \varrho \times \nabla \zeta_f = \frac{\dot{\Psi}_{\text{pol}}^r}{2\pi\sqrt{g_f}} \mathbf{e}_{\theta_f} + \frac{\dot{\Psi}_{\text{tor}}}{2\pi\sqrt{g_f}} \mathbf{e}_{\zeta_f} \quad (6.6.1)$$

and

$$\mathbf{B} = \left(-\mu_0 \tilde{\eta} + \frac{\partial \tilde{\Phi}}{\partial \varrho} \right) \nabla \varrho + \left(\frac{\mu_0}{2\pi} I_{\text{tor}} + \frac{\partial \tilde{\Phi}}{\partial \theta_f} \right) \nabla \theta_f + \left(\frac{\mu_0}{2\pi} I_{\text{pol}}^d + \frac{\partial \tilde{\Phi}}{\partial \zeta_f} \right) \nabla \zeta_f . \quad (6.6.2)$$

Since $\mathbf{B} \cdot \mathbf{B} = (B)^2 = B^i B_i$, we obtain from the dot product of these equations:

$$\begin{aligned} (B)^2 &= \frac{\dot{\Psi}_{\text{pol}}^r}{2\pi\sqrt{g_f}} \left(\frac{\mu_0}{2\pi} I_{\text{tor}} + \frac{\partial \tilde{\Phi}}{\partial \theta_f} \right) + \frac{\dot{\Psi}_{\text{tor}}}{2\pi\sqrt{g_f}} \left(\frac{\mu_0}{2\pi} I_{\text{pol}}^d + \frac{\partial \tilde{\Phi}}{\partial \zeta_f} \right) \\ &= \frac{1}{\sqrt{g_f}} \frac{\mu_0}{4\pi^2} (\dot{\Psi}_{\text{pol}}^r I_{\text{tor}} + \dot{\Psi}_{\text{tor}} I_{\text{pol}}^d) + \frac{\dot{\Psi}_{\text{pol}}^r}{2\pi\sqrt{g_f}} \frac{\partial \tilde{\Phi}}{\partial \theta_f} + \frac{\dot{\Psi}_{\text{tor}}}{2\pi\sqrt{g_f}} \frac{\partial \tilde{\Phi}}{\partial \zeta_f} \end{aligned}$$

or

$$(B)^2 = \frac{1}{\sqrt{g_f}} \frac{\mu_0}{4\pi^2} (\dot{\Psi}_{\text{pol}}^r I_{\text{tor}} + \dot{\Psi}_{\text{tor}} I_{\text{pol}}^d) + \mathbf{B} \cdot \nabla \tilde{\Phi} . \quad (6.6.3)$$

Recall that $\mathbf{B} \cdot \nabla = B^i \partial/\partial u^i$, the B^i being given in (6.1.25). After solving this equation for $\sqrt{g_f}$ we find:

$$\sqrt{g_f} = \frac{\mu_0}{4\pi^2} \frac{(\dot{\Psi}_{\text{pol}}^r I_{\text{tor}} + \dot{\Psi}_{\text{tor}} I_{\text{pol}}^d)}{(B)^2 - \mathbf{B} \cdot \nabla \tilde{\Phi}} . \quad (6.6.4a)$$

Somewhat more symmetrically, this is

$$\sqrt{g_f} = \frac{\mu_0}{4\pi^2} \frac{(\dot{\Psi}_{\text{tor}} I_{\text{pol}}^d - \dot{\Psi}_{\text{pol}}^r I_{\text{tor}})}{(B)^2 - \mathbf{B} \cdot \nabla \tilde{\Phi}} . \quad (6.6.4b)$$

Boozer takes $\tilde{\Phi} \equiv 0$. Therefore, Boozer's Jacobian, with ϱ still used as a radial coordinate, is

$$\sqrt{g_B} = \frac{\mu_0}{4\pi^2} \frac{(\dot{\Psi}_{\text{pol}}^r I_{\text{tor}} + \dot{\Psi}_{\text{tor}} I_{\text{pol}}^d)}{(B)^2} . \quad (6.6.5)$$

Boozer (1981) uses the toroidal flux as a flux label, $\varrho \equiv \Psi_{\text{tor}}$. The Jacobian in this case is denoted by $(\sqrt{g})_B$ and is given by

$$(\sqrt{g})_B^{-1} \equiv \nabla \Psi_{\text{tor}} \cdot \nabla \theta_B \times \nabla \zeta_B . \quad (6.6.6)$$

It relates to the Jacobian $\sqrt{g_B}$ with ϱ as a coordinate, as follows:

$$(\sqrt{g})_B^{-1} = \dot{\Psi}_{\text{tor}} (\nabla \varrho \cdot \nabla \theta_B \times \nabla \zeta_B) = \dot{\Psi}_{\text{tor}} (\sqrt{g_B})^{-1} . \quad (6.6.7)$$

Making use of the definition of iota-bar, $\bar{\tau} = \dot{\Psi}_{\text{pol}}^r / \dot{\Psi}_{\text{tor}}$, we obtain the Jacobian as used by Boozer:

$$(\sqrt{g})_B = \dot{\Psi}_{\text{tor}}^{-1} \frac{\mu_0}{4\pi^2} \left(\frac{\dot{\Psi}_{\text{pol}}^r I_{\text{tor}} + \dot{\Psi}_{\text{tor}} I_{\text{pol}}^d}{(B)^2} \right) \quad (6.6.8a)$$

or

$$(\sqrt{g})_B = \frac{\mu_0}{4\pi^2} \frac{\bar{\tau} I_{\text{tor}} + I_{\text{pol}}^d}{B^2} ; \quad \varrho \equiv \Psi_{\text{tor}} . \quad (6.6.8b)$$

To find the toroidal Boozer flux coordinates θ_B and ζ_B (in which $\tilde{\Phi} \equiv 0$) in terms of other flux coordinates θ_f and ζ_f (in which, in general, $\tilde{\Phi} \neq 0$), we use the freedom provided by (6.1.28):

$$\theta_B = \theta_f + \dot{\Psi}_{\text{pol}}^r G_B(\varrho, \theta_f, \zeta_f) \quad (6.6.9a)$$

$$\zeta_B = \zeta_f + \dot{\Psi}_{\text{tor}} G_B(\varrho, \theta_f, \zeta_f) . \quad (6.6.9b)$$

Recall that this transformation preserves the straightness of the magnetic field lines. A proper choice of G_B gives the coordinate system extra properties such as a zero periodic scalar magnetic potential in Boozer's case.

The new Jacobian $\sqrt{g_B}$ was calculated in (6.6.5), while the old one is written in (6.6.4). Now we look for an alternative relationship between $\sqrt{g_f}$ and $\sqrt{g_B}$ in order to determine G_B . By definition, the inverse Jacobian is

$$(\sqrt{g_B})^{-1} = \nabla \varrho \cdot (\nabla \theta_B \times \nabla \zeta_B) . \quad (6.6.10)$$

Upon substitution of the Boozer angles θ_B and ζ_B from (6.6.9), this becomes

$$\begin{aligned}
 (\sqrt{g_B})^{-1} &= \nabla\varrho \cdot [\nabla(\theta_f + \dot{\Psi}_{pol}^r G_B) \times \nabla(\zeta_f + \dot{\Psi}_{tor} G_B)] \\
 &= \nabla\varrho \cdot (\nabla\theta_f \times \nabla\zeta_f) + \dot{\Psi}_{pol}^r \nabla\varrho \cdot (\nabla G_B \times \nabla\zeta_f) \\
 &\quad + \dot{\Psi}_{tor} \nabla\varrho \cdot (\nabla\theta_f \times \nabla G_B) \\
 &= (\sqrt{g_f})^{-1} + \dot{\Psi}_{pol}^r (\nabla\zeta_f \times \nabla\varrho) \cdot \nabla G_B + \dot{\Psi}_{tor} (\nabla\varrho \times \nabla\theta_f) \cdot \nabla G_B \\
 &= (\sqrt{g_f})^{-1} + \frac{\dot{\Psi}_{pol}^r}{2\pi\sqrt{g_f}} (2\pi)e_{\theta_f} \cdot \nabla G_B + \frac{\dot{\Psi}_{tor}}{2\pi\sqrt{g_f}} (2\pi)e_{\zeta_f} \cdot \nabla G_B
 \end{aligned}$$

or

$$(\sqrt{g_B})^{-1} = (\sqrt{g_f})^{-1} + 2\pi \mathbf{B} \cdot \nabla G_B . \quad (6.6.11)$$

Here we used (6.1.25) for B^i .

This is a magnetic differential equation for G_B . If we replace the Jacobians in (6.6.11) by their computed values given in (6.4.4) and (6.6.5), we obtain

$$\mathbf{B} \cdot \nabla G_B = \frac{2\pi}{\mu_0} \frac{(B)^2}{\dot{\Psi}_{pol}^r I_{tor} + \dot{\Psi}_{tor} I_{pol}^d} - \frac{2\pi}{\mu_0} \frac{(B)^2 - \mathbf{B} \cdot \nabla \tilde{\Phi}}{\dot{\Psi}_{pol}^r I_{tor} + \dot{\Psi}_{tor} I_{pol}^d}$$

or

$$\mathbf{B} \cdot \nabla G_B = \frac{2\pi}{\mu_0} \frac{\mathbf{B} \cdot \nabla \tilde{\Phi}}{\dot{\Psi}_{pol}^r I_{tor} + \dot{\Psi}_{tor} I_{pol}^d} \quad (6.6.12)$$

We find that the simple choice

$$G_B = \frac{2\pi}{\mu_0} \frac{\tilde{\Phi}}{\dot{\Psi}_{pol}^r I_{tor} + \dot{\Psi}_{tor} I_{pol}^d} \quad (6.6.13)$$

allows the unique determination of the toroidal Boozer flux coordinates θ_B and ζ_B via (6.6.9).

We can now relate the Boozer toroidal flux coordinates, $(\varrho, \theta_B, \zeta_B)$, to the Boozer-Grad Clebsch coordinates, (ϱ, v, χ) , given above in (6.1.23) and (6.5.39) with $\tilde{\Phi} \equiv 0$:

$$\varrho = \varrho \quad (6.6.14a)$$

$$v = \frac{\dot{\Psi}_{tor}}{2\pi} \theta_B - \frac{\dot{\Psi}_{pol}^r}{2\pi} \zeta_B \quad (6.6.14b)$$

$$\chi = \frac{\mu_0}{2\pi} I_{tor} \theta_B + \frac{\mu_0}{2\pi} I_{pol}^d \zeta_B . \quad (6.6.14c)$$

The Jacobian $\sqrt{g_B}$ is given in (6.6.5). Note that we still use ϱ as a flux-surface label, whereas Boozer normally uses Ψ_{tor} . The only difference is that $\nabla\Psi_{tor} =$

$\dot{\Psi}_{tor} \nabla\varrho$ and that the $\dot{\Psi}_{tor}$ is absorbed in Boozer's v ; it is for this reason that his v contains an t .

It is interesting to investigate what happens to Boozer's covariant form for \mathbf{B} in vacuum. To do this, we start from the Boozer-Grad form given in (6.5.40), $\mathbf{B} = \nabla\chi + \lambda \nabla\varrho$ with χ and λ given by (6.6.14c) and (6.5.36), respectively. From our discussion in Sect. 5.4., we recall that $\lambda = 0$ when $\mathbf{J} = 0$ (see (5.4.37)). With the replacement $I_{pol}^d = I_{pol}^{plas} + \sum I_{coils} - I_{pol}^r(\varrho) \rightarrow \sum I_{coils}$ in vacuum, we obtain a simple form for the vacuum field:

$$\mathbf{B} = \sum (I_{coils}) \nabla\zeta_B . \quad (6.6.15)$$

The interpretation of this equation is important. It looks like the magnetic field has only a "toroidal" component in vacuum; however, this is not necessarily the case. For a tokamak, $\sum I_{coils}$ stands for the toroidal-field coil currents that encircle the torus poloidally. In the vacuum case, the poloidal and toroidal coordinate curves intersect at right angles (see Fig. 6.7a). In an axisymmetric device, the tangent-basis vector e_ζ and the reciprocal basis vector $\nabla\zeta$ are parallel such that indeed the vacuum field is in the toroidal direction. In a stellarator, $\sum (I_{coils})$ also consists only of the toroidal-field coil currents, since the currents in the helical windings add up to zero. However, we know that stellarator field lines

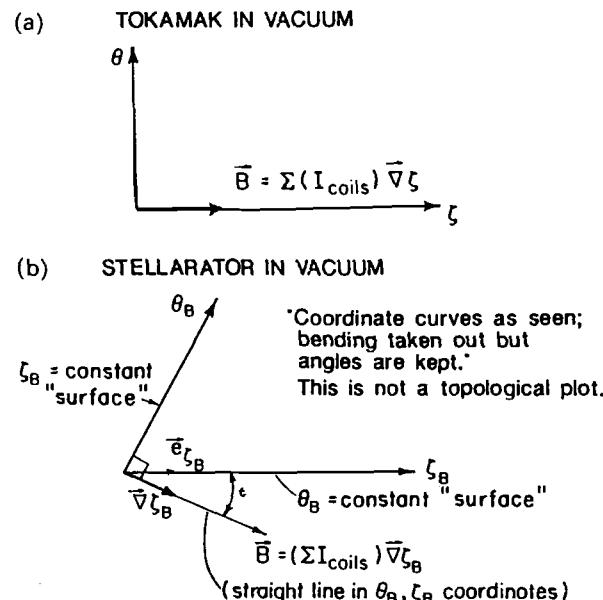


Fig. 6.7a-b. Comparison of the vacuum field in a tokamak and a stellarator. (a) In a tokamak \mathbf{B} points in the $\nabla\zeta \propto \hat{\zeta}$ direction. (b) In a stellarator \mathbf{B} is still in the $\nabla\zeta_B$ direction, but $\nabla\zeta_B$ makes an angle t with the $\hat{\zeta}_B$ direction. (The subscript B refers to Boozer coordinates.)

follow spiralling paths in vacuum and are thus not in the usual toroidal direction. Yet Boozer's expression in (6.6.15) is correct. The explanation lies in Fig. 6.7b. Boozer has adjusted his coordinates so that the periodic scalar potential $\tilde{\Phi}$ vanishes while the \mathbf{B} lines remain straight. As a consequence, the coordinate surfaces $\theta_B = \text{constant}$ and $\zeta_B = \text{constant}$ do not intersect at right angles nor do the coordinate curves. This has the important consequence that the contravariant-basis vector $\nabla\zeta_B$ does not lie in the e_{ζ_B} direction, the latter normally being called the toroidal direction. As shown in Fig. 6.7b, the field line lies in the $\nabla\zeta_B$ direction which makes an angle with the ζ_B coordinate curve. This angle is related to the rotational transform.

6.7 Ideal-MHD-Equilibrium Conditions for Toroidally Confined Plasmas

To find an equation for $\tilde{\eta}$, the periodic part of the current stream function η , we substitute the contravariant components of \mathbf{B} and \mathbf{J} , given in (6.1.25) and the LHS of (6.5.19) in the MHD equilibrium equation $\nabla p = \mathbf{J} \times \mathbf{B}$. The first component of this force balance is, in vector form,

$$\dot{p} \nabla Q = J^{\theta_t} B^{\zeta_t} e_{\theta_t} \times e_{\zeta_t} + J^{\zeta_t} B^{\theta_t} e_{\zeta_t} \times e_{\theta_t}$$

or

$$\dot{p} \nabla Q = J^{\theta_t} B^{\zeta_t} \nabla Q \sqrt{g_f} - J^{\zeta_t} B^{\theta_t} \nabla Q \sqrt{g_f} . \quad (6.7.1)$$

Using

$$B^i = \frac{1}{2\pi\sqrt{g_f}} (0, \dot{\Psi}_{\text{pol}}^r, \dot{\Psi}_{\text{tor}}^r) \quad (6.7.2)$$

$$J^j = \frac{1}{2\pi\sqrt{g_f}} \left(0, \dot{I}_{\text{pol}}^r - 2\pi \frac{\partial \tilde{\eta}}{\partial \zeta_f}, \dot{I}_{\text{tor}}^r + 2\pi \frac{\partial \tilde{\eta}}{\partial \theta_f} \right), \quad (6.7.3)$$

we obtain

$$\frac{1}{4\pi^2\sqrt{g_f}} \left\{ -\dot{I}_{\text{tor}} \dot{\Psi}_{\text{pol}}^r + \dot{I}_{\text{pol}}^r \dot{\Psi}_{\text{tor}}^r - 2\pi \dot{\Psi}_{\text{pol}}^r \frac{\partial \tilde{\eta}}{\partial \theta_f} - 2\pi \dot{\Psi}_{\text{tor}}^r \frac{\partial \tilde{\eta}}{\partial \zeta_f} \right\} = \dot{p} . \quad (6.7.4)$$

Since $\dot{\Psi}_{\text{pol}}^r \sim B^{\theta_t}$, $\dot{\Psi}_{\text{tor}}^r \sim B^{\zeta_t}$ and $\partial \tilde{\eta} / \partial \theta_f = (\nabla \tilde{\eta})_{\theta_f}$, $\partial \tilde{\eta} / \partial \zeta_f = (\nabla \tilde{\eta})_{\zeta_f}$ and $B^e \equiv 0$, the last two terms inside the curly brackets of (6.7.4) represent the dot product of $(4\pi^2\sqrt{g_f} \mathbf{B})$ with $\nabla \tilde{\eta}$. We obtain

$$\frac{-\dot{I}_{\text{tor}} \dot{\Psi}_{\text{pol}}^r + \dot{I}_{\text{pol}}^r \dot{\Psi}_{\text{tor}}^r}{4\pi^2\sqrt{g_f}} - \mathbf{B} \cdot \nabla \tilde{\eta} = \dot{p} . \quad (6.7.5)$$

Thus, we find a magnetic differential equation for $\tilde{\eta}$:

$$\mathbf{B} \cdot \nabla \tilde{\eta} = \frac{\dot{I}_{\text{pol}}^r \dot{\Psi}_{\text{tor}}^r - \dot{I}_{\text{tor}}^r \dot{\Psi}_{\text{pol}}^r - \dot{p}}{4\pi^2\sqrt{g_f}} . \quad (6.7.6)$$

The second and third components of $\nabla p = \mathbf{J} \times \mathbf{B}$ give no useful information because $J^e = B^e = 0$.

In order for (6.7.6) to have a solution, the source term must satisfy two constraint conditions ((4.10.7) and (4.10.4))

$$\iiint_V \{ \} d^3R = 0 \quad (6.7.7)$$

and

$$\oint \frac{dl}{B} \{ \} = 0 . \quad (6.7.8)$$

To impose the first condition, we note that $d^3R = \sqrt{g_f} d\varrho d\theta_f d\zeta_f$. The volume integral must be taken over the infinitesimal volume dV between two flux surfaces. Hence, we obtain for the first condition:

$$d\varrho \int_0^{2\pi} \int_0^{2\pi} \sqrt{g_f} d\theta_f d\zeta_f \left\{ \frac{\dot{I}_{\text{pol}}^r \dot{\Psi}_{\text{tor}}^r - \dot{I}_{\text{tor}}^r \dot{\Psi}_{\text{pol}}^r}{4\pi^2\sqrt{g_f}} - \dot{p} \right\} = 0 , \quad (6.7.9)$$

or, since all the quantities involved are flux-surface quantities,

$$\dot{I}_{\text{pol}}^r \dot{\Psi}_{\text{tor}}^r - \dot{I}_{\text{tor}}^r \dot{\Psi}_{\text{pol}}^r = \dot{p} \int_0^{2\pi} \int_0^{2\pi} \sqrt{g_f} d\theta_f d\zeta_f . \quad (6.7.10)$$

The double integral is proportional to the volume of the thin shell we integrated over

$$dV = d\varrho \int_0^{2\pi} \int_0^{2\pi} \sqrt{g_f} d\theta_f d\zeta_f . \quad (6.7.11)$$

We thus have for this solvability condition:

$$\dot{I}_{\text{pol}}^r \dot{\Psi}_{\text{tor}}^r - \dot{I}_{\text{tor}}^r \dot{\Psi}_{\text{pol}}^r = \dot{p} \frac{dV}{d\varrho} . \quad (6.7.12)$$

This is a frequently used form of the ideal MHD equilibrium condition in toroidally confined plasmas.

The second solvability condition must be taken over a closed field line. From the equation of a field line, we have that $dl/B = d\zeta_f / B^{\zeta_f}$ which with $B^{\zeta_f} = \dot{\Psi}_{\text{tor}}^r / (2\pi\sqrt{g_f})$ equals $2\pi d\zeta_f \sqrt{g_f} / \dot{\Psi}_{\text{tor}}^r$. If we integrate over a field line that closes upon itself after N toroidal transits, we can write

$$\dot{p} \oint \frac{dl}{B} = \frac{\dot{I}_{\text{pol}}^r \dot{\Psi}_{\text{tor}}^r - \dot{I}_{\text{tor}}^r \dot{\Psi}_{\text{pol}}^r}{2\pi \dot{\Psi}_{\text{tor}}^r} \int_0^{2\pi N} d\zeta = \frac{(\dot{I}_{\text{pol}}^r \dot{\Psi}_{\text{tor}}^r - \dot{I}_{\text{tor}}^r \dot{\Psi}_{\text{pol}}^r)N}{\dot{\Psi}_{\text{tor}}^r} .$$

If we now make use of the first constraint in (6.7.12), we obtain

$$\dot{p} \oint \frac{dl}{B} = N \frac{dV}{d\varrho} \dot{\Psi}_{\text{tor}}$$

or with $\dot{p} \equiv dp/d\varrho$ and $\dot{\Psi}_{\text{tor}} \equiv d\Psi_{\text{tor}}/d\varrho$

$$\frac{dp}{d\varrho} \left[\frac{dV}{d\Psi_{\text{tor}}} - \frac{1}{N} \oint \frac{dl}{B} \right] = 0. \quad (6.7.13)$$

For $\dot{p} \equiv dp/d\varrho \neq 0$ this condition states that the loop integral of dl/B must be constant on rational surfaces:

$$\frac{1}{N} \oint \frac{dl}{B} = \frac{dV}{d\Psi_{\text{tor}}}. \quad (6.7.14)$$

The loop integral on the left-hand side is called the “proper length of a field line”. This condition requires that the proper lengths of all the field lines on a rational surface must be the same. This condition is often referred to as the *current closure condition*. If it is satisfied, the device is said to possess true flux surfaces. In axisymmetric devices, this condition is satisfied. To see this, consider a particular field line at a point A as shown in Fig. 6.8 and compute its $(1/N) \oint dl/B$. Do the same for a different field line in the same poloidal plane starting at a point B . Now move a toroidal distance $\Delta\zeta_f = q\Delta\theta_f$ backwards, where q is the safety factor. The field line arriving at point C is equivalent to that arriving at point B because of axisymmetry. However, the former field line goes to point A because it has a rotational transform $t = 1/q$. Hence, the proper length of both field lines must be the same. In nonaxisymmetric systems such as stellarators, the argument does not hold, and condition (6.7.14) is *not* satisfied. If this condition is violated, the Fourier coefficients of $\tilde{\eta}$ will exhibit resonant denominators of the form $(mt - n)$ leading to “infinite” Pfirsch-Schlüter currents and magnetic islands having fiber-like structures. We will discuss in Chap. 12 in what sense the two quantities on either side of (6.7.14) differ from each other in general.

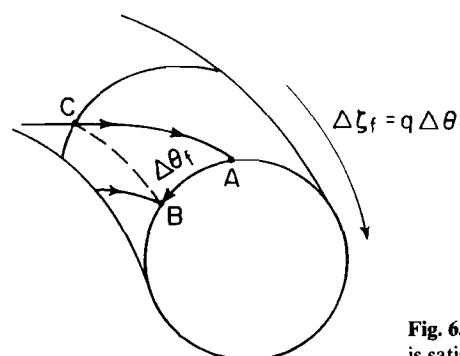


Fig. 6.8. Illustrating that the current-closure condition is satisfied in an axisymmetric tokamak. See text

It is customary to put the magnetic differential equation for $\tilde{\eta}$ in a simpler form by substituting the equilibrium condition given in (6.7.12) into (6.7.6). We then have

$$\mathbf{B} \cdot \nabla \tilde{\eta} = \dot{p} \left[\frac{1}{4\pi^2 \sqrt{g_f}} \frac{dV}{d\varrho} - 1 \right] \quad (6.7.15)$$

or with $dV/d\varrho \equiv \dot{V}$

$$\mathbf{B} \cdot \nabla \tilde{\eta} = \dot{p} \left[\frac{\dot{V}}{4\pi^2 \sqrt{g_f}} - 1 \right]. \quad (6.7.16)$$

In (θ_f, ζ_f) coordinates, the LHS can be found from $\mathbf{B}^i (\nabla \tilde{\eta})_i$ with the B^i given in (6.1.25). Equation (6.7.16) then becomes

$$\dot{\Psi}_{\text{pol}} \frac{\partial \tilde{\eta}}{\partial \theta_f} + \dot{\Psi}_{\text{tor}} \frac{\partial \tilde{\eta}}{\partial \zeta_f} = \frac{\dot{p}}{2\pi} (\dot{V} - 4\pi^2 \sqrt{g_f}). \quad (6.7.17)$$

From our expression for the flux-surface average given in (4.9.11) and using (6.7.11), we can write the flux-surface average of $1/\sqrt{g_f}$

$$\langle (\sqrt{g_f})^{-1} \rangle = \frac{\int_0^{2\pi} \int_0^{2\pi} d\theta_f d\zeta_f}{\int_0^{2\pi} \int_0^{2\pi} d\theta_f d\zeta_f \sqrt{g_f}} = \frac{4\pi^2}{dV/d\varrho} = \frac{4\pi^2}{\dot{V}}. \quad (6.7.18)$$

(Note that the average of $\sqrt{g_f}$ that Shafranov (1968) uses is not the same as a flux-surface average. For Shafranov, $\langle \sqrt{g_f} \rangle = (1/4\pi^2) \int \int d\theta_f d\zeta_f \sqrt{g_f}$.) The specific flux volume can thus alternatively be written as

$$\dot{V} = \frac{4\pi^2}{\langle (\sqrt{g_f})^{-1} \rangle} = 4\pi^2 \langle (\sqrt{g_f})^{-1} \rangle^{-1} \quad (6.7.19)$$

and consequently, the magnetic differential equation for $\tilde{\eta}$ of (6.7.16) reads

$$\mathbf{B} \cdot \nabla \tilde{\eta} = \dot{p} \left[\frac{\langle (\sqrt{g_f})^{-1} \rangle^{-1}}{\sqrt{g_f}} - 1 \right]. \quad (6.7.20)$$

It is interesting to consider the alternative path that Boozer (1981) uses to arrive at the same differential equation for $\tilde{\eta}$. Starting from his Clebsch form for the current $\mu_0 \mathbf{J} = \nabla \lambda \times \nabla \varrho$ and expanding $\nabla \lambda$ in its covariant components as $\nabla \lambda = (\partial \lambda / \partial \varrho) \nabla \varrho + (\partial \lambda / \partial \theta_f) \nabla \theta_f + (\partial \lambda / \partial \zeta_f) \nabla \zeta_f$, he writes an expression for the current density

$$\mathbf{J} = \frac{1}{\mu_0} \left[(\nabla \theta_f \times \nabla \varrho) \frac{\partial \lambda}{\partial \theta_f} + (\nabla \zeta_f \times \nabla \varrho) \frac{\partial \lambda}{\partial \zeta_f} \right]. \quad (6.7.21)$$

Substitution of this form for \mathbf{J} in the MHD equilibrium equation $\nabla p = \mathbf{J} \times \mathbf{B}$ results in a differential equation for λ as will be shown presently. We use the contravariant form for \mathbf{B}

$$\mathbf{B} = \frac{\dot{\Psi}_{\text{tor}}}{2\pi} (\nabla\varrho \times \nabla\theta_f) - \frac{\dot{\Psi}_{\text{pol}}}{2\pi} (\nabla\varrho \times \nabla\zeta_f) \quad (6.7.22)$$

and obtain for the MHD force balance

$$\begin{aligned} \mu_0 \dot{p} \nabla\varrho &= \left[-\frac{\partial\lambda}{\partial\theta_f} (\nabla\theta_f \times \nabla\varrho) \times (\nabla\varrho \times \nabla\zeta_f) \frac{\dot{\Psi}_{\text{pol}}}{2\pi} \right. \\ &\quad \left. + \frac{\partial\lambda}{\partial\zeta_f} (\nabla\zeta_f \times \nabla\varrho) \times (\nabla\varrho \times \nabla\theta_f) \frac{\dot{\Psi}_{\text{tor}}}{2\pi} \right]. \end{aligned} \quad (6.7.23)$$

Upon application of the *bac-cab* rule, this becomes

$$\mu_0 \dot{p} \nabla\varrho = \left(\frac{\dot{\Psi}_{\text{pol}}}{2\pi} \frac{\partial\lambda}{\partial\theta_f} + \frac{\dot{\Psi}_{\text{tor}}}{2\pi} \frac{\partial\lambda}{\partial\zeta_f} \right) [-\nabla\varrho (\nabla\zeta_f \cdot \nabla\theta_f \times \nabla\varrho)]. \quad (6.7.24)$$

With the definition of the Jacobian $(\sqrt{g_f})^{-1} = \nabla\varrho \cdot (\nabla\theta_f \times \nabla\zeta_f) = -\nabla\zeta_f \cdot (\nabla\theta_f \times \nabla\varrho)$, we obtain

$$\dot{p} = \frac{1}{2\pi\mu_0\sqrt{g_f}} \left(\dot{\Psi}_{\text{pol}} \frac{\partial\lambda}{\partial\theta_f} + \dot{\Psi}_{\text{tor}} \frac{\partial\lambda}{\partial\zeta_f} \right). \quad (6.7.25)$$

With λ given in (6.5.36) and (6.5.37),

$$\lambda = -\mu_0\eta = -\mu_0\tilde{\eta} - \frac{\mu_0}{2\pi} \hat{I}_{\text{tor}}\theta_f + \frac{\mu_0}{2\pi} \hat{I}_{\text{pol}}\zeta_f,$$

the above differential equation for λ reduces to (6.7.6).

To solve for the unknown function $\tilde{\eta}$, we must solve its magnetic differential equation, following the usual procedure (see Sect. 4.10). The functions $\tilde{\eta}$ and $\sqrt{g_f}$ appearing in (6.7.17) are expanded in a double Fourier series in θ_f and ζ_f and use is made of the equation of a field line $\theta_{f0} = \theta_f - t\zeta_f$ to eliminate one of the variables. The two solubility conditions put constraints on the Fourier expansion coefficients and allow proper treatment of resonant denominators of the form $(m_t - n)$ on rational surfaces.

6.8 Hamada Coordinates

In Sect. 6.4., we discussed the condition for straight current lines. The current is represented by $\mathbf{J} = \nabla\varrho \times \nabla\eta$, where $\eta = (1/2\pi)\hat{I}_{\text{tor}}\theta - (1/2\pi)\hat{I}_{\text{pol}}\zeta + \tilde{\eta}$, which implies that they are straight only when $\tilde{\eta}$ is a flux function or simply zero.

We wish to use the freedom of (6.1.28) to find a new set of flux coordinates in which the \mathbf{J} lines are straight in addition to the \mathbf{B} lines being straight. This will result in a coordinate system with a Jacobian equal to a constant (which can be made unity if desired). These new coordinates are known as *Hamada coordinates* or *natural coordinates* and are represented by θ_H and ζ_H .

The existence of Hamada coordinates was first proved by Hamada (1962). An independent proof was later established by Greene and Johnson (1962). Most treatments (and ours) follow this latter reference. See also Solov'ev and Shafranov (1970); and Boozer (1981).

The requirement that $\tilde{\eta}$ be a flux function or zero implies that $\mathbf{B} \cdot \nabla\tilde{\eta}$ must vanish. From (6.7.20), we observe that this is the case (for nonzero \dot{p}) if the inverse of the flux-surface average of the inverse Jacobian equals the Jacobian itself. In other words, the *Jacobian* must be a flux function. (There is a confusing misprint in Solov'ev and Shafranov, above (2.87). D is the total Jacobian from Cartesian coordinates (x, y, z) to Hamada coordinates $(\varrho, \theta_H, \zeta_H)$. It is not the Jacobian from θ, ζ to θ' (our θ_H) and ζ' (our ζ_H) as stated above (2.87).)

Equation (6.7.19) suggests that the choice of flux label,

$$\varrho \equiv V, \quad (6.8.1)$$

provides a very simple expression for the flux-surface average of the inverse Jacobian. Since then $\dot{V} = 1$, we have

$$\langle(\sqrt{g_f})^{-1}\rangle^{-1} = \frac{1}{4\pi^2}. \quad (6.8.2)$$

Hence, to make \mathbf{B} and \mathbf{J} lines straight, we must take the inverse Jacobian equal to $4\pi^2$ when $\varrho \equiv V$. Thus

$$(\sqrt{g})_H \equiv \frac{1}{4\pi^2}, \quad \varrho_H \equiv V. \quad (6.8.3)$$

$\varrho = V$ is not necessary for the \mathbf{J} lines to be straight, but it is very convenient because the ubiquitous \dot{V} equals one. If $\varrho \neq V$, then $\sqrt{g_H} = f(\varrho)$ is needed. (As before, we denote the Jacobian with V as flux label by $(\sqrt{g})_H$; the symbol \sqrt{g}_H is the “Hamada Jacobian” for the flux label ϱ .)

Since we began from the magnetic differential equation for $\tilde{\eta}$, in (6.7.20), we may conclude that a Hamada coordinate system exists when the solvability conditions of the $\mathbf{B} \cdot \nabla\tilde{\eta}$ equation are satisfied. The first implies $\langle \text{RHS} \rangle \equiv 0$ as stated in (4.10.5). This is equivalent to the volume integral condition; cf. (4.10.7). It is clear that with $\sqrt{g_f} \equiv \sqrt{g_H}$ equal to a flux function, the flux-surface average of the RHS of (6.7.20) is identically zero. Thus this condition reduces to an identity. This is not surprising because (6.7.20) is a rewritten form of (6.7.15), which was, in turn, obtained by subtracting the solvability constraint (6.7.12) from the original $\mathbf{B} \cdot \nabla\tilde{\eta}$ equation (6.7.6).

The other condition is the “current closure condition”

$$\frac{1}{N} \oint \frac{dl}{B} = \frac{dV}{d\Psi_{\text{tor}}}. \quad (6.8.4)$$

Thus, Hamada coordinates exist when the proper equilibrium conditions are satisfied.

In terms of Hamada coordinates θ_H and ζ_H , the contravariant components of \mathbf{B} are flux functions themselves. With $\sqrt{g_f} \rightarrow (\sqrt{g})_H = 1/4\pi^2$, we find from (6.1.25):

$$\mathbf{B}^i = 2\pi \left(0, \frac{d\Psi_{\text{pol}}^r}{dV}, \frac{d\Psi_{\text{tor}}}{dV} \right). \quad (6.8.5)$$

\mathbf{B} itself keeps the same Clebsch form as before, since the Jacobian did not appear in the expression for v in (6.1.23):

$$\mathbf{B} = \nabla V \times \nabla \left(\frac{1}{2\pi} \frac{d\Psi_{\text{tor}}}{dV} \theta_H - \frac{1}{2\pi} \frac{d\Psi_{\text{pol}}^r}{dV} \zeta_H \right). \quad (6.8.6)$$

The contravariant current-density components follow from (6.4.13)

$$\mathbf{J}^i = 2\pi \left(0, \frac{dI_{\text{pol}}^r}{dV}, \frac{dI_{\text{tor}}}{dV} \right), \quad (6.8.7)$$

while its Clebsch form, using (6.4.5) and (6.4.7) with $\tilde{\eta} = 0$, or (6.4.12), reads

$$\mathbf{J} = \nabla V \times \nabla \left(\frac{1}{2\pi} \frac{dI_{\text{tor}}}{dV} \theta_H - \frac{1}{2\pi} \frac{dI_{\text{pol}}^r}{dV} \zeta_H \right). \quad (6.8.8)$$

All other formulae and expressions discussed so far can be transformed to the Hamada set (V, θ_H, ζ_H) without any difficulties as long as we recall that $(\sqrt{g})_H = 1/4\pi^2$.

The expression for the Jacobian given in (6.6.4) allows us to compare the Hamada system with Boozer's system. With the Hamada choice of Jacobian, as given in (6.8.3), we have

$$(B)^2 - \mathbf{B} \cdot \nabla \tilde{\Phi} = \mu_0 (\dot{\Psi}_{\text{pol}}^r I_{\text{tor}} + \dot{\Psi}_{\text{tor}} I_{\text{pol}}^d) \quad (6.8.9a)$$

or

$$\frac{\partial \tilde{\Phi}}{\partial l} = (B)^2 - \mu_0 (\dot{\Psi}_{\text{pol}}^r I_{\text{tor}} + \dot{\Psi}_{\text{tor}} I_{\text{pol}}^d), \quad (6.8.9b)$$

with the dot representing here d/dV .

This equation says that θ_H and ζ_H are chosen such that the variation of $\tilde{\Phi}$ along a field line satisfies (6.8.9). Recall that Boozer chooses his θ_B and ζ_B such that $\tilde{\Phi} \equiv 0$ in his system. His Jacobian goes like $(B)^{-2}$ as shown by (6.6.5).

To actually find the new coordinates θ_H and ζ_H , we use (6.1.28)

$$\theta_H = \theta_f + \dot{\Psi}_{\text{pol}}^r G_H(\varrho, \theta_f, \zeta_f) \quad (6.8.10a)$$

$$\zeta_H = \zeta_f + \dot{\Psi}_{\text{tor}} G_H(\varrho, \theta_f, \zeta_f). \quad (6.8.10b)$$

We must find the function G_H to ensure that the Jacobian $(\sqrt{g_H})^{-1} = \nabla \varrho \cdot (\nabla \theta_H \times \nabla \zeta_H)$ is a flux function. The procedure is analogous to the derivation of (6.6.13) for Boozer coordinates. We substitute the expressions for θ_H and ζ_H in the Hamada Jacobian $\sqrt{g_H}$ and find a magnetic differential equation for $G_H(\varrho, \theta_f, \zeta_f)$

$$2\pi \mathbf{B} \cdot \nabla G_H = \frac{1}{\sqrt{g_H}} - \frac{1}{\sqrt{g_f}}. \quad (6.8.11)$$

(The B^i used to arrive at (6.8.11) are those in terms of the old system (θ_f, ζ_f) , i.e., given in (6.1.25).) As we discussed above (see also Solov'ev and Shafranov (1970)), this magnetic differential equation is subject to the usual solvability conditions. To actually solve for $G_H(\varrho \equiv V, \theta_f, \zeta_f)$, we write the $\mathbf{B} \cdot \nabla$ operator in terms of the old coordinates (θ_f, ζ_f) but with $\varrho \equiv V$. This last choice is not part of the transformation, it is merely to make things easy:

$$\frac{1}{\sqrt{g_f}} \frac{d\Psi_{\text{pol}}^r}{dV} \frac{\partial G_H}{\partial \theta_f} + \frac{1}{\sqrt{g_f}} \frac{d\Psi_{\text{tor}}}{dV} \frac{\partial G_H}{\partial \zeta_f} = \frac{1}{(\sqrt{g})_H} - \frac{1}{\sqrt{g_f}}$$

or

$$\left\{ t \frac{\partial G_H}{\partial \theta_f} + \frac{\partial G_H}{\partial \zeta_f} \right\} = (4\pi^2 \sqrt{g_f} - 1) \frac{dV}{d\Psi_{\text{tor}}}. \quad (6.8.12)$$

Then expanding the expressions for G_H and $\sqrt{g_f}$ in double Fourier series, and taking measures to deal with the resonant denominators of the form $(m_f - n)$ allows us to find the Fourier coefficients for G_H . (See Sect. 4.10 on Magnetic Differential Equations, or the review by Solov'ev and Shafranov (1970) for details.) Once G_H is known, the new coordinates θ_H and ζ_H have been found.

Knowledge of the angle functions θ_H and ζ_H allows determination of the basis vectors and thus the coefficients of the metric tensor g_{ij} . Solov'ev (1968, 1975) has indicated an alternative route for obtaining the basis vectors of a Hamada system. For simplicity, we shall use the indexed notation:

$$(u^1, u^2, u^3) \equiv (V, \theta_H, \zeta_H). \quad (6.8.13)$$

(Solov'ev (1968, 1975) uses index 3 to denote the radial variable V . Also, his Jacobian equals unity, whereas ours is $1/4\pi^2$; we have a factor of 2π in (6.8.5) and (6.8.7), whereas he divides through by 2π .)

The magnetic field \mathbf{B} and the current density \mathbf{J} are to be written in terms of Hamada coordinates as

$$\mathbf{B} = B^2 \mathbf{e}_2 + B^3 \mathbf{e}_3 = 2\pi \dot{\Psi}_{\text{pol}}^r \mathbf{e}_2 + 2\pi \dot{\Psi}_{\text{tor}} \mathbf{e}_3 \quad (6.8.14)$$

$$\mathbf{J} = J^2 \mathbf{e}_2 + J^3 \mathbf{e}_3 = 2\pi I_{\text{pol}}^r \mathbf{e}_2 + 2\pi I_{\text{tor}} \mathbf{e}_3. \quad (6.8.15)$$

Again, the dot here stands for d/dV .

Recall that $B^1 = J^1 = 0$. The systems of equations, (6.8.14) and (6.8.15) can be solved for the unknowns e_2 and e_3 in terms of the known vector fields \mathbf{B} and \mathbf{J} . With the help of the equilibrium condition, (6.7.12), where ϱ is identified with V , $\dot{I}_{\text{pol}}^r \dot{\Psi}_{\text{tor}} - \dot{I}_{\text{tor}} \dot{\Psi}_{\text{pol}}^r = \dot{p}$, we find:

$$e_2 = \frac{1}{2\pi\dot{p}} (\dot{\Psi}_{\text{tor}} \mathbf{J} - \dot{I}_{\text{tor}} \mathbf{B}) \quad (6.8.16)$$

and

$$e_3 = \frac{1}{2\pi\dot{p}} (\dot{I}_{\text{pol}}^r \mathbf{B} - \dot{\Psi}_{\text{pol}}^r \mathbf{J}) . \quad (6.8.17)$$

To find the first covariant-basis vector e_1 , we expand $e^1 \equiv \nabla V$ along the covariant set e_i as follows:

$$\begin{aligned} \nabla V &= (\nabla V \cdot e^1) e_1 + (\nabla V \cdot e^2) e_2 + (\nabla V \cdot e^3) e_3 \\ &= |\nabla V|^2 e_1 + g^{12} e_2 + g^{13} e_3 . \end{aligned} \quad (6.8.18)$$

Here we left g^{11} as $|\nabla V|^2$. After solving this equation for e_1 , we obtain

$$e_1 \equiv e_V = \frac{\nabla V}{|\nabla V|^2} - \frac{g^{12}}{|\nabla V|^2} e_2 - \frac{g^{13}}{|\nabla V|^2} e_3 . \quad (6.8.19)$$

For clarity, we define two new symbols for the unknown coefficients of e_2 and e_3 in (6.8.19):

$$\begin{aligned} \omega &= -\frac{g^{12}}{|\nabla V|^2} \quad \text{and} \quad \mu = -\frac{g^{13}}{|\nabla V|^2} , \\ e_1 &= \frac{\nabla V}{|\nabla V|^2} + \omega e_2 + \mu e_3 . \end{aligned} \quad (6.8.20)$$

After some straightforward but somewhat tedious vector manipulations (the main points of which are given below), one obtains two magnetic differential equations for ω and μ , respectively. They are a consequence of $\mu_0 \mathbf{J} = \nabla \times \mathbf{B}$ and read:

$$\mathbf{B} \cdot \nabla \omega = -2\pi \dot{\Psi}_{\text{pol}}^r - \frac{4\pi^2}{|\nabla V|^2} e_3 \cdot \left[\mu_0 \mathbf{J} - \frac{2}{|\nabla V|^2} (\nabla V \times (\nabla V \cdot \nabla) \mathbf{B}) \right] \quad (6.8.21)$$

$$\mathbf{B} \cdot \nabla \mu = -2\pi \dot{\Psi}_{\text{tor}} + \frac{4\pi^2}{|\nabla V|^2} e_2 \cdot \left[\mu_0 \mathbf{J} - \frac{2}{|\nabla V|^2} (\nabla V \times (\nabla V \cdot \nabla) \mathbf{B}) \right] . \quad (6.8.22)$$

When these equations have been solved, ω , μ and e_1 , and thus all three covariant basis vectors are known in terms of the known magnetic field and current density.

Proof of (6.8.21):¹

We now indicate how the magnetic differential equations can be obtained. We focus on the first one, (6.8.21).

- 1) From the vector identity $\nabla(\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla) \mathbf{B} + (\mathbf{B} \cdot \nabla) \mathbf{A}$, with $\mathbf{A} \equiv \nabla V$, and $\mathbf{B} \equiv \mathbf{B}$, $\mathbf{B} \cdot \nabla V = 0$ and $\mu_0 \mathbf{J} = \nabla \times \mathbf{B}$, we find that

$$\mu_0 \mathbf{J} \times \nabla V = (\nabla V \cdot \nabla) \mathbf{B} + (\mathbf{B} \cdot \nabla) \nabla V . \quad (6.8.23)$$

- 2) The perpendicular component of a vector \mathbf{A} with respect to a directional vector $\hat{\mathbf{a}}$ is $-\hat{\mathbf{a}} \times (\hat{\mathbf{a}} \times \mathbf{A})$. Since $\mathbf{J} \perp \nabla V$, it follows that $\nabla V \times$ (6.8.23) equals

$$\mu_0 |\nabla V|^2 \mathbf{J} = \nabla V \times (\nabla V \cdot \nabla) \mathbf{B} + \nabla V \times (\mathbf{B} \cdot \nabla) \nabla V . \quad (6.8.24)$$

To obtain terms similar to those of (6.8.21), we dot multiply (6.8.24) by e_3 :

$$\mu_0 |\nabla V|^2 e_3 \cdot \mathbf{J} = e_3 \cdot [\nabla V \times (\nabla V \cdot \nabla) \mathbf{B}] + e_3 \cdot [\nabla V \times (\mathbf{B} \cdot \nabla) \nabla V] \quad (6.8.25)$$

- 3) We concentrate on the last term of (6.8.25). Write the first ∇V of this term as e^1 and replace the second ∇V by $|\nabla V|^2 (e_1 - \omega e_2 - \mu e_3)$ according to (6.8.18) or (6.8.20). The term proportional to $\mathbf{B} \cdot \nabla |\nabla V|^2$ vanishes because $e^1 \times e^1 / |\nabla V|^2 \equiv 0$. Next, we expand $\mathbf{B} \cdot \nabla$ as $B^i \partial/\partial u^i$. The terms involving $\partial e_k / \partial u^i$ can be expanded in contravariant components as $\partial e_k / \partial u^i = e_j (e^j \cdot \partial e_k / \partial u^i)$. Then writing e_j in terms of the e^a , via the “back-and-forth” relations of (2.5.25) with a Jacobian equal to $1/4\pi^2$, and application of the *bac-cab* rule leads to $e_3 \cdot e^1 \times e_j = (-g^{12}/4\pi^2, |\nabla V|^2/4\pi^2, 0)$ for $j = (1, 2, 3)$, respectively. These rules imply that the term $-|\nabla V|^2 (e_3 \cdot e^1 \times e_3) \mathbf{B} \cdot \nabla \mu$ vanishes, while $-|\nabla V|^2 (e_3 \cdot e^1 \times e_2) \mathbf{B} \cdot \nabla \omega = -|\nabla V|^4 \mathbf{B} \cdot \nabla \omega / 4\pi^2$ survives. After using the property of the Christoffel symbol, given in (2.6.22), $e^j \cdot \partial e_k / \partial u^i = -e_k \cdot \partial e^j / \partial u^i$, and (6.8.20), we have

$$e_3 \cdot [\nabla V \times (\mathbf{B} \cdot \nabla) \nabla V] = \frac{-|\nabla V|^4}{4\pi^2} \mathbf{B} \cdot \nabla \omega - B^i \frac{|\nabla V|^2}{4\pi^2} \nabla V \cdot \frac{\partial e^2}{\partial u^i} + B^i \left(\frac{g^{12}}{4\pi^2} \right) \nabla V \cdot \frac{\partial e^1}{\partial u^i} . \quad (6.8.26)$$

- 4) The last two terms of (6.8.26) can be manipulated into a term proportional to $e_3 \cdot \nabla V \times (\nabla V \cdot \nabla) \mathbf{B}$, leading to the factor 2 in (6.8.21) and a term proportional to $\dot{\Psi}_{\text{pol}}^r$. The easiest way to prove this is to show that $e_3 \cdot \nabla V \times (\nabla V \cdot \nabla) \mathbf{B}$ leads to something similar to (6.8.26). Because $(\nabla V \cdot \nabla) \equiv g^{1k} \partial/\partial u^k$ and since $B^1 \equiv 0$, we obtain by using the above mentioned expressions for $e_3 \cdot e^1 \times e_j$:

$$e_3 \cdot [\nabla V \times (\nabla V \cdot \nabla) \mathbf{B}] = \frac{|\nabla V|^4}{2\pi} \dot{\Psi}_{\text{pol}}^r - \frac{g^{12}}{4\pi^2} g^{1k} B^i \left(e^1 \cdot \frac{\partial e_i}{\partial u^k} \right) + \frac{|\nabla V|^2}{4\pi^2} g^{1k} B^i \left(e^2 \cdot \frac{\partial e_i}{\partial u^k} \right) . \quad (6.8.27)$$

We now replace $e^j \cdot \partial e_i / \partial u^k = e^j \cdot \partial e_k / \partial u^i$ by $-e_k \cdot \partial e^j / \partial u^i$ (for $j = 1, 2$ in the last two terms) and write g^{1k} as $e^1 \cdot e^k$. After recognizing that we are then left with the covariant expansion of $\partial e^j / \partial u^i$ as $e^k e_k \cdot \partial e^j / \partial u^i$, we obtain for (6.8.27):

$$e_3 \cdot [\nabla V \times (\nabla V \cdot \nabla) \mathbf{B}] = \frac{|\nabla V|^4}{2\pi} \dot{\Psi}_{\text{pol}}^r + \frac{g^{12}}{4\pi^2} B^i \left(e^1 \cdot \frac{\partial e^1}{\partial u^i} \right) - \frac{|\nabla V|^2}{4\pi^2} B^i \left(e^1 \cdot \frac{\partial e^2}{\partial u^i} \right) . \quad (6.8.28)$$

- 5) Comparing (6.8.28) with (6.8.26) gives

$$e_3 \cdot [\nabla V \times (\mathbf{B} \cdot \nabla) \nabla V] = -\frac{|\nabla V|^4}{4\pi^2} \mathbf{B} \cdot \nabla \omega + e_3 \cdot [\nabla V \times (\nabla V \cdot \nabla) \mathbf{B}] - \frac{|\nabla V|^4}{2\pi} \dot{\Psi}_{\text{pol}}^r . \quad (6.8.29)$$

- 6) After substitution of (6.8.29) into (6.8.25), we recover (6.8.21).

¹ Can be skipped without loss of continuity.

7. Conversion from Clebsch Coordinates to Toroidal Flux Coordinates

In this Chapter, we discuss the conversion formulae from toroidal flux coordinates to Clebsch functions. First, we deal with the (ϱ, v, l) system, and after that, we discuss the Boozer-Grad (ϱ, v, χ) system.

7.1 The Generic Clebsch Coordinate System (ϱ, v, l)

It is quite inconvenient to transform from angle flux coordinates to the Clebsch function v and the “along the field line” coordinate l . The major problem has to do with the fact that the l coordinate curve does not close upon itself after one toroidal transit. For a qualitative discussion of the problems involved, see Sect. 4.6.1d.

In angle flux coordinates, the magnetic field is written as

$$\mathbf{B} = \nabla \varrho \times \nabla v = \frac{\dot{\Psi}_{\text{tor}}}{2\pi} \theta_f - \frac{\dot{\Psi}_{\text{pol}}^r}{2\pi} \zeta_f . \quad (7.1.1)$$

The coordinates are $(\varrho, \theta_f, \zeta_f)$. In the Clebsch (ϱ, v, l) representation, it is

$$\mathbf{B} = \nabla \varrho \times \nabla v = B \hat{l} . \quad (7.1.2)$$

It is not difficult to see that

$$\varrho \equiv \varrho \quad (7.1.3)$$

$$v \equiv \frac{1}{2\pi} (\dot{\Psi}_{\text{tor}} \theta_f - \dot{\Psi}_{\text{pol}}^r \zeta_f) . \quad (7.1.4)$$

Finding l as a function of θ_f and ζ_f is not easy. The field lines are straight in flux coordinates, but the θ_f and ζ_f coordinate curves do not intersect at right angles. We must use the metric coefficients g_{ij} to relate l to θ_f and ζ_f , which results in an integral relationship. One way to see this is to compute the square of the differential arc length $(dl)^2$:

$$(dl)^2 = g_{ij} du^i du^j \\ = g_{\theta_f \theta_f} (d\theta_f)^2 + 2g_{\theta_f \zeta_f} (d\theta_f)(d\zeta_f) + g_{\zeta_f \zeta_f} (d\zeta_f)^2 . \quad (7.1.5)$$

The equation of a field line, expressed in flux coordinates, $d\theta_f = t d\zeta_f$, simplifies this a little. We can express $(dl)^2$ in terms of only one angle ζ_f ,

$$(dl)^2 = (t^2 g_{\theta_f \theta_f} + 2t g_{\theta_f \zeta_f} + g_{\zeta_f \zeta_f}) (d\zeta_f)^2$$

or

$$dl = \sqrt{t^2 g_{\theta_f \theta_f} + 2t g_{\theta_f \zeta_f} + g_{\zeta_f \zeta_f}} d\zeta_f . \quad (7.1.6)$$

Since the expression under the square root is in general a function of ζ_f , analytic integration is not generally possible. The best we can do is to write an integral relationship

$$l = \int_0^{\zeta_f} \sqrt{t^2 g_{\theta_f \theta_f} + 2t g_{\theta_f \zeta_f} + g_{\zeta_f \zeta_f}} d\zeta_f . \quad (7.1.7)$$

An alternative expression for $l \equiv l(\theta_f, \zeta_f)$ can be obtained by using the general equation of a field line as follows:

$$\frac{dl}{B} = \frac{d\zeta_f}{B^{\zeta_f}} \quad \text{or} \quad dl = \frac{B}{B^{\zeta_f}} d\zeta_f . \quad (7.1.8)$$

In flux coordinates $B^{\zeta_f} = \dot{\Psi}_{\text{tor}} / (2\pi\sqrt{g_f})$ so that dl is given by

$$dl = \frac{2\pi B \sqrt{g_f}}{\dot{\Psi}_{\text{tor}}} d\zeta_f . \quad (7.1.9)$$

$B \equiv |\mathbf{B}|$ as well as $\sqrt{g_f}$ are generally functions of θ_f and ζ_f so again an integral formula is the best we can do:

$$l = \frac{2\pi}{\dot{\Psi}_{\text{tor}}} \int_0^{\zeta_f} B \sqrt{g_f} d\zeta_f . \quad (7.1.10)$$

A second use of the equation of a field line, $d\theta_f = t d\zeta_f$ or $\theta_{f0} = \theta_f - t\zeta_f$ can remove the θ_f dependence from B and $\sqrt{g_f}$. Even in Hamada coordinates, the problem is not any simpler. Although the Jacobian can be taken out of the integral, $B \equiv B(\theta_f, \zeta_f)$ must remain under it.

7.2 Boozer-Grad Coordinates (ϱ, v, χ)

The covariant and contravariant forms of \mathbf{B} in the (ϱ, v, χ) coordinate frame are, respectively,

$$\mathbf{B} = \nabla \chi + \lambda \nabla \varrho . \quad (7.2.1)$$

and

$$\mathbf{B} = \nabla \varrho \times \nabla v . \quad (7.2.2)$$

As discussed above, the covariant Boozer-Grad form reduces to a very simple form in vacuum: $\mathbf{B} = \nabla\chi$, where χ can *then* be thought of as a scalar magnetic potential. A second important feature of the “along the field line” label χ is the fact that it can very easily be expressed in terms of the flux coordinate angles θ_f and ζ_f . This is probably the major advantage of the variable χ over the length l .

The variables ϱ and v do not cause any difficulty. As in the previous section, they obey (7.1.3) and (7.1.4). Concerning the “along the field line” variable χ , we shall now argue that starting from the covariant form of \mathbf{B} , there is only one possible form that χ can have in terms of θ_f and ζ_f . (Here, we temporarily disregard the fact that we have already found χ in terms of ϱ , θ_f and ζ_f in (6.5.39), which, as one recalls, was obtained by finding the covariant components of \mathbf{B} in flux coordinates and identifying the covariant expression for \mathbf{B} with the Boozer-Grad form. What we are about to undertake now is a somewhat more direct approach.) The argument is closely analogous to that which we presented for the stream function v in (6.1.13).

The function χ in (7.2.1) need not be single valued (i.e., periodic) in θ_f and ζ_f . Even $\nabla\chi$ is exempted from the periodicity property, as long as the $\lambda\nabla\varrho$ term cancels off any non-single valuedness so that \mathbf{B} is single valued. However, the term $\lambda\nabla\varrho$ can obviously only affect the terms in $\nabla\chi$ proportional to $\nabla\varrho$.

As was the case with v , any non-periodic function which is not a linear polynomial in θ_f or ζ_f must be rejected.

The “constant” in front of the linear term can be a function of ϱ . Although that term does produce a secular term in θ_f or ζ_f , it has $\nabla\varrho$ in its gradient which can be cancelled by the $\lambda\nabla\varrho$ term: $\nabla(C(\varrho)\theta_f) = C\nabla\theta_f + \theta_f C\nabla\varrho$.

There is a possibility that a periodic function must be included, $\tilde{\Phi}(\varrho, \theta_f, \zeta_f)$. Thus χ must have the form

$$\chi = K(\varrho) + C(\varrho)\theta_f + D(\varrho)\zeta_f + \tilde{\Phi}(\varrho, \theta_f, \zeta_f). \quad (7.2.3)$$

If we absorb the $K(\varrho)$ in the function $\tilde{\Phi}$, we obtain

$$\chi = C(\varrho)\theta_f + D(\varrho)\zeta_f + \tilde{\Phi}(\varrho, \theta_f, \zeta_f). \quad (7.2.4)$$

With this χ , the covariant representation of \mathbf{B} of (7.2.1) is

$$\mathbf{B} = \left(C + \frac{\partial \tilde{\Phi}}{\partial \theta_f} \right) \nabla \theta_f + \left(D + \frac{\partial \tilde{\Phi}}{\partial \zeta_f} \right) \nabla \zeta_f + \left(\lambda + \dot{C}\theta_f + \dot{D}\zeta_f + \frac{\partial \tilde{\Phi}}{\partial \varrho} \right) \nabla \varrho. \quad (7.2.5)$$

To find the functions $C(\varrho)$ and $D(\varrho)$, we use Ampère’s law along a toroidal coordinate curve and along a poloidal coordinate curve,

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 I_{\text{enclosed}}. \quad (7.2.6)$$

This has been explained in (6.5.22) and (6.5.23). We find that

$$C(\varrho) \equiv \frac{\mu_0}{2\pi} I_{\text{tor}}(\varrho) \quad (7.2.7)$$

$$D(\varrho) \equiv \frac{\mu_0}{2\pi} (I_{\text{pol}}^{\text{plas}} - I_{\text{pol}}^r(\varrho)) + \frac{\mu_0}{2\pi} \sum (I_{\text{coils}}) = \frac{\mu_0}{2\pi} I_{\text{pol}}^d(\varrho). \quad (7.2.8)$$

The symbols are all explained in Sect. 6.6.5. As expected, χ has the form already given in (6.5.39).

If we choose θ_f and ζ_f such that $\tilde{\Phi}$ vanishes in that system, which angles we now call θ_B and ζ_B (for Boozer), we recover (6.6.14c). Apart from the arbitrariness in v and χ mentioned in (5.2.13) and (5.4.39), respectively, the functions v and χ are completely determined.

It is clear that the expression for χ given in (7.2.4) is much simpler than that for l in (7.1.7) or (7.1.10). Thus, it appears that the (ϱ, v, χ) system is preferable to (ϱ, v, l) . However, the flux coordinate angle system (θ_f, ζ_f) seems to have an advantage over both of these other systems because of the “ambiguity” in the constant v surfaces as explained in Sect. 4.6.1d.

8. Establishment of the Flux-Coordinate Transformation; A Summary

The material in this section on finding the transformation from a particular coordinate system to flux coordinates has to a large extent already been discussed above, although the details were scattered throughout several sections. Here we summarize briefly the methods used and provide some practical references.

First, it should be understood that use of flux coordinates does not require knowledge of the transformation equations. In fact, the transformation itself is rarely undertaken. The whole idea of flux coordinates is based on computational convenience: they give a helpful decoupling of certain physics problems one wishes to focus on (e.g., transport across magnetic surfaces or along magnetic field lines) from complicated geometrical features of the device under investigation. To gain a theoretical understanding of a variety of physics issues this is sufficient. However, when we wish to obtain numbers which can be related to experimental data, quantities expressed in flux coordinates must be transformed.

A crucial assumption in the methods of establishing flux coordinates is that *the magnetic field and the current density are known everywhere* in a finite pressure plasma. Knowledge of \mathbf{B} allows the construction of magnetic surfaces via analytic formulae in systems with symmetry and via approximate analytic calculations or via numerical field-line following in more general devices. The determination of the flux surfaces is part of the more general problem of finding the equilibrium state of the plasma. This will be discussed very briefly below. These flux surfaces then serve as our first coordinate surfaces. The fact that \mathbf{B} is known at every point on the flux surface permits a choice of other suitable coordinates which make the field lines appear as straight lines. J lines also lie in flux surfaces and similarly, we can make the current-density lines straight (in addition to having straight magnetic-field lines or independently of the shape of the magnetic-field lines).

8.1 Toroidal Systems

In toroidal systems, we consider the components of \mathbf{B} to be known as functions of the flux variable ϱ and two angular variables θ and ζ . For example, if \mathbf{B} is known in terms of the Cartesian coordinates, a first transformation to an “ideal” toroidal-cylindrical or “elementary” toroidal coordinate system (r, θ_e, ζ_e) is easily performed. As soon as the flux surfaces are known the radial variable r can be

replaced by a flux function $\varrho \equiv \varrho(\Psi)$. So far, we have $\mathbf{B} \equiv \mathbf{B}(\varrho, \theta_e, \zeta_e)$. If we keep the coordinate surfaces $\theta_e = \text{constant}$ and $\zeta_e = \text{constant}$ to start with, the contravariant components of \mathbf{B} , $\mathbf{B} \cdot \nabla \theta_e = B^{\theta_e}$ and $\mathbf{B} \cdot \nabla \zeta_e = B^{\zeta_e}$ are straightforwardly determined in terms of $(\varrho, \theta_e, \zeta_e)$.

In Sect. 6.1, we explained how from B^θ and B^ζ (in the above example, B^{θ_e} and B^{ζ_e}) the periodic part $\tilde{v}(\varrho, \theta, \zeta)$ of the streaming function $v(\varrho, \theta, \zeta)$ can be found (see underneath (6.1.13)). The original ζ ($\equiv \zeta_f$) coordinate, together with a deformed θ -coordinate, now called θ_f , and equal to $\theta + 2\pi\tilde{v}/\Psi_{\text{tor}}$, make the magnetic field lines straight (see (6.1.22) to (6.1.24)).

For Boozer and Hamada coordinates, we go one step further. Having established the flux coordinates θ_f and ζ_f , we deform them further to get rid of the periodic part of the “magnetic scalar potential” in Boozer’s case, or to make the J lines straight in the Hamada system, in addition to having straight \mathbf{B} lines. (The Hamada system also has the advantage of a unity or constant Jacobian.) In both cases, $G(\varrho, \theta_f, \zeta_f)$ must be found and the new “angles” are then determined by (6.1.28). Equation (6.6.13) shows the solution for $G_B \equiv G_{\text{Boozer}}$, whereas (6.8.11) gives the differential equation that should be solved for $G_H \equiv G_{\text{Hamada}}$.

If a Clebsch coordinate system is desired, the streaming functions are the flux surfaces labeled by ϱ and the function $v = \text{constant}$: $\mathbf{B} = \nabla \varrho \times \nabla v$. The third variable is either l or χ , where $\chi = \int B dl$ (see (5.4.26)). To find l from the available information, one could use the equation of a field line in the coordinate system in which \mathbf{B} was originally specified.

A flux-coordinate system is completely specified (in terms of the variables of another system) when the transformation equations or the metric coefficients g_{ij} have been found.

Specific examples of establishing a flux-coordinate system have been treated in the literature. In one approach, Shafranov (1968) derives the metric coefficients starting from a Frenet-Serret coordinate system based upon the magnetic axis. The actual computation is performed via an expansion procedure in terms of the curvature of the magnetic axis. These axial coordinate systems are discussed in considerable detail in the review papers by Solov’ev and Shafranov (1970) and by Solov’ev (1975). The expansion procedure in terms of the curvature is described by Shafranov (1966).

In his 1982 paper, Boozer describes how to construct a magnetic coordinate system via field-line following and Fourier expansion. To limit the extent of the field line integration, a Gaussian window function is used. This seems to be a widely used technique in electronics when dealing with the fast Fourier transform. (The fast Fourier transform technique and the desirability of window functions are discussed in Bergland (1969). A variety of window functions has been reviewed in the article by Harris (1978).) For further details on field line integration for flux coordinate purposes, see Kuo-Petravic, Boozer, Rome and Fowler (1983), Kuo-Petravic (1984) and Beasley, Rome, Attenberger and Hirshman (1987). An improved field-line-following method, which requires integration along two field lines, has been devised by Rome (1989). Reiman and Pomphrey

(1989) have developed an algorithm which is a combination of iterative and trajectory-following methods. In the next section, we shall say more about Boozer's "Fourier field-line-following" method.

We shall now comment on the computation of the flux surfaces. This calculation is not a trivial exercise, even when the magnetic field is known everywhere. An analytic result can be obtained for symmetric devices (such as tokamaks or helically symmetric stellarators) or in a non-axisymmetric device an approximate solution may be found by using the method of averaging or an expansion technique near a closed field line. These methods are explained by Morozov and Solov'ev (1966). In addition, the review paper by Solov'ev and Shafranov (1970) covers a solution procedure for flux surfaces close to the magnetic axis by means of a perturbation technique analogous to the stellarator expansion introduced by Greene and Johnson (1961). For more general cases (e.g., where a suitable expansion parameter does not exist), one has to resort to numerical methods. This merely amounts to integrating the equations of a field line.

Recently, Kuo-Petravic and Boozer (1986) described a numerically economical method in which the Hamiltonian character of magnetic-field lines is utilized. According to the authors, this is an effective way to evaluate the quality of the magnetic surfaces. We shall discuss a few aspects of the magnetic field-line Hamiltonian in the next Chapter as this is a rather important concept.

Frequently, the determination of the flux surfaces is undertaken as part of a self-consistent equilibrium calculation. In devices with a symmetry direction, a substantial amount of information is summarized in the solution of an equation known as the Grad-Schlüter-Shafranov equation (after its discoverers, in alphabetical order). This is a second-order differential equation for the flux variable $\Psi \propto \Psi_{\text{pol}}$ in terms of the flux functions $I \equiv I(\Psi) \equiv RB_T$ and the pressure $p \equiv p(\Psi)$. It is obtained by substituting $J = (\nabla \times B)/\mu_0$ into $J \times B = \nabla p = p(\Psi) \nabla \Psi$, after which B is replaced by $I\nabla\zeta + (\nabla\zeta \times \nabla\Psi)$. Because of axisymmetry, $I \equiv I(\Psi)$ is a flux function and one obtains

$$R^2 \nabla \cdot (R^{-2} \nabla \Psi) = -I \frac{dI}{d\Psi} - R^2 \mu_0 \frac{dp}{d\Psi}. \quad (8.1.1)$$

The operator on the left may be written as:

$$\Delta^* \Psi \equiv \frac{\partial^2 \Psi}{\partial R^2} - \frac{1}{R} \frac{\partial \Psi}{\partial R} + \frac{\partial^2 \Psi}{\partial z^2} = R \frac{\partial}{\partial R} \left(\frac{1}{R} \frac{\partial \Psi}{\partial R} \right) + \frac{\partial^2 \Psi}{\partial z^2}. \quad (8.1.2)$$

Here z is the vertical coordinate of the cylindrical system, (R, ζ_c, z) . Δ^* is reminiscent of the notation for the Laplacian $\Delta \equiv \nabla^2$. Δ^* is a "magnetic differential" operator similar to the non-angular part of the Laplacian in cylindrical coordinates:

$$\left(\Delta - \frac{1}{R^2} \frac{\partial^2}{\partial \zeta_c^2} \right)_{\text{cyl}} \equiv \left(\nabla^2 - \frac{1}{R^2} \frac{\partial^2}{\partial \zeta_c^2} \right)_{\text{cyl}} = \frac{1}{R} \frac{\partial}{\partial R} \left(R \frac{\partial}{\partial R} \right) + \frac{\partial^2}{\partial z^2}. \quad (8.1.3)$$

The Grad-Schlüter-Shafranov equation was derived independently by Shafranov (1958) in 1957, by Lüst and Schlüter (1957) and by Grad and Rubin (1958). The derivation of the equation can also be found in the review papers by Shafranov (1966), Freidberg (1982), Zakharov and Shafranov (1986) and Hazeltine and Meiss (1985). See also, Callen and Dory (1972). For simple textbook treatments, we refer to Bateman (1978), Stacey (1981) and Freidberg (1987).

Since p and I are not necessarily linear functions of Ψ and since the boundary conditions can be nonlinear as well, the Grad-Schlüter-Shafranov equation is in general a second-order nonlinear differential equation. Since one must specify $p(\Psi)$ and $I(\Psi)$ before one knows what the flux surfaces are, there is some ambiguity in modelling a plasma. In most calculations of equilibria for experimental devices, the Grad-Schlüter-Shafranov equation is solved by numerical iteration. See e.g. Lackner (1976). A valuable variational method is that by Lao, Hirshman and Wieland (1981). (See also Lao, 1984). Certain idealized cases seem to be tractable analytically, as has been discussed by Solov'ev (1975) and Zakharov and Shafranov (1986).

The existence of a Grad-Schlüter-Shafranov equation is crucially dependent on the presence of a symmetry direction because this makes I in the expression for B a flux function. As a consequence, there also exists a G-S-S equation for a helically symmetric stellarator. This equation is derived in Zakharov and Shafranov (1986); see also, Pytte and Boozer (1981) and White and Chance (1984).

In devices without a symmetry direction such as toroidal stellarators, there is no Grad-Schlüter-Shafranov equation, and the equilibrium problem is even more complicated. To treat this problem, Greene and Johnson (1961) developed the so-called stellarator expansion. This procedure solves the equation $B \cdot \nabla \Psi = 0$, where Ψ is expanded as $\sum_i \Psi^{(i)}$ such that higher order $\Psi^{(i)}$ are considered as perturbations with regard to the lowest order Ψ . To lowest order it models a uniform axial magnetic field. The first-order field is a superposition of multi-polar fields, and the second order consists of fields which take into account the effects of the curvature of the system and the currents in the plasma. This expansion procedure was suggested in an earlier paper by Johnson, Oberman, Kulsrud and Frieman (1958). The method has been further utilized and/or discussed by Anania, Johnson and Weimer (1983a), and by Anania and Johnson (1983b), by Johnson, et al., (1984) and by Johnson (1986). Dobrott and Frieman (1971) have computed the vacuum magnetic surfaces of an $l = 3$ stellarator using an "optimal stellarator expansion". This same expansion was used by Fielding and Hitchon (1980) to compute the equilibrium flux surfaces and fields and, in particular, the equilibrium value of the vertical field, in an $l = 3$ stellarator. Later, this was extended to arbitrary l -numbers (Hitchon and Fielding, 1981). Related procedures have been described by Solov'ev and Shafranov (1970) and Johnson (1986).

Besides analytic (mostly asymptotic) results which can provide physical insight, numerical equilibrium codes provide results which are asymptotic in the mesh size and attack more complicated cases. Reviews of numerical methods

used to compute the plasma equilibrium are given in the papers of Zakharov and Shafranov (1986), and Johnson (1986). More details can be found in Methods in Computational Physics, Killeen (1976). In addition, we should mention the 3-D equilibrium codes by Chodura and Schlüter (1981) and by Bauer, Betancourt and Garabedian (1978 and 1984). See also Hirshman and Whitson (1983), Lao (1984), Schwenn (1984), Bhattacharjee et al. (1984) and Hirshman and Lee (1986), among others. Island formation and destruction of flux surfaces in three-dimensional equilibria have been investigated by Boozer (1984a, b) and Reiman and Boozer (1984).

Further discussion of toroidal equilibrium can be found in the papers by Grad (1967), Kruskal and Kulsrud (1958), Spies and Nelson (1974), and Boozer (1981).

8.2 Open-Ended Systems

The flux coordinates of concern in open-ended systems are of the Clebsch or the Boozer-Grad type. To obtain the transformation equations, the same equations as in toroidal systems must be solved. There is one crucial difference concerning the flux label and the magnetic surfaces. As pointed out above, in mirrors there are no flux surfaces which are traced out by a single field line. One must *define* a closed contour somewhere in the plasma, and the flux surface is that surface traced out by all the magnetic field lines intersecting the chosen contour. Conventionally, for the vacuum field, the curve is chosen to be a circle (concentric about the axis) located in the middle of the device at $z = 0$, when the device is bounded by $z = \pm L$. (In finite- β plasmas, the originally circular flux surfaces tend to acquire a "square" shape.) The vacuum flux surfaces for tandem mirrors in the "long-thin cylinder" or "paraxial" approximation have been computed by Furth and Rosenbluth (1964) and by Cordey and Watson (1969). See also the review paper by Baldwin (1977). More recent papers dealing with mirror equilibrium include those by Hall and McNamara (1975), Pearlstein, Kaiser and Newcomb (1981), Newcomb (1981), and Bulmer, et al., (1983).

9. Canonical Coordinates or "Generalized Magnetic Coordinates"

9.1 Flux Coordinates Versus Canonical Coordinates

As an extension to the subject of flux coordinates, we discuss concisely the more general topic of canonical coordinates for magnetic fields. The word canonical derives from the Hamiltonian character of the magnetic field, which will be demonstrated presently. Canonical coordinates for \mathbf{B} , and the canonical representation of \mathbf{B} exist even if there are no flux surfaces. Flux coordinates, on the contrary, only exist in the presence of magnetic surfaces. The canonical coordinates are sometimes called "generalized magnetic coordinates". It must be understood, however, that this is *not* a synonym for flux coordinates. The reason for the name "generalized magnetic coordinates" is the superficial resemblance of the \mathbf{B} -field representation in canonical coordinates and flux coordinates.

9.2 On the Existence of Flux Surfaces, Revisited

In the section on flux surfaces, we pointed out that they only rigorously exist in devices with an ignorable coordinate. In such a case, the irrational field lines cover a two-dimensional *surface* ergodically. In more general systems, the magnetic-field lines do not all stay on a surface but in general ergodically fill a finite *volume*.

We are interested in devices (such as stellarators or real tokamaks) in which the field lines generate "surfaces" with very limited ergodic volumes. The magnetic field in such devices is considered as one that has the required symmetry, plus a perturbation field. One might expect that there would be no closed surfaces left at all due to coupling of the perturbation with the original field structure, giving rise to singularities or resonances. However, the KAM theorem (after Kolmogorov, Arnold and Moser – references are given below) guarantees the existence of certain invariant surfaces in some circumstances.

Loosely speaking, this theorem states that there are two regions of non-vanishing volume (technically, sets of non-zero measure), one of which is small compared to the other and shrinks to zero as the perturbation goes to zero. The (hopefully) larger region has the familiar structure of embedded tori (flux surfaces in the proper sense) covered with dense field-line trajectories. Interlaced between

these “good” flux surfaces, there is a possibly dense set of surfaces which open up into ergodic regions or islands, where the field lines behave erratically and may move far from nearby confined trajectories. However, even if a given magnetic field line is ergodic within an entire annular volume, it is absolutely contained, provided only that there is a closed (KAM) surface on either side. Further, the KAM theorem guarantees that for small perturbations, the chaotic regions will be thin, albeit of finite volume.

In conclusion, magnetic surfaces in non-symmetric devices exist only in the sense that the perturbed invariant toroidal surfaces of KAM theory exist as a set of non-zero measure (Grad 1967; Balescu 1975; Whiteman 1977). The degree of flux-surface destruction and island formation depends on the vacuum field structure and on the plasma pressure (β) via the diamagnetic and Pfirsch-Schlüter equilibrium currents (Boozer 1984 a, b; Reiman and Boozer 1984).

Stochastic motion of field lines and other KAM related issues have been studied by means of modern Hamiltonian mechanics. Although this is not the place to discuss these points in detail, it seems warranted to point out the Hamiltonian character of the magnetic-field lines and to show how the transformation to canonical coordinates is done in practice. This has been the subject of a series of papers by Boozer (1983, 1984c, 1984d, 1985) and by Kuo-Petric and Boozer (1986). We shall base our discussion on their work. An alternative route for describing magnetic-field lines by means of *non-canonical* Hamiltonian mechanics has been advocated by Cary and Littlejohn (1983), and Littlejohn (1984). Hamiltonian mechanics in general is authoritatively discussed in the classic book by Goldstein (1981). Hamiltonian theory applied to stochasticity is treated in Lichtenberg and Lieberman (1983). Also recommended are the review paper by Whiteman (1977), the appendix in Balescu’s (1975) book, and White’s book (1989).

Before dealing with the canonical coordinates and the Hamiltonian, we must first reconsider flux coordinates as a preliminary.

9.3 Flux Coordinates

From our discussion of flux coordinates, we recall that the magnetic field can be written as

$$\begin{aligned} \mathbf{B} &= \nabla\varrho \times \nabla \left(\frac{\dot{\Psi}_{\text{tor}}}{2\pi} \theta_f + \frac{\dot{\Psi}_{\text{pol}}^d}{2\pi} \zeta_f \right) \\ &= \frac{\dot{\Psi}_{\text{tor}}}{2\pi} \nabla\varrho \times \nabla\theta_f + \frac{\dot{\Psi}_{\text{pol}}^d}{2\pi} \nabla\varrho \times \nabla\zeta_f . \end{aligned} \quad (9.3.1)$$

Here, ϱ is a flux-surface label, and Ψ_{tor} and Ψ_{pol}^d measure the toroidal flux enclosed by, and the poloidal flux outside, the surface labeled by ϱ , respectively. Both Ψ_{tor}

and Ψ_{pol}^d depend only on ϱ . For simplicity of notation, we introduce two new flux functions:

$$\psi_t \equiv \frac{\Psi_{\text{tor}}}{2\pi} \quad (9.3.2a)$$

$$\psi_p \equiv \frac{\Psi_{\text{pol}}^d}{2\pi} . \quad (9.3.2b)$$

In terms of these functions, we have

$$\mathbf{B} = \nabla\psi_t \times \nabla\theta_f + \nabla\psi_p \times \nabla\zeta_f . \quad (9.3.3)$$

The rotational transform t , defined by $t \equiv -\dot{\Psi}_{\text{pol}}^d/\dot{\Psi}_{\text{tor}} \equiv -d\psi_p/d\psi_t$, allows us to write the above equation as

$$\mathbf{B} = \nabla\psi_t \times \nabla\theta_f - t \nabla\psi_t \times \nabla\zeta_f \quad (9.3.4a)$$

$$= \nabla\psi_t \times \nabla(\theta_f - t\zeta_f) . \quad (9.3.4b)$$

This is the best known representation. The flux coordinates are ψ_t , θ_f and ζ_f . t is a flux function, $t \equiv t(\psi_t)$. Equation (9.3.4) is the contravariant representation of \mathbf{B} , not because it is written in terms of (cross products of) contravariant-basis vectors ∇u^i , but because it deals with the (somewhat hidden) contravariant components. The more transparent contravariant form is revealed if we make use of the $e_i \leftrightarrow e^j$ “back and forth transformation” relations, where $e^j \equiv \nabla u^j$ and $e_i \equiv \partial R/\partial u^i$:

$$e_i = \frac{e^j \times e^k}{e^i \cdot e^j \times e^k} = \sqrt{g} e^j \times e^k . \quad (9.3.5)$$

Therefore, we can write instead of (9.3.4a):

$$\mathbf{B} = \frac{1}{\sqrt{g}} \left(\frac{\partial R}{\partial \zeta_f} + t(\psi_t) \frac{\partial R}{\partial \theta_f} \right) \quad (9.3.6a)$$

$$= \frac{t}{\sqrt{g}} e_{\theta_f} + \frac{1}{\sqrt{g}} e_{\zeta_f} . \quad (9.3.6b)$$

The contravariant components are thus $B^{\psi_t} = 0$, $B^{\theta_f} = t/\sqrt{g}$ and $B^{\zeta_f} = 1/\sqrt{g}$. B^{ψ_t} is zero because \mathbf{B} is tangent to the flux surfaces, i.e., $\mathbf{B} \cdot \nabla\psi_t \equiv 0$ and \sqrt{g} is the Jacobian:

$$\sqrt{g} = \frac{1}{\nabla\psi_t \cdot \nabla\theta_f \times \nabla\zeta_f} = \frac{\partial R}{\partial \psi_t} \cdot \frac{\partial R}{\partial \theta_f} \times \frac{\partial R}{\partial \zeta_f} = \frac{1}{\mathbf{B} \cdot \nabla\zeta_f} . \quad (9.3.7)$$

(Note the distinction between this Jacobian in terms of the toroidal flux and that defined in (6.1.38) which is a function of poloidal flux.)

From (9.3.4), we see that \mathbf{B} is uniquely determined if $\tau(\psi_t)$, ψ_t , θ_f and ζ_f are known as functions of the position vector \mathbf{R} such that the gradients can be calculated. According to Boozer (1983, 1985), it is more useful to define the field by (9.3.6) as far as a practical evaluation is concerned. The magnetic field can be constructed if $\mathbf{R}(\psi_t, \theta_f, \zeta_f)$, i.e., $x(\psi_t, \theta_f, \zeta_f)$, $y(\psi_t, \theta_f, \zeta_f)$, and $z(\psi_t, \theta_f, \zeta_f)$, in addition to $\tau(\psi_t)$ are known functions. Here, we assume a finite value of the Jacobian $(\sqrt{g})^{-1} = \mathbf{B} \cdot \nabla \zeta_f$. In other words, the toroidal field is nonzero.

The establishment of the transformation $\mathbf{R}(\psi_t, \theta_f, \zeta_f)$ by means of numerical field line following and Fourier decomposition was described by Boozer (1982, 1984b), Kuo-Petravic, Boozer, Rome and Fowler (1983), Kuo-Petravic (1984), Kuo-Petravic and Boozer (1986) and Beasley, Rome, Attenberger and Hirshman (1987). (Improvements to this method were made by Rome (1989) and by Reiman and Pomphrey (1989).) Because of the double periodicity, the three Cartesian components, x , y and z of \mathbf{R} can be written as a double Fourier series. We treat x , y and z simultaneously via the vector symbol $\mathbf{R}(\psi_t, \theta_f, \zeta_f) \equiv \mathbf{R}[x(\psi_t, \theta_f, \zeta_f), y(\psi_t, \theta_f, \zeta_f), z(\psi_t, \theta_f, \zeta_f)]$:

$$\mathbf{R}(\psi_t, \theta_f, \zeta_f) = \sum_{m,n} \mathbf{R}_{mn}(\psi_t) e^{i(m\theta_f - n\zeta_f)}. \quad (9.3.8)$$

Boozer's straight- \mathbf{B} coordinate system (Boozer 1980, 1981) makes use of a function χ as a parameter to measure the "length" along a field line.

$$\chi = \int_0^l B dl'. \quad (9.3.9)$$

χ reduces to the magnetic scalar potential if the plasma is currentless (i.e., $\mathbf{B} = \nabla \chi$) and is equal to (6.6.14c):

$$\chi = \frac{\mu_0}{2\pi} (I_{\text{tor}} \theta_f + I_{\text{pol}}^d \zeta_f). \quad (9.3.10)$$

I_{tor} and I_{pol}^d are flux functions. For flux coordinates where (9.3.10) holds, we earlier used a subscript B , i.e., $\theta_f \equiv \theta_B$ and $\zeta_f \equiv \zeta_B$.

Since the equation of a field line in flux coordinates is given by (4.9.15),

$$\theta_{f0} = \theta_f - \tau \zeta_f, \quad (9.3.11)$$

we can rewrite the exponent of the Fourier series in terms of θ_{f0} and χ :

$$\mathbf{R}(\psi_t, \theta_f, \zeta_f) = \sum_{m,n} \mathbf{R}_{mn}(\psi_t) \exp \left[i \frac{(mI_{\text{pol}}^d + nI_{\text{tor}})\theta_{f0} + (m\tau - n)2\pi\chi/\mu_0}{I_{\text{tor}}\tau + I_{\text{pol}}^d} \right] \quad (9.3.12)$$

To evaluate the Cartesian coefficients $x_{mn}(\psi_t)$, $y_{mn}(\psi_t)$ and $z_{mn}(\psi_t)$ of $\mathbf{R}_{mn}(\psi_t)$, we perform a field line integration, starting from an arbitrary point. We choose χ and θ_{f0} zero at the starting point. Since θ_{f0} remains constant along the field line

trajectory, θ_{f0} remains zero everywhere. χ is determined by the integration via (9.3.9), and thus, this field line integration determines $\mathbf{R}(\psi_t, \chi) \equiv \mathbf{R}(\psi_t, \theta_{f0} = 0, \chi)$. We then have for (9.3.12), where the L.H.S. is now known numerically from the field-line integration:

$$\mathbf{R}(\psi_t, \theta_f, \zeta_f) = \sum_{m,n} \mathbf{R}_{mn}(\psi_t) \exp \left\{ i \frac{(m\tau - n)2\pi\chi}{\mu_0(I_{\text{tor}}\tau + I_{\text{pol}}^d)} \right\}. \quad (9.3.13)$$

This Fourier decomposition of $\mathbf{R}(\psi_t, \theta_f, \zeta_f) \equiv \mathbf{R}(\psi_t, \chi)$ has distinct peaks. The amplitudes of the peaks determine the $\mathbf{R}_{mn}(\psi_t)$, and their locations determine m , n and $\tau(\psi_t)$. (The practical numerical decomposition is done with the help of a suitable window function. For references, see Sect. 8.1.) In practice, it turns out that "about 5 poloidal and toroidal circuits of a field line must be followed to carry out the Fourier decomposition to reasonable accuracy" (Boozer 1982). This way, $\mathbf{R}(\psi_t, \theta_f, \zeta_f)$ and $\tau(\psi_t)$ are known, and thus, the flux coordinate representation of (9.3.6) can be formed. The method outlined above is only applicable in the Boozer coordinate system, where (9.3.10) holds and where $\zeta_f \equiv \zeta_B$ and $\theta_f \equiv \theta_B$. A method valid for other straight- \mathbf{B} systems is explained below.

The above explanation is in principle only valid for irrational surfaces, i.e., surfaces that are traced out by a single field line. This is not the case for rational surfaces on which τ is the ratio of two integers. However, as long as the number of poloidal and toroidal transits is sufficiently large (e.g., larger than 5), the surface can be considered as being adequately covered for all practical purposes. When τ is the ratio of low integers, we invoke the fact that the real number system is dense and thus that any rational number (rational surface) can be approximated infinitesimally closely by an irrational number (irrational surface).

In a later paper, Kuo-Petravic and Boozer (1986) perform the field line integration and Fourier decomposition without mentioning the function χ (be aware that in that paper, the symbol χ is used with a different meaning). There, they start out with a definite choice of the toroidal angle ζ_f , which for convenience is taken to be (minus) the azimuthal angle of the cylindrical system $(R, -\zeta_f, z)$. This is effectively the angle that we called ζ_e . In this case the field line integration is based on the equations

$$\frac{dx}{d\zeta_f} = \frac{B^x(x, y, z)}{\mathbf{B} \cdot \nabla \zeta_f}, \quad \frac{dy}{d\zeta_f} = \frac{B^y(x, y, z)}{\mathbf{B} \cdot \nabla \zeta_f}, \quad \frac{dz}{d\zeta_f} = \frac{B^z(x, y, z)}{\mathbf{B} \cdot \nabla \zeta_f}$$

or

$$\frac{d\mathbf{R}}{d\zeta_f} = \frac{\mathbf{B}(\mathbf{R})}{\mathbf{B} \cdot \nabla \zeta_f}. \quad (9.3.14)$$

This allows determination of the function $\mathbf{R}(\zeta_f)$ on a particular flux-surface. Again, $\mathbf{R}(\psi_t, \theta_f, \zeta_f)$ and $\tau(\psi_t)$ can be obtained via a Fourier decomposition

$$\mathbf{R}(\psi_t, \theta_f, \zeta_f) = \sum_{m,n} \mathbf{R}_{mn}(\psi_t) e^{i(m\theta_f - n\zeta_f)}$$

or

$$\mathbf{R}(\psi_t, \theta_f, \zeta_f) \equiv \mathbf{R}(\zeta_f) = \sum_{m,n} R_{mn}(\psi_t) e^{i(m+n)\zeta_f}. \quad (9.3.15)$$

The notation $\mathbf{R}(\zeta_f)$ indicates that we have chosen a fixed flux surface. As before, θ_{f0} has been chosen equal to zero. The expansion in (9.3.15) performed on a variety of flux surfaces determines R_{mn} and τ .

9.4 Canonical Coordinates; The Field-Line Hamiltonian

We are now ready to treat the magnetic-field representation in the more general case where field lines are not restricted to a flux surface but cover a finite volume ergodically. The types of systems that we are interested in for magnetic confinement and whose description is still more or less tractable analytically are those where the magnetic field $\mathbf{B}(\mathbf{R})$ does not differ very much from a field $\mathbf{B}_o(\mathbf{R})$ which has perfect surfaces. The difference between the two fields can then be considered as a perturbation to the "perfect" field.

When a field is represented by means of a "magnetic field line Hamiltonian", the field lines can be identified with the phase-space trajectories produced by the Hamiltonian. Mathematically, the issue of magnetic-surface quality then reduces to a problem in Hamiltonian mechanics, meaning that a whole wealth of methods, such as Canonical Transformations, Hamilton-Jacobi theory, Canonical Perturbation theory, mappings, etc. are at our disposal.

Hamilton's equations for magnetic fields were indicated by Kerst (1962). The Hamiltonian character of the magnetic field was implicitly used by Spitzer (1958) in the paper in which he presented the stellarator concept, and by Kruskal and Kulsrud (1958). More explicit references to the Hamiltonian relationship were made by Rosenbluth, Sagdeev and Taylor (1966), Filonenko, Sagdeev and Zaslavsky (1967), and by Grad (1967). The field-line problem in the context of Hamiltonian mechanics is discussed by Lichtenberg and Lieberman (1983), who provide more references. Other references to the Hamiltonian nature of field lines were given in Sect. 9.2.

The simplest proof that field line trajectories are derivable from a Hamiltonian was suggested by H. Weitzner (according to Boozer (1985)). A globally divergence free vector field can be represented by a single-valued vector potential $\mathbf{A}(\mathbf{R})$. If we consider a set of variables (ϱ, θ, ζ) , where ϱ is a radius-like variable characterized by $\varrho = 0$ along the magnetic axis, we can write for the vector field $\mathbf{A}(\mathbf{R})$:

$$\mathbf{A}(\mathbf{R}) \equiv \mathbf{A}(\varrho, \theta, \zeta) = A_\varrho \nabla \varrho + A_\theta \nabla \theta + A_\zeta \nabla \zeta. \quad (9.4.1)$$

Here A_ϱ , A_θ , and A_ζ are single valued functions of (ϱ, θ, ζ) . Note that ϱ is in general not a flux-surface label because flux surfaces may not exist.

We know that a magnetic vector potential is determined up to an additive term, which is the gradient of a scalar function, $\nabla G(\varrho, \theta, \zeta)$. Indeed, we have that $\mathbf{B} = \nabla \times \mathbf{A} = \nabla \times (\mathbf{A} + \nabla G)$. Therefore, we could choose a G such that $\partial G / \partial \varrho = A_\varrho$, $G = 0$ at $\varrho = 0$, $\partial G / \partial \theta = A_\theta - \Psi$ and $\partial G / \partial \zeta = A_\zeta - \Phi$. Ψ and Φ are two single-valued functions that will be interpreted below. With this choice for G , we have

$$\mathbf{A} = \Psi \nabla \theta + \Phi \nabla \zeta + \nabla G. \quad (9.4.2)$$

Upon taking the curl of this vector potential, we obtain for \mathbf{B} :

$$\mathbf{B} = \nabla \Psi \times \nabla \theta + \nabla \Phi \times \nabla \zeta. \quad (9.4.3)$$

This representation of \mathbf{B} is known as the *canonical form*. An alternative proof (for which $\mathbf{B} \cdot \nabla \zeta$ must be non-zero) was given by Boozer (1985).

The expression in (9.4.3) resembles the flux-coordinate representation of (9.3.3), but the similarity is superficial; (9.3.3) is only valid for a \mathbf{B} field with proper magnetic surfaces everywhere, whereas (9.4.3) is valid for any magnetic field in a toroidal device. The difference is that in the flux-surface representation, both ψ_p and ψ_t are functions of only the flux label ϱ . In other words, $\psi_p \equiv \psi_p(\psi_t)$ only. The functions Ψ and Φ of (9.4.3), on the contrary, are both functions of the three variables, ϱ , θ , and ζ : $\Phi \equiv \Phi(\varrho, \theta, \zeta)$, $\Psi \equiv \Psi(\varrho, \theta, \zeta)$. Because Ψ and Φ depend on all three coordinates ϱ, θ, ζ , it follows that θ and ζ are arbitrary (poloidal and toroidal) angles, whereas the θ_f and ζ_f in (9.3.3) must be flux coordinates for (9.3.3) to be valid.

From the knowledge of $\mathbf{B}(\mathbf{R})$ everywhere, it is possible to construct functions Ψ and Φ such that \mathbf{B} is represented by (9.4.3). This will be discussed below. We first give a physical interpretation of Ψ and Φ after which we shall indicate the connection with Hamiltonian mechanics.

It is possible to infer the following statement from the formal similarity between (9.3.3) and (9.4.3), but a rigorous proof can also be constructed. The function $\Psi(\varrho, \theta, \zeta)$ measures the toroidal flux through a constant Ψ surface: the toroidal flux equals $2\pi\Psi$. Similarly, $2\pi\Phi$ is the poloidal flux outside a constant Φ surface (i.e., through a disk tangent to a constant Φ surface). The formal proof makes use of (9.4.3), the expression for a surface element in a constant u^i surface, $dS(i) = \sqrt{g} du^i du^k \nabla u^i$, the definition of (flux through surface $u^i = \text{constant}$) = $\iint \mathbf{B} \cdot dS(i)$, and the fact that $\Phi \equiv \Phi(\text{disk})$ with $d\Phi = -d\Phi(\text{ribbon})$. Note that the surfaces $\Psi(\varrho, \theta, \zeta) = \text{constant}$ and $\Phi(\varrho, \theta, \zeta) = \text{constant}$ do not coincide, in general. Also, in general, $\mathbf{B} \cdot \nabla \Psi \neq 0 \neq \mathbf{B} \cdot \nabla \Phi$. Only in a system where flux surfaces exist are Φ and Ψ equivalent. In that case, Ψ can be taken equal to ψ_t while $\Phi \equiv \psi_p$.

As mentioned above, (9.4.3) is called the canonical representation of \mathbf{B} (Boozer 1985). The reason is that the four functions, Φ , Ψ , θ , ζ can be considered as canonical variables in a Hamiltonian context. Φ will be identified with the

"Hamiltonian", ζ with the canonical "time", Ψ with the "momentum", and θ with the "coordinate".

To see this, we first eliminate the "auxiliary" variable ϱ from the problem and consider Φ as a function of the coordinates, Ψ , θ , ζ ; so $\Phi \equiv \Phi(\Psi, \theta, \zeta)$. In terms of Ψ , θ , ζ coordinates, the equation of a field line reads:

$$\frac{d\Psi}{B \cdot \nabla \Psi} = \frac{d\theta}{B \cdot \nabla \theta} = \frac{d\zeta}{B \cdot \nabla \zeta}. \quad (9.4.4)$$

To find $B \cdot \nabla \Psi$, we dot multiply (9.4.3) with $\nabla \Psi$ and obtain

$$B \cdot \nabla \Psi = \nabla \Psi \cdot \nabla \Phi \times \nabla \zeta. \quad (9.4.5)$$

But with $\nabla \Phi = (\partial \Phi / \partial \Psi) \nabla \Psi + (\partial \Phi / \partial \theta) \nabla \theta + (\partial \Phi / \partial \zeta) \nabla \zeta$, this last equation becomes

$$B \cdot \nabla \Psi = \frac{\partial \Phi}{\partial \theta} \nabla \Psi \cdot \nabla \theta \times \nabla \zeta. \quad (9.4.6)$$

Therefore, we have:

$$\frac{d\Psi}{d\zeta} = \frac{B \cdot \nabla \Psi}{B \cdot \nabla \zeta} = \frac{\partial \Phi}{\partial \theta} \frac{\nabla \Psi \cdot \nabla \theta \times \nabla \zeta}{\nabla \zeta \cdot \nabla \Psi \times \nabla \theta}$$

or

$$\frac{d\Psi}{d\zeta} = \frac{\partial \Phi}{\partial \theta}. \quad (9.4.7)$$

Analogously, from the second and third equations, we have

$$\frac{d\theta}{d\zeta} = -\frac{\partial \Phi}{\partial \Psi}. \quad (9.4.8)$$

Thus, after identification of ζ with t , of Φ with $-H$, of θ with the coordinate q , and of Ψ with the momentum p , (9.4.7) and (9.4.8) are Hamilton's equations of motion: $\dot{p} = -\partial H / \partial q$ and $\dot{q} = \partial H / \partial p$. To emphasize this result, $-\Phi$ is called the *field-line Hamiltonian*. Because they satisfy Hamilton's equations, (9.4.7) and (9.4.8), Ψ and θ are canonical variables or conjugate variables.

It is sometimes useful to consider the time as a coordinate and the negative Hamiltonian as its conjugate momentum. This then allows us to consider ζ and Φ as canonical variables as well; hence the name canonical representation for (9.4.3).

The underlying idea is that in the general case, our magnetic-field Hamiltonian is dependent on toroidal angle ζ or on "time". We are thus dealing with a "time" dependent Hamiltonian problem in one degree of freedom (characterized by one coordinate and one momentum). It is now possible by means of canonical

transformation to "transform away" the time dependence by considering an "extended phase space", where Hamilton's equations are parameterized by a new "time" τ (Lichtenberg and Lieberman 1983):

$$\begin{array}{ll} q_1 \equiv \theta & q_1 \equiv \theta \\ q_2 \equiv t & q_2 \equiv \zeta \\ p_1 \equiv \Psi & p_1 \equiv \Psi \\ p_2 \equiv -H & p_2 \equiv \Phi \end{array} \quad \text{or} \quad \begin{array}{ll} & q_1 \equiv \theta \\ & q_2 \equiv \zeta \\ & p_1 \equiv \Psi \\ & p_2 \equiv \Phi \end{array}$$

The new Hamiltonian is time independent.

We are now dealing with an equivalent time-independent Hamiltonian in two degrees of freedom (2 coordinates and 2 momenta). This permits use of all the computational tools that have been devised for conservative systems. It also makes it clear that the field-line problem is characterized by a similar level of complexity as a two degree-of-freedom system, and this implies the existence of stochastic field-line behavior.

When flux surfaces exist, the magnetic field line Hamiltonian is only a function of Ψ : $-\Phi_0 \equiv -\Phi_0(\Psi)$, since then $\Phi_0 \equiv \psi_p(\varrho) \equiv \psi_p(\psi_t)$. This means that the Hamiltonian is in action-angle form and that the action Ψ (denoted J in Hamiltonian mechanics) is a conserved quantity, while the angle $\theta = \theta_f$ depends linearly on the toroidal angle $\zeta = \zeta_f$ (Hamiltonian time t). The latter relationship is exactly the equation of a field line in flux coordinates

$$\theta_f = t(\Psi)\zeta_f + \theta_{f0}. \quad (9.4.9)$$

The rotational transform t plays the role of a constant "frequency" in the action-angle picture.

In Hamiltonian language, a magnetic field with perfect magnetic surfaces everywhere is *integrable* (Lichtenberg and Lieberman (1983); Goldstein (1981) uses the word "separable"). Since we are interested in magnetic field configurations that have "reasonably good" surfaces in the sense explained in Sect. 9.2., we also wish to consider *near-integrable* systems, where the canonical variables of the field-line Hamiltonian are chosen so that the Hamiltonian is a function of the canonical momentum (\equiv action) alone, plus a small perturbation. Such Hamiltonians are said to be in "near action-angle form." In our field-line case considered above, it would look like:

$$\Phi(\Psi, \theta, \zeta) = \Phi_0(\Psi) + \Phi_1(\Psi, \theta, \zeta). \quad (9.4.10)$$

The field-line Hamiltonian contains all the information on the existence of surfaces, islands and stochastic regions. The near-integrable Hamiltonian problem is treated by means of the canonical perturbation theory of Hamiltonian mechanics (see Goldstein (1981), and Lichtenberg and Lieberman (1983)).

9.5 Practical Evaluation of the Field-Line Hamiltonian

Critical to this evaluation of the Hamiltonian function $\Phi(\Psi, \theta, \zeta)$ is the assumption that the given magnetic field $\mathbf{B}(\mathbf{R})$ is close to one with perfect magnetic surfaces $\mathbf{B}_0(\mathbf{R})$. The method to be discussed was developed by Boozer (1983), and the numerical implementation of it has been considered by Kuo-Petravic and Boozer (1986).

By analogy to the case where perfect surfaces exist, we are interested in the transformation equation $\mathbf{R}(\psi_i, \theta_f, \zeta_f)$. Here (ϱ, θ, ζ) are set equal to $(\psi_i, \theta_f, \zeta_f)$, being the flux coordinates for the magnetic field related to the unperturbed Hamiltonian $\Phi_0(\Psi) \equiv \Phi_0(\psi_i)$. Note that the angles of the canonical set are arbitrary, thus it is perfectly legitimate to make this choice.

We can no longer evaluate $\mathbf{R}_{mn}(\psi_i)$ and $t(\psi_i)$ for all values of ψ_i . However, it can still be done over much of the plasma volume, for many ψ_i values of interest. Smooth interpolation between these values then leads to $\mathbf{R}_{mn}(\psi_i)$ or $\mathbf{R}(\psi_i, \theta_f, \zeta_f)$ and $t(\psi_i)$. Knowledge of these quantities allows the construction of the "gedanken" field with perfect surfaces \mathbf{B}_0 , according to (9.3.6a), in which the left-hand side now reads \mathbf{B}_0 . Keep in mind that $(\psi_i, \theta_f, \zeta_f)$ are not flux coordinates for \mathbf{B} , only for \mathbf{B}_0 . Thus $\mathbf{B}_0 \cdot \nabla \psi_i \equiv 0$, but $\mathbf{B} \cdot \nabla \psi_i \neq 0$. According to Boozer (1983), $(\psi_i, \theta_f, \zeta_f)$ form an appropriate system for determining Ψ and Φ .

When we consider (9.4.3) with $\theta = \theta_f$, $\zeta = \zeta_f$, $\Psi \equiv \Psi(\psi_i, \theta_f, \zeta_f)$ and $\Phi \equiv \Phi(\psi_i, \theta_f, \zeta_f)$, and expand $\nabla \Psi$ and $\nabla \Phi$ along the basis vectors $\nabla \psi_i$, $\nabla \theta_f$ and $\nabla \zeta_f$ via the chain rule, we obtain

$$\mathbf{B} = \left(\frac{\partial \Phi}{\partial \theta_f} - \frac{\partial \Psi}{\partial \zeta_f} \right) (\nabla \theta_f \times \nabla \zeta_f) - \frac{\partial \Phi}{\partial \psi_i} (\nabla \zeta_f \times \nabla \psi_i) + \frac{\partial \Psi}{\partial \psi_i} (\nabla \psi_i \times \nabla \theta_f). \quad (9.5.1)$$

The usual way to write \mathbf{B} in terms of contravariant components is

$$\mathbf{B} = (\mathbf{B} \cdot \nabla \psi_i) \mathbf{e}_{\psi_i} + (\mathbf{B} \cdot \nabla \theta_f) \mathbf{e}_{\theta_f} + (\mathbf{B} \cdot \nabla \zeta_f) \mathbf{e}_{\zeta_f}. \quad (9.5.2)$$

If we now substitute the $e_i \leftrightarrow e^i$ "back and forth transformation", given in (9.3.5) into (9.5.2), and use its dual form,

$$\mathbf{e}^i = \nabla u^i = \frac{\mathbf{e}_j \times \mathbf{e}_k}{\mathbf{e}_i \cdot \mathbf{e}_j \times \mathbf{e}_k} = \frac{1}{\sqrt{g}} \left(\frac{\partial \mathbf{R}}{\partial u^j} \times \frac{\partial \mathbf{R}}{\partial u^k} \right) \quad (9.5.3)$$

in the $\mathbf{B} \cdot \nabla u^i$ expressions of (9.5.2), we obtain after comparing with (9.5.1):

$$\frac{\partial \Phi}{\partial \theta_f} - \frac{\partial \Psi}{\partial \zeta_f} = \mathbf{B} \cdot \left(\frac{\partial \mathbf{R}}{\partial \theta_f} \times \frac{\partial \mathbf{R}}{\partial \zeta_f} \right), \quad (9.5.4a)$$

$$\frac{\partial \Phi}{\partial \psi_i} = \mathbf{B} \cdot \left(\frac{\partial \mathbf{R}}{\partial \psi_i} \times \frac{\partial \mathbf{R}}{\partial \zeta_f} \right), \quad (9.5.4b)$$

and

$$\frac{\partial \Psi}{\partial \psi_i} = \mathbf{B} \cdot \left(\frac{\partial \mathbf{R}}{\partial \psi_i} \times \frac{\partial \mathbf{R}}{\partial \theta_f} \right). \quad (9.5.4c)$$

The right-hand side of these equations is known, since $\mathbf{B}(\mathbf{R})$ is given and $\mathbf{R}(\psi_i, \theta_f, \zeta_f)$ is known after field-line integration and Fourier decomposition. In fact, the last two equations determine Ψ and Φ . The first one can be dropped as it does not provide any new information; it is consistent with the other two equations, since its derivation was based on the general form of \mathbf{B} in (9.4.3), which in turn was based on (9.4.2) and $\mathbf{B} = \nabla \times \mathbf{A}$ and thus on $\nabla \cdot \mathbf{B} = 0$.

Integration of (9.5.4b) and (9.5.4c) allows determination of $\Phi(\psi_i, \theta_f, \zeta_f)$ and $\Psi(\psi_i, \theta_f, \zeta_f)$. In practice, one picks a number of different θ_f and ζ_f values and uses a fast Fourier transform. After elimination of the "auxiliary" radial variable ψ_i , one obtains the field-line Hamiltonian $-\Phi$ in terms of its variables $(\Psi, \theta_f, \zeta_f)$ in Fourier decomposed form.

The nature of θ_f and ζ_f , being flux coordinates for \mathbf{B}_0 , is not important for the canonical representation of \mathbf{B} . Other coordinates, ϱ , θ , and ζ can do the job as well. This alternative and more general set was used by Kuo-Petravic and Boozer (1986). In terms of these general coordinates, the field-line Hamiltonian, in Fourier decomposed form, is

$$\Phi(\Psi, \theta, \zeta) = \Phi_0(\Psi) + \sum_{m,n} \Phi_{mn}(\Psi) e^{i(m\theta - n\zeta)}. \quad (9.5.5)$$

Φ_0 is the θ, ζ averaged part of Φ and is the Hamiltonian of \mathbf{B}_0 . Since Ψ is known as a function of (ψ_i, θ, ζ) it is possible to construct the canonical transformation equations from Cartesian coordinates $\mathbf{R}(x, y, z)$ to (Ψ, θ, ζ) as $\mathbf{R}(\Psi, \theta, \zeta)$. All the information regarding the magnetic field is contained in the functions $\Phi(\Psi, \theta, \zeta)$ and $\mathbf{R}(\Psi, \theta, \zeta)$. A typical example based on the theory outlined above has been worked out by Kuo-Petravic and Boozer (1986).

Part III

Selected Topics

In the third and last part of this book, we deal with five subjects that are related to some issues we addressed in the earlier chapters, but which, in the interest of not interrupting the flow of reading, have been postponed until the end of this treatise. In some cases, the topics to be covered now are clarifications of material discussed earlier; in other cases, they should be seen as additional information, given for the sake of completeness. (The topics covered do not necessarily reflect a preference by the authors; the subjects presented themselves automatically during the course of writing.)

In Chap. 10, we treat “proper” toroidal coordinates. This coordinate system is to be considered as a three-dimensional version of the bipolar coordinate system. Although this system is not frequently used for toroidally-confined-plasma computations, we think it is important that the reader should be aware of its existence, and understand the fundamental distinction between this system and those treated in earlier chapters.

Chapter 11 deals with the dynamic equilibrium of an ideal tokamak plasma. Fully realizing that the notion of an infinitely conducting plasma only has meaning in a scaling sense (i.e., that the zero-resistivity plasma is an asymptotic case), we interpret $\sigma \rightarrow \infty$ literally, and consider the consequences for a tokamak “through” which the transformer is pumping magnetic flux. We show that both the plasma and the flux surfaces move radially inward due to the presence of a perpendicular electric field, confirming the frozen-flux theorem.

In Chap. 12, we clarify the relationship and the distinction between the specific flux volume $dV/d\Psi_{\text{tor}}$ and the “proper” length of a field line $\int dl/B$.

Chapter 13 treats the transformation properties of vector and tensor components. It clarifies the names covariant and contravariant. It also makes clear why the normally used Levi-Civita symbols and the Christoffel symbols do not qualify as tensors.

Finally, in Chap. 14, we give two alternative derivations of the divergence formula. Although the derivation given in Chap. 2 suffices from the “result” point of view, it is interesting to see how (more elaborate) approaches, used frequently in the mathematical-physics literature, lead to the same result.

10. “Proper” Toroidal Coordinates

10.1 Introduction

“Proper” toroidal coordinates are not frequently utilized for describing the physics of toroidally confined plasmas. However, they have been employed occasionally (Shafranov 1960 and Fielding and Hitchon 1980). Furthermore, these coordinates have their own intrinsic value, since many problems in mathematical physics in toroidal geometry have made use of them (see e.g. Morse and Feshbach 1953).

“Proper” toroidal coordinates are constructed by three-dimensionalizing the two-dimensional “bipolar” coordinate system by means of a rotation, just like spherical coordinates are the three-dimensional generalization of the two-dimensional polar system. We will first introduce the bipolar coordinate system in an intuitive way, explain its meaning, and simply write down the equations of the constant coordinate curves. Then we show the analogy between the 3-D spherical system and the 2-D polar system. Next, we will exploit this analogy to construct the “real toroidal” coordinate system. Finally, we derive the equations of the 2-D bipolar system, a task that we postpone until the end in order not to interrupt the logic.

10.2 Bipolar Coordinates: Intuitive Considerations

In two dimensions, we always work with curves, not with surfaces. The position of a point is determined by the intersection of two constant-coordinate curves $u^1 = c^1$ and $u^2 = c^2$, where c^1 and c^2 are constants. The two-dimensional bipolar coordinate system is based on two intersecting families of circles, $\eta = \text{constant}$ and $\xi = \text{constant}$, as shown in Fig. 10.1. These curves resemble the loci of constant electrostatic potential and electric-field lines of two point charges of opposite signs located at points A and B . The coordinates ξ and η of a point P are defined as

$$\xi = \alpha_1 - \alpha_2 \quad \text{and} \quad \eta = \ln\left(\frac{d_2}{d_1}\right) \quad (10.2.1)$$

where d_1 and d_2 represent the distances from the fixed points A and B to the point P , respectively, and α_1 and α_2 are the angles between the x -axis and the

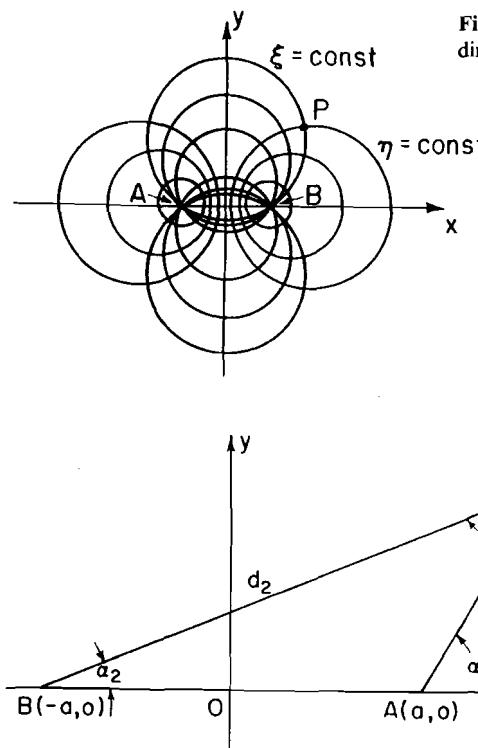


Fig. 10.1. Constant coordinate curves of the two-dimensional bipolar coordinate system

lines AP and BP . These quantities are shown in Fig. 10.2. The transformation formulae are

$$x = \frac{a \sinh \eta}{\cosh \eta - \cos \xi} \quad 0 \leq \xi < 2\pi \quad (10.2.2a)$$

$$y = \frac{a \sin \xi}{\cosh \eta - \cos \xi} \quad -\infty < \eta < +\infty \quad (10.2.2b)$$

and the constant coordinate curves are given by

$$x^2 + (y - a \cot \xi)^2 = a^2 \csc^2 \xi \quad (10.2.3)$$

$$(x - a \coth \eta)^2 + y^2 = a^2 \operatorname{csch}^2 \eta \quad (10.2.4)$$

where a is defined in Fig. 10.2. Both constant coordinate curves are circles. One family has its center on the y -axis (10.2.3), the other on the x -axis (10.2.4). Their radii are dependent on the values of ξ and η . Note that the family represented by (10.2.3) passes through the points A and B for all values of ξ .

10.3 The Relationship of 3-D Spherical Coordinates to 2-D Polar Coordinates

Now, we compare the transformation formulae of the common 2-D polar coordinate system of Fig. 10.3 with the 3-D spherical system of Fig. 10.4. The latter is obtained by rotating the 2-D system about its y -axis, after which we rename the axes: the 2-D y -axis becomes the z -axis of the 3-D system, and the 2-D x -axis is now identified with the axis denoted by an asterisk (*) in Fig. 10.4, which we now call the ϱ -axis. Thus, the 2-D x - y plane can be thought of as the z - ϱ plane in 3-D. Then the 3-D x - and y -axes are constructed such that they make a right angle with each other, and \hat{x} , \hat{y} , \hat{z} form a right-handed basis. Note that $\varrho = \sqrt{x^2 + y^2}$ in the 3-D system. The location of the ϱ -axis is fixed by prescribing an angle ϕ such that $x = \varrho \cos \phi$ and $y = \varrho \sin \phi$. If we now observe that the angle Θ in 3-D is measured from the 3-D z -axis (our former 2-D y -axis), whereas angle θ in the polar system is measured from the 2-D x -axis (the ϱ -axis in 3-D), then we know that $\sin \theta$ in 2-D should be replaced by $\cos \Theta$ in 3-D and vice-versa.

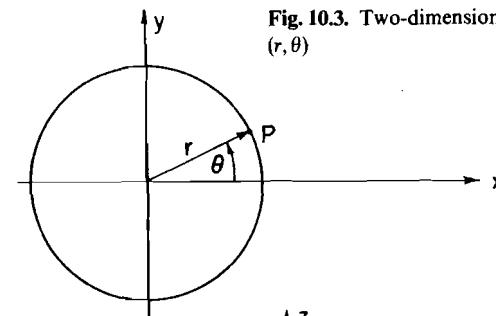


Fig. 10.3. Two-dimensional polar coordinate system. P has coordinates (r, θ)

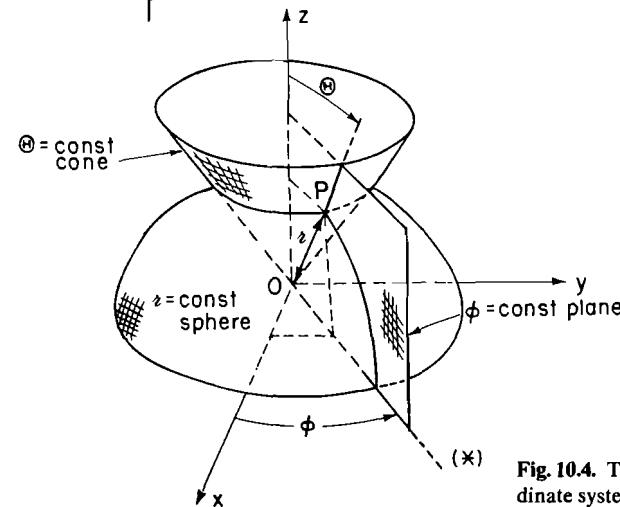


Fig. 10.4. Three-dimensional spherical coordinate system. P has coordinates (r, Θ, ϕ)

Thus, the 2-D equations $x = r \cos \theta$ and $y = r \sin \theta$, become in 3-D: $\varrho = z \sin \Theta$ $z = z \cos \Theta$ (r and z are defined in Figs. 10.3 and 10.4, respectively; in the present context, $z \equiv r$). By using then the angle ϕ , we also obtain x and y in the 3-D system. Schematically, we have thus

$$\begin{array}{lll} \text{2-D Polar} & & \text{3-D Spherical} \\ x = r \cos \theta & \rightarrow & \varrho = z \sin \Theta \quad \rightarrow \quad x = z \sin \Theta \cos \phi \\ y = r \sin \theta & & z = z \cos \Theta \quad \quad \quad y = z \sin \Theta \sin \phi \\ & & \quad \quad \quad x = \varrho \cos \phi \\ & & \quad \quad \quad z = z \cos \Theta \\ & & \quad \quad \quad y = \varrho \sin \phi \end{array}$$

The constant coordinate curves are "transformed" in a similar way to constant coordinate surfaces. We see that the 2-D $x^2 + y^2 = r^2 = \text{constant}$ circle becomes a sphere: $\varrho^2 + z^2 = z^2 = \text{constant}$ or $x^2 + y^2 + z^2 = z^2 = \text{constant}$. The 2-D straight line $\theta = \text{constant}$ is transformed into a cone $\Theta = \text{constant}$. The third constant coordinate surface is a new one, originating from the (rotation) angle $\phi = \text{constant}$.

10.4 "Proper" Toroidal Coordinates as a 3-D Version of Bipolar Coordinates

Completely analogously, we can now establish the "proper" toroidal coordinate system by rotating the 2-D bipolar system of Fig. 10.1 about its y -axis, such that the $\xi = \text{constant}$ circles go over into spheres, while the $\eta = \text{constant}$ circles become tori. With similar replacements as above, i.e., $x(2\text{-D}) \rightarrow \varrho = \sqrt{x^2 + y^2}$ (3-D), $y(2\text{-D}) \rightarrow z(3\text{-D})$, a (rotation) angle ζ such that $x(3\text{-D}) = \varrho \cos \zeta$ and $y(3\text{-D}) = \varrho \sin \zeta$, and with (the old bipolar variables) ξ and η now retaining their original meaning in contrast to the $\Theta \leftrightarrow \theta$ relationship, we obtain

$$x = \frac{a \sinh \eta}{\cosh \eta - \cos \xi} (\cos \zeta) \quad (10.4.1a)$$

and

$$y = \frac{a \sinh \eta}{\cosh \eta - \cos \xi} (\sin \zeta) \quad (10.4.1b)$$

from (10.2.2a). According to (10.2.2b), we have then for z :

$$z = \frac{a \sin \xi}{\cosh \eta - \cos \xi}. \quad (10.4.1c)$$

If we take the toroidal coordinates as (ξ, η, ζ) in that order, then a right-handed

system implies that ζ should be measured from the x -axis towards the y -axis. This clarifies the above relationships between x , y , and ζ .

The equations for the constant coordinate surfaces are easily obtained from (10.2.3) and (10.2.4) by using our prescription above. Thus,

$$x^2 + y^2 + (z - a \cot \xi)^2 = a^2 \csc^2 \xi ; \quad \xi = \text{constant} , \quad (10.4.2a)$$

is the equation of a family of spheres with center on the z -axis (at $z = a \cot \xi$) and radius ($a \csc \xi$). Also,

$$((x^2 + y^2)^{1/2} - a \coth \eta)^2 + z^2 = a^2 \operatorname{csch}^2 \eta ; \quad \eta = \text{constant} , \quad (10.4.2b)$$

represents the equation of a toroidal surface, with the minor axis having a radius ($a \coth \eta$). The cross sections with the $z - \varrho$ plane ($\zeta = \text{constant}$ plane) are circles with radius ($a \operatorname{csch} \eta$). Finally, we have a plane $\zeta = \text{constant}$ (the $z - \varrho$ plane) with equation

$$y = x \tan \zeta ; \quad \zeta = \text{const} . \quad (10.4.2c)$$

The coordinate system is an orthogonal one, and its scale factors h_i are

$$h_\xi = h_\eta = \frac{a}{\cosh \eta - \cos \xi} \quad (10.4.3a)$$

$$h_\zeta = \frac{|a \sinh \eta|}{\cosh \eta - \cos \xi} . \quad (10.4.3b)$$

10.5 The Bipolar Coordinate System: A Detailed Analysis

Our last task is to derive (10.2.2) for the bipolar system. (See e.g. Margenau and Murphy 1976.) We find it convenient to introduce a complex variable u , defined by $u = x + iy$. Its complex conjugate is denoted by $u^* = x - iy$. Since $x = (u + u^*)/2$ and $y = i(u^* - u)/2$, we would like to find expressions for u and u^* as functions of the bipolar variables ξ and η defined in (10.2.1). From Fig. 10.2, we obtain that

$$u = a + d_1 e^{i\alpha_1} = -a + d_2 e^{i\alpha_2} . \quad (10.5.1)$$

If we write the first equation as $u - a$, the second as $u + a$, and divide the two by each other, we find

$$\frac{u + a}{u - a} = \frac{d_2}{d_1} e^{-i(\alpha_1 - \alpha_2)} = e^{+\eta} e^{-i\xi} = e^{-i(\xi + i\eta)} . \quad (10.5.2)$$

Here, we have used the definition equations of ξ and η given in (10.2.1). Then from (10.5.2), it follows that

$$u = a \frac{e^{-i(\xi+i\eta)} + 1}{e^{-i(\xi+i\eta)} - 1} = a \frac{e^{-iw} + 1}{e^{-iw} - 1} \quad (10.5.3)$$

with $w = \xi + i\eta$. From the Euler formula, $e^{iw} = \cos w + i \sin w$, and thus, $\tan w = \sin w / \cos w = i(e^{-iw} - e^{iw})/(e^{-iw} + e^{iw}) = i(1 - e^{2iw})/(1 + e^{2iw})$, so we can rewrite (10.5.3):

$$u = +ia \cot(w/2) \quad (10.5.4a)$$

$$u^* = -ia \cot(w^*/2) . \quad (10.5.4b)$$

Now, we substitute (10.5.4) into $x = (u + u^*)/2$, and use the identities $\cot p = \cos p / \sin p$, $w = \xi + i\eta$, $w^* = \xi - i\eta$, $\cos p - \cos q = -2 \sin(p+q)/2 \sin(p-q)/2$, and $\sin(p-q) = \sin p \cos q - \sin q \cos p$, in order to obtain:

$$x = ia \frac{\sin i\eta}{\cos \xi - \cos i\eta} .$$

Since $\sin(i\eta) = i \sinh \eta$ and $\cos(i\eta) = \cosh \eta$, this equation leads to (10.2.2a). The computation of (10.2.2b) is completely analogous.

Finally, we compute the constant coordinate curves given in (10.2.3) and (10.2.4). From the trigonometric identity $\tan p = i(1 - e^{2ip})/(1 + e^{2ip})$ given above, we derive easily that

$$2ip = \ln \left(\frac{i - \tan p}{i + \tan p} \right) . \quad (10.5.5)$$

If we now observe that $\tan \alpha_1 = y/(x-a)$ and $\tan \alpha_2 = y/(x+a)$, we obtain upon using (10.5.5) with $p \equiv \alpha_1$, and subsequently, $p \equiv \alpha_2$:

$$\begin{aligned} \xi = \alpha_1 - \alpha_2 &= \frac{i}{2} \left[\ln \left(\frac{i + y/(x-a)}{i - y/(x-a)} \right) - \ln \left(\frac{i + y/(x+a)}{i - y/(x+a)} \right) \right] \\ &= \frac{i}{2} \ln \frac{(ix - ia + y)(ix + ia - y)}{(ix - ia - y)(ix + ia + y)} \\ &= \frac{i}{2} \ln \frac{x^2 + y^2 - a^2 - 2iay}{x^2 + y^2 - a^2 + 2iay} \end{aligned} \quad (10.5.6)$$

After taking the anti-logarithm, and factoring, we find

$$x^2 + y^2 - a^2 + 2iay \left(\frac{1 + e^{2i\xi}}{1 - e^{2i\xi}} \right) = 0 \quad (10.5.7a)$$

or

$$x^2 + y^2 - a^2 - 2ay \cot \xi = 0 . \quad (10.5.7b)$$

With $\csc^2 \xi = 1 + \cot^2 \xi$, we retrieve (10.2.3).

Analogously, with the definition of η , and the Pythagorean theorem, we easily obtain

$$\begin{aligned} e^{2\eta} &= \frac{(d_2)^2}{(d_1)^2} = \frac{(x+a)^2 + y^2}{(x-a)^2 + y^2} \\ &= \frac{x^2 + y^2 + a^2 + 2ax}{x^2 + y^2 + a^2 - 2ax} . \end{aligned} \quad (10.5.8)$$

After some rearranging, (10.2.4) is recovered.

11. The Dynamic Equilibrium of an Ideal Tokamak Plasma

11.1 Introduction

To acquire a good understanding of the relationship between the different flux measures in a tokamak (Ψ_{tor} , Ψ_{pol}^d , Ψ_{pol}^r), a careful investigation of the time evolution of the flux surfaces in an ideal plasma (with $\sigma \rightarrow \infty$) is called for. Plasma and flux surfaces will be shown to move in response to a time-varying transformer flux.

In this Chapter, we do not intend to “prove” relationships, but only to make certain statements plausible. We shall make use of certain ideas that are familiar to most readers, but which have not been discussed in detail in previous chapters, where we referred to this Chapter.

11.2 A Useful Identity in Tokamak Geometry

The representation of the magnetic field in terms of the contravariant-basis vectors ∇u^i is discussed in Sects. 6.1–3. In an axisymmetric tokamak, with symmetry angle ζ_o , the magnetic field can be written (see (6.3.10)) as:

$$\mathbf{B} = I \nabla \zeta_o + \nabla \zeta_o \times \nabla \Psi = \mathbf{B}_T + \mathbf{B}_P . \quad (11.2.1)$$

We assume that the toroidal field \mathbf{B}_T and the poloidal field \mathbf{B}_P are in the positive ζ_o and θ directions, respectively. Ψ is the radial variable and is related to Ψ_{pol}^d by $\Psi = -\Psi_{\text{pol}}^d/2\pi$ (apart from an additive constant which is unimportant here but which will be mentioned at the end of this Chapter); $\nabla \Psi_{\text{pol}}^d$ points radially inward, whereas $\nabla \Psi$ points radially outward. I is a flux-surface quantity equal to $B_T R$ and $\nabla \zeta_o \cdot \nabla \Psi = 0$. As a consequence of this last property, it follows from (11.2.1) that $|\nabla \Psi|/R = B_P$ or $|\nabla \Psi| = RB_P$.

A very useful geometric relationship that we shall use below can be derived from the \mathbf{B} representation in (11.2.1): we wish to decompose the component perpendicular to \mathbf{B} , which is tangent to the flux surface ($\propto \mathbf{B} \times \nabla \Psi$), into a toroidal component ($\propto \nabla \zeta_o$) and a parallel component ($\propto \mathbf{B}$). Take the cross product of $\nabla \Psi$ with (11.2.1):

$$\nabla \Psi \times \mathbf{B} = I(\nabla \Psi \times \nabla \zeta_o) + \nabla \Psi \times (\nabla \zeta_o \times \nabla \Psi) . \quad (11.2.2)$$

The expression between brackets of the first term equals $(I \nabla \zeta_o - \mathbf{B})$. To the last term we apply the *bac-cab* rule and take into account that $\nabla \zeta_o \cdot \nabla \Psi = 0$. We obtain

$$\begin{aligned} \nabla \Psi \times \mathbf{B} &= -IB + I^2 \nabla \zeta_o + \nabla \zeta_o |\nabla \Psi|^2 \\ &= -IB + \nabla \zeta_o (R^2 B_T^2 + R^2 B_P^2) \end{aligned}$$

or

$$\nabla \Psi \times \mathbf{B} = -IB + B^2 R^2 \nabla \zeta_o . \quad (11.2.3)$$

The unit vector perpendicular to \mathbf{B} , but tangent to the flux surface is then

$$\frac{\nabla \Psi \times \hat{\mathbf{B}}}{|\nabla \Psi|} = -\frac{I \hat{\mathbf{B}}}{|\nabla \Psi|} + \frac{BR^2}{|\nabla \Psi|} \nabla \zeta_o . \quad (11.2.4)$$

Recalling that $|\nabla \Psi| = RB_P$,

$$\frac{\nabla \Psi \times \hat{\mathbf{B}}}{RB_P} = -\frac{B_T}{B_P} \hat{\mathbf{B}} + \frac{B}{B_P} \hat{\zeta}_o . \quad (11.2.5)$$

11.3 The Existence of an Electric Field in a Tokamak Plasma with Infinite Conductivity

A time-varying magnetic flux through the core of a transformer “generates” a non-electrostatic electric field $\mathbf{E}^{(A),\text{tr}}$ which is tangent to any circle symmetric about the core of the transformer (see Fig. 11.1). The superscript (A) indicates that we are dealing with an \mathbf{E} due to a $\partial \mathbf{A}/\partial t$ (and not a $\nabla \Phi$). The value of $\mathbf{E}^{(A),\text{tr}}$ at any position can be found from the law of Faraday-Lenz.

Next we consider a torus encircling the transformer core (Fig. 11.2). Suppose that the torus is an ideal conductor, $\sigma \rightarrow \infty$. There cannot exist an electric field in this ideal conductor, and Faraday’s law is now obeyed via the generation of a surface current in response to a $d\Psi^{\text{tr}}/dt$ such that $\Psi_{\text{tot},\text{disk}}$ remains constant. Said differently, the changing current density \mathbf{J} generates a changing poloidal magnetic field whose flux change through the hole of the disk cancels $d\Psi^{\text{tr}}/dt$ exactly. Still another way of looking at it is that the electric field $\mathbf{E}^{(A),\text{tr}}$ due to the transformer is cancelled by an equally large, but reverse electric field $\mathbf{E}^{(A),\text{react}}$ generated by the changing poloidal field so that at any moment $\mathbf{E}^{(A),\text{tot}} \equiv 0$. In terms of loop voltages, $e = |\mathbf{E}^{(A),\text{tr}} 2\pi R|$ equals $|d\Psi^{\text{tr}}/dt|$, whereas, $|\mathbf{E}^{(A),\text{react}} 2\pi R|$ is related to $|L di/dt|$, where L is the self-inductance of the torus ring, and i is the current. Hence, the current i changes proportionally to Ψ^{tr} .

The situation is different in a tokamak plasma. In Fig. 11.3, we consider a typical flux surface containing an ideal plasma at an instant $t = t_0$. At this time $t = t_0$, a spiraling magnetic field is present ($\mathbf{B}_0 = \mathbf{B}_{T_0} + \mathbf{B}_{P_0}$) as well as a non-zero

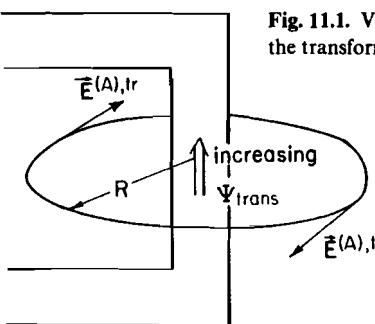


Fig. 11.1. Vacuum electric field $\vec{E}^{(A),tr}$, induced by a changing flux in the transformer: $d\Psi^tr/dt$

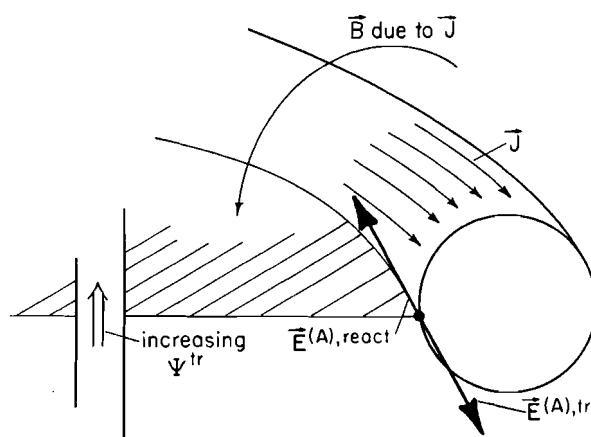


Fig. 11.2. The net electric field in an ideal conductor encircling a transformer is zero: $\vec{E}^{(A),react} = -\vec{E}^{(A),tr}$. See text

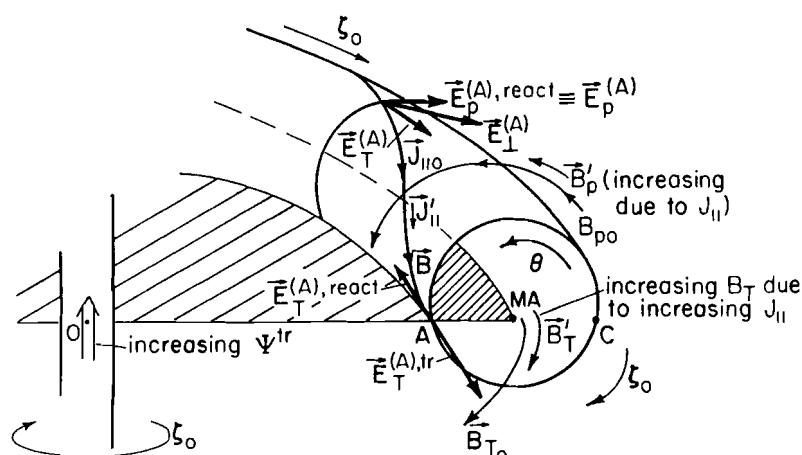


Fig. 11.3. Electric-field components, current-density components and magnetic-field components in an ideal plasma in response to an increasing transformer flux. See text

current density $J_{||0}$. We assume that the transformer flux has been increasing – say, linearly – from $t = 0$ to $t = t_0$, and will continue to do so for $t > t_0$. (The (perpendicular) diamagnetic current which is necessary for equilibrium does not influence our explanation here and is not mentioned henceforth.) The current density $J_{||0}$ is not only a surface current, but must exist inside the plasma to generate the poloidal magnetic field.

The requirement in an ideal plasma is that there *cannot* exist an electric field parallel to \mathbf{B} , since that would cause the current to blow up. In response to a linearly increasing Ψ^tr , the current density $J_{||}$ will also increase linearly, somewhat similarly to the case considered in Fig. 11.2. In simplistic terms, as far as the current response is concerned, the tokamak case that we are dealing with now, could be compared with an ideal conductor in which the current is forced to flow helically; $\sigma_{||} \rightarrow \infty$ along the helix, $\sigma_{\perp} \rightarrow 0$ perpendicular to the helix. (We shall show below, however, that $J_{||}$ not only increases on the flux surface that we isolated, but also further inside the plasma.) The increase in $J_{||}$ during a small time interval Δt is represented by $J'_||$. The increasing current density $J_{||} = J_{||0} + J'_||$ has components in the toroidal and poloidal directions. Its toroidal component opposes the increase of flux through the hole (disk) of the toroidal surface (in much the same way as was the case in Fig. 11.2). The electric-field component “generated” by this changing toroidal component of $J_{||}$ (and thus the changing poloidal field B_p) is labeled by $\vec{E}_T^{(A),react}$. However, because the current density \mathbf{J} is spiraling around, it will not be able to cancel the total electric-field component $\vec{E}_T^{(A),tr}$. (The subscripts T specify that the E-fields are in the toroidal direction.) So we might expect that a net toroidal component of E will survive. This is explained momentarily.

The poloidal component of the increasing $J_{||}$ increases the toroidal magnetic field. The change in \mathbf{B}_T during Δt is denoted by \mathbf{B}'_T . Application of Faraday’s law along a poloidal loop shows that an increasing $\mathbf{B}_T = \mathbf{B}_{T0} + \mathbf{B}'_T$ sets up a poloidal electric field. In Fig. 11.3, this field is indicated as $\vec{E}_P^{(A),react}$, since it follows from the “reaction current” $J_{||}$. Since a poloidal E-field has a parallel component, and no net parallel electric field can be tolerated, there must be another electric-field component that cancels this parallel component of $\vec{E}_P^{(A),react}$. As shown in Fig. 11.4, that field must be toroidal, so that the only surviving electric field is in the

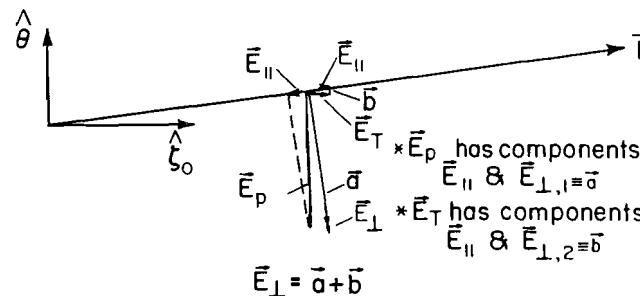


Fig. 11.4. Perpendicular and parallel components of \vec{E}_p and \vec{E}_T ; see text

perpendicular direction. This is exactly the surviving toroidal electric field we alluded to in the former paragraph (see also, Hinton and Hazeltine (1976), p. 252). In contrast thus to an ideal conductor (as in Fig. 11.2), there can be a toroidal electric field component because the magnetic field spirals (as long as the total E is in the perpendicular direction). The fields $E_T^{(A)}$ and $E_p^{(A)}$ are constant in time for a linearly increasing transformer flux. Although we did not say so, they obviously were already present from $t = 0$ to $t = t_0$.

The fact that a perpendicular E field can exist in an ideal plasma should be no surprise, since in a fully ionized plasma, the “perpendicular conductivity” defined as the proportionality constant between current density and electric field is zero. (The perpendicular conductivity found in the literature is the proportionality constant between the perpendicular current and the perpendicular friction force in analogy with the parallel one.) Rather than causing a current, a perpendicular electric field causes an $E \times B$ fluid motion (see Golant, Zhilinsky and Sakharov (1980) and Braginskii (1965)). Note that, even if we had assumed that J_\parallel were only a surface current (in analogy with the ideal conductor), a field $E_\perp^{(A)}$ would still have existed *inside* the plasma.

11.4 Motion of the Plasma and the Flux Surfaces

11.4.1 Plasma Motion due to $E \times B$ Drift

When an electric field is present in a plasma, the latter will move with a fluid velocity $E \times B/B^2$. In the case we are considering here, we find that the plasma *pinches radially inward*. Letting v denote the plasma fluid velocity, we compute the radial “velocity” as follows:

$$v \cdot \nabla \Psi = \nabla \Psi \cdot \frac{E \times B}{B^2} = E \cdot \frac{B \times \nabla \Psi}{B^2}. \quad (11.4.1)$$

By means of the geometric identity of (11.2.3), this gives

$$v \cdot \nabla \Psi = E \cdot B \frac{I}{B^2} - E \cdot \nabla \zeta_o R^2, \quad (11.4.2)$$

or with

$$E_\parallel = E \cdot \hat{B} \text{ and } E_T = E \cdot \hat{\zeta}_o = E \cdot R \nabla \zeta_o,$$

$$v \cdot \nabla \Psi = \frac{I}{B} E_\parallel - R E_T. \quad (11.4.3)$$

In our ideal plasma, the infinite conductivity requires that E_\parallel vanishes, so we obtain

$$v \cdot \nabla \Psi = -R E_T^{(A)}. \quad (11.4.4)$$

In the last expression, the label *(A)* has been added to the electric-field symbol.

$E_T^{(A)}$ is the *net* toroidal electric field, part of which is caused by the transformer, and a (counteracting) part is due to the reaction currents dJ_\parallel/dt . $\nabla \Psi$ is radially outward, and $E_T^{(A)}$ is a positive number (see Fig. 11.3); it thus follows that the plasma pinches *inward*.

11.4.2 Motion of Flux Surfaces Defined by the Poloidal Disk Flux Ψ_{pol}^d

Consider now the flux surfaces labeled by Ψ_{pol}^d , the *poloidal disk flux*. We define the flux surfaces by $\Psi_{\text{pol}}^d = \text{constant}$. This implies that at that surface

$$\frac{d \Psi_{\text{pol}}^d}{dt} = 0. \quad (11.4.5)$$

After expanding d/dt as a convective derivative, we get

$$\left. \frac{\partial \Psi_{\text{pol}}^d}{\partial t} \right|_R + \mathbf{u}_s \cdot \nabla \Psi_{\text{pol}}^d = 0. \quad (11.4.6)$$

Here \mathbf{u}_s is the velocity of a flux surface $\Psi_{\text{pol}}^d = \text{constant}$. Note that only radial flux surface motion makes sense physically.

To find an expression for the partial derivative in (11.4.6), we start from the “radial” contravariant component of Faraday’s law:

$$\nabla \Psi \cdot \frac{\partial \mathbf{B}}{\partial t} = -\nabla \Psi \cdot \nabla \times \mathbf{E}. \quad (11.4.7)$$

Since \mathbf{B} is by definition tangent to a flux surface, we know that $\mathbf{B} \cdot \nabla \Psi \equiv 0$. It then follows from $\partial(\mathbf{B} \cdot \nabla \Psi)/\partial t = 0$ that $\nabla \Psi \cdot \partial \mathbf{B}/\partial t = -\mathbf{B} \cdot \partial \nabla \Psi/\partial t$. Since we are dealing with smooth quantities, we can interchange the ∇ and $\partial/\partial t$ operators to obtain

$$\nabla \Psi \cdot \frac{\partial \mathbf{B}}{\partial t} = -\mathbf{B} \cdot \nabla \frac{\partial \Psi}{\partial t}. \quad (11.4.8)$$

For the right side of (11.4.7), we observe that

$$\nabla u^1 \cdot (\nabla \times \mathbf{E}) \equiv (\nabla \times \mathbf{E})^1 = \frac{1}{\sqrt{g}} \left(\frac{\partial E_3}{\partial u^2} - \frac{\partial E_2}{\partial u^3} \right).$$

Because $u^3 = \zeta_o$ and $\partial/\partial \zeta_o \equiv 0$ as a consequence of axisymmetry, we see that only the E_{ζ_o} (covariant) component is of concern. Thus,

$$\begin{aligned} -\nabla \Psi \cdot (\nabla \times \mathbf{E}) &= -\nabla \Psi \cdot (\nabla \times E_{\zeta_o} \nabla \zeta_o) \\ &= -\nabla \Psi \cdot (\nabla E_{\zeta_o} \times \nabla \zeta_o) \\ &= -(\nabla \zeta_o \times \nabla \Psi) \cdot \nabla E_{\zeta_o} \\ &= -\mathbf{B} \cdot \nabla E_{\zeta_o} \\ &= -\mathbf{B} \cdot \nabla (R E_T) . \end{aligned} \quad (11.4.9)$$

In this derivation we have used the fact that

$$\mathbf{B}_T \cdot \nabla E_{\zeta_0} = I \nabla \zeta_0 \cdot \nabla E_{\zeta_0} = I \nabla \zeta_0 \cdot \left(\frac{\partial E_{\zeta_0}}{\partial \Psi} \nabla \Psi + \frac{\partial E_{\zeta_0}}{\partial \theta} \nabla \theta + \frac{\partial E_{\zeta_0}}{\partial \zeta_0} \nabla \zeta_0 \right) \equiv 0 ,$$

since $\partial/\partial \zeta_0 = 0$ and $\nabla \zeta_0 \cdot \nabla \Psi = 0 = \nabla \zeta_0 \cdot \nabla \theta$, all because of axisymmetry. We employed $E = E_T \hat{\zeta}_0 = E_{\zeta_0} \nabla \zeta_0$, where $\nabla \zeta_0 = \hat{\zeta}_0/R$. Faraday's law is then, from (11.4.8) and (11.4.9),

$$\mathbf{B} \cdot \nabla \left(-\frac{\partial \Psi}{\partial t} + RE_T \right) = 0 . \quad (11.4.10)$$

This means that

$$\frac{\partial \Psi}{\partial t} = RE_T + \text{flux function} . \quad (11.4.11)$$

With $\Psi \equiv -\Psi_{\text{pol}}^d/2\pi$, and with $E_T \equiv E_T^{(A)}$, the net electric field, we see that the flux function can be chosen to be zero. (This follows from the integral form of Ampère's law where a ζ_0 loop is taken through point A of Fig. 11.3.) The time rate of change of the poloidal flux through a disk bounded by a fixed ζ_0 loop is thus

$$\frac{\partial \Psi_{\text{pol}}^d}{\partial t} \Big|_{\text{disk}} = -2\pi R E_T^{(A)} . \quad (11.4.12)$$

After substitution of this result in (11.4.6), the latter can be solved for the radial flux-surface "velocity":

$$\mathbf{u}_s^d \cdot \nabla \Psi_{\text{pol}}^d = -\frac{\partial \Psi_{\text{pol}}^d}{\partial t} \Big|_{\text{d}} = 2\pi R E_T^{(A)} . \quad (11.4.13)$$

(In \mathbf{u}_s^d , s stands for surface and d for disk.) Taking into account that $\Psi \equiv -\Psi_{\text{pol}}^d/2\pi$, we observe from (11.4.4) and (11.4.13) that the plasma $\mathbf{E} \times \mathbf{B}$ inward pinch velocity is exactly equal to the radial velocity of the flux surfaces. This is a statement of the frozen flux theorem: plasma and flux surfaces are tied together.

An observer sitting on (and moving with) the plasma will find that the fluid-pinch motion generates a motional EMF. The electric field connected with this motional EMF is exactly the negative of the electric field that caused the $\mathbf{E} \times \mathbf{B}$ inward motion in the first place. This is as it should be: an observer moving with the plasma, or thus the plasma itself, does not see any electric field. The electric field is only present in the lab frame.

11.4.3 Motion of Flux Surfaces Defined by the Poloidal Ribbon Flux Ψ_{pol}^r

Suppose now that we wish to label a flux surface by the *poloidal ribbon flux* (see Fig. 11.3). To actually show that the flux surfaces labeled by $\Psi_{\text{pol}}^r = \text{constant}$ move with the same radially inward velocity as the plasma, we return to (11.4.10) or (11.4.11). When we now identify $\Psi = \Psi_{\text{pol}}^r/2\pi$, we must recognize that the

ribbon flux does not know about the transformer; all it contains is the "reaction flux" due to an increase of B_p and $(J_{\parallel})_T$. The integration constant that we called the "flux function" in (11.4.11) is the ubiquitous space independent transformer loop voltage (devided by 2π) $e(t)/2\pi = RE^{(A),\text{tr}} = \oint \mathbf{E} \cdot d\mathbf{l}/2\pi = |d\Psi^{\text{tr}}/dt|/2\pi$ that we encountered in Fig. 11.1 ($e(t)$ is spatially uniform, but $E^{(A),\text{tr}} \propto 1/R$). Instead of (11.4.12), we obtain for the case of the ribbon flux:

$$\begin{aligned} \frac{\partial \Psi_{\text{pol}}^r}{\partial t} \Big|_{\text{fixed ribbon}} &= +2\pi R E_T^{(A),\text{react}} + e(t) \\ &= 2\pi R [+E_T^{(A),\text{react}} + E_T^{(A),\text{tr}}] . \end{aligned} \quad (11.4.14)$$

The appropriate signs can be understood from (11.4.12): $\Psi = \Psi_{\text{pol}}^r/2\pi$ and $E_T^{(A),\text{react}} = E_T^{(A),\text{react}} \hat{\zeta}_0 = -|E_T^{(A),\text{react}}| \hat{\zeta}_0$; thus $E_T^{(A),\text{react}}$ is a negative quantity (see Fig. 11.3).

With the flux surfaces, $\Psi_{\text{pol}}^r = \text{constant}$ defined in a similar way to (11.4.5), we find for their radial "velocity"

$$\mathbf{u}_s^r \cdot \nabla \Psi_{\text{pol}}^r = -\frac{\partial \Psi_{\text{pol}}^r}{\partial t} \Big|_{\mathbf{R}} = -2\pi R [E_T^{(A),\text{react}} + E_T^{(A),\text{tr}}] . \quad (11.4.15)$$

From Fig. 11.3, it is clear that $E_T^{(A),\text{tr}} + E_T^{(A),\text{react}} = E_T^{(A),\text{tr}} - |E_T^{(A),\text{react}}| = E_T^{(A)}$, where $E_T^{(A)}$ is the magnitude of the net toroidal electric field ($E_T^{(A)} > 0$). Thus

$$\mathbf{u}_s^r \cdot \nabla \Psi_{\text{pol}}^r = -2\pi R E_T^{(A)} . \quad (11.4.16)$$

A comparison with (11.4.4) and (11.4.13) shows that, indeed,

$$\mathbf{v} \cdot \nabla \Psi = \mathbf{u}_s^d \cdot \nabla \Psi = \mathbf{u}_s^r \cdot \nabla \Psi . \quad (11.4.17)$$

Thus plasma and flux surfaces labeled by a constant poloidal (ribbon or disk) flux move together.

11.4.4 Motion of Flux Surfaces Defined by the Toroidal Flux Ψ_{tor}

Finally, we show that the *toroidal-flux label* Ψ_{tor} leads to a similar conclusion. We could invoke the general applicability of the frozen-flux theorem: the plasma and the flux surfaces (regardless of their label) move together. Simple manipulations with formulae, however, provide some physical insight: the toroidal-flux change should be connected with the *poloidal* electric field.

To find the inward $\mathbf{E} \times \mathbf{B}$ plasma velocity, we first derive a geometric relationship for the component perpendicular to \mathbf{B} in terms of poloidal and toroidal components. From (11.2.2) we obtain:

$$\begin{aligned} \frac{\nabla \Psi \times \hat{\mathbf{B}}}{|\nabla \Psi|} &= \frac{|\nabla \Psi|}{B} \nabla \zeta_0 - \frac{I}{B |\nabla \Psi|} (\nabla \zeta_0 \times \nabla \Psi) \\ &= \frac{|\nabla \Psi|}{RB} \hat{\zeta}_0 - \frac{B_T}{B} \left(\hat{\zeta}_0 \times \frac{\nabla \Psi}{|\nabla \Psi|} \right) . \end{aligned} \quad (11.4.18)$$

The only induced \mathbf{E} -field component present is in the perpendicular direction

and points in the $+(\nabla\Psi \times \hat{\mathbf{B}})$ direction. The $\mathbf{E} \times \mathbf{B}$ fluid flow is

$$\frac{\mathbf{E}^{(4)} \times \mathbf{B}}{B^2} = \frac{\mathbf{E}_\perp^{(4)} \times \mathbf{B}}{B^2} = \mathbf{E}_\perp^{(4)} \frac{(\nabla\Psi \times \hat{\mathbf{B}})}{|\nabla\Psi|} \times \frac{\hat{\mathbf{B}}}{B} = -\frac{\mathbf{E}_\perp^{(4)}}{B} \frac{\nabla\Psi}{|\nabla\Psi|}. \quad (11.4.19)$$

After multiplication of (11.4.18) with $\mathbf{E}_\perp^{(4)}$, we find that the poloidal component of the electric field is given by

$$\mathbf{E}_P^{(4)} = -\frac{B_T}{B} \mathbf{E}_\perp^{(4)}. \quad (11.4.20)$$

With our conventions $\mathbf{E}_\perp^{(4)}$ is positive (see also Fig. 11.4). The radial fluid velocity, in terms of the poloidal electric field is thus

$$\mathbf{v} \cdot \nabla\Psi_{\text{tor}} = \mathbf{E}_P^{(4)} \frac{|\nabla\Psi_{\text{tor}}|}{B_T}. \quad (11.4.21)$$

In going from (11.4.18) and (11.4.19) to (11.4.21), we have converted Ψ into $\Psi_{\text{tor}}/2\pi$. (We could have written down this result immediately from (11.4.4); multiply its RHS by B_P/B_P and after replacing RB_P by $|\nabla\Psi|$, set $E_T^{(4)} = -(B_P/B_T)\mathbf{E}_P^{(4)}$.) Note that $\nabla\Psi$ in the relations (11.2.1), (11.2.3), (11.4.4) and (11.4.18), in fact, is the gradient of a poloidal flux (which is why we could set $|\nabla\Psi| = RB_P$). However, $\nabla\Psi_{\text{tor}} \propto \nabla\Psi_{\text{pol}}$, and the proportionality constant cancels in (11.4.21).

For the motion of the flux surfaces, we go back to (11.4.7), (11.4.8) and (11.4.9). Because $\mathbf{B} \cdot \nabla\Psi_{\text{tor}} \equiv 0$, (11.4.8) holds for $\Psi = \Psi_{\text{tor}}/2\pi$. Noting that $\tau = |\nabla\Psi_{\text{pol}}|/|\nabla\Psi_{\text{tor}}|$, the magnetic field \mathbf{B} , in terms of $\nabla\Psi_{\text{tor}}$ equals

$$\mathbf{B} = I \nabla\zeta_o + \frac{\tau}{2\pi} (\nabla\zeta_o \times \nabla\Psi_{\text{tor}}). \quad (11.4.22)$$

If we go through the exercise of (11.4.9) for Ψ now replaced by $\Psi_{\text{tor}}/2\pi$, and replace $(\nabla\zeta_o \times \nabla\Psi_{\text{tor}})$ by $2\pi(\mathbf{B} - I \nabla\zeta_o)/\tau$, we obtain

$$-\nabla\Psi_{\text{tor}} \cdot (\nabla \times \mathbf{E}) = -\frac{\mathbf{B}}{\tau} \cdot \nabla(2\pi R E_T). \quad (11.4.23)$$

The combination of the adjusted (11.4.8) and (11.4.23) leads to

$$\mathbf{B} \cdot \nabla \left(\frac{\partial \Psi_{\text{tor}}}{\partial t} - 2\pi \frac{R E_T}{\tau} \right) \equiv 0. \quad (11.4.24)$$

Although the poloidal electric field component is completely generated by “reaction” currents, it is related to the net toroidal \mathbf{E} -field via $E_P^{(4)} = -(B_T/B_P)E_T^{(4)}$. Equation (11.4.24) holds for E_T equal to both $E_T^{(4)}$ and $E_T^{(4),\text{react}}$ because of the operator $\mathbf{B} \cdot \nabla$. It gives, upon integration,

$$\frac{\partial \Psi_{\text{tor}}}{\partial t} = -\frac{2\pi R B_P}{B_T \tau} E_P^{(4)}. \quad (11.4.25)$$

For circular (idealized) flux surfaces, $\tau = 1/q = RB_P/rB_T$, where r is the minor radius. Thus with $RB_P/B_T \tau = r$, it is clear that (11.4.25) is in agreement with intuition.

The flux surface motion follows from $d\Psi_{\text{tor}}/dt \equiv 0$, expanded via the convective derivative,

$$\begin{aligned} \mathbf{U}_s \cdot \nabla\Psi_{\text{tor}} &= \frac{2\pi R B_P}{\tau B_T} E_P^{(4)} = \frac{|\nabla\Psi_{\text{pol}}^{\text{d,r}}|}{\tau} \frac{E_P^{(4)}}{B_T} \\ \mathbf{U}_s \cdot \nabla\Psi_{\text{tor}} &= |\nabla\Psi_{\text{tor}}| \frac{E_P^{(4)}}{B_T}. \end{aligned} \quad (11.4.26)$$

Again, the flux-surface motion (11.4.26) equals the radial plasma motion given in (11.4.21). For mnemonic purposes, one should think of $|\nabla\Psi_{\text{tor}}|/B_T \sim 2\pi r$.

The physical picture is consistent. Plasma and flux surfaces, regardless of their labeling move radially inward together. By shrinking, the flux surfaces provide smaller areas for their (constant) contained fluxes. Consequently, both B_P and B_T increase (as we argued in Fig. 11.3). The increase of B_P within the plasma is the reason why the parallel current must also increase inside the plasma and not only on the surface. (Recall that a toroidal current on a toroidal surface of radius r_o influences only the magnetic field on the outside, for $r > r_o$, not that on the inside.)

11.5 Constancy of the Rotational-Transform Flux Function

From our discussion above, we learned that the magnetic fluxes within the flux surfaces remain constant and therefore that the concept of flux surfaces as flux tubes retains its physical meaning. This also implies that the poloidal and toroidal fluxes for an infinitesimal shell of flux volume remain constant. The toroidal and poloidal components of the magnetic field increase proportionally so that τ remains constant. Because of potential (notational) confusion when interpreting the literature, we shall now demonstrate the constancy of τ .

First, we need a relationship for the rotational transform in terms of the quantities appearing in the \mathbf{B} -field expression, (11.2.1). By definition, the rotational transform is

$$\tau = \frac{\delta \Psi_{\text{pol}}}{\delta \Psi_{\text{tor}}} = \frac{\nabla\Psi_{\text{pol}} \cdot d\mathbf{R}}{\nabla\Psi_{\text{tor}} \cdot d\mathbf{R}}. \quad (11.5.1)$$

Equation (11.5.1) stresses that τ is related to the spatial variations of Ψ_{pol} and Ψ_{tor} . The vector $d\mathbf{R}$ is an arbitrary differential position vector. The toroidal flux Ψ_{tor} equals

$$\begin{aligned}\Psi_{\text{tor}} &= \frac{1}{2\pi} \iiint_V d^3R \mathbf{B} \cdot \nabla \zeta_o \\ &= \frac{1}{2\pi} \iiint_V d^3R I |\nabla \zeta_o|^2 \\ &= \frac{1}{2\pi} \iiint_V d^3R \frac{I}{R^2}.\end{aligned}\quad (11.5.2)$$

The differential toroidal flux $\delta \Psi_{\text{tor}}$ in terms of the poloidal ribbon flux label Ψ , i.e., $\delta \Psi = \delta \Psi_{\text{pol}}/2\pi$, is then

$$\begin{aligned}\delta \Psi_{\text{tor}} &= \frac{\delta \Psi}{2\pi} \int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta_o \sqrt{g_o} \frac{I}{R^2} \\ &= \frac{\delta \Psi}{2\pi} I \int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta_o \sqrt{g_o} \left(\frac{1}{R^2} \right).\end{aligned}\quad (11.5.3)$$

Recall that $I \equiv I(\Psi)$. Since the flux-surface average of a quantity Φ is defined as

$$\langle \Phi \rangle \equiv \frac{\int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta_o \sqrt{g_o} \Phi}{\int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta_o \sqrt{g_o}} = \frac{\int_0^{2\pi} \int_0^{2\pi} d\theta d\zeta_o \sqrt{g_o} \Phi}{\delta V / \delta \Psi}, \quad (11.5.4)$$

we obtain for (11.5.3):

$$\frac{\delta \Psi_{\text{tor}}}{\delta \Psi} = \frac{1}{2\pi} \frac{\delta V}{\delta \Psi} I \langle R^{-2} \rangle. \quad (11.5.5)$$

With $q \equiv 1/t = \delta \Psi_{\text{tor}}/(2\pi \delta \Psi)$, we get for the safety factor:

$$q \equiv \frac{1}{t} = \frac{1}{4\pi^2} \frac{\delta V}{\delta \Psi} \langle IR^{-2} \rangle = \frac{I}{4\pi^2} \frac{\delta V}{\delta \Psi} \left\langle \frac{1}{R^2} \right\rangle. \quad (11.5.6)$$

The time rate of change of the safety factor q on a (moving) constant- Ψ flux surface can be related to its change at a fixed point R in space via the chain rule:

$$\frac{\partial q}{\partial t} \Big|_{\Psi} = \frac{\partial q}{\partial t} \Big|_R - \frac{\partial q}{\partial \Psi} \Big|_t \frac{\partial \Psi}{\partial t} \Big|_R = \frac{\partial q}{\partial t} \Big|_R - \frac{\partial}{\partial \Psi} \Big|_t \left[q \frac{\partial \Psi}{\partial t} \Big|_R \right]. \quad (11.5.7)$$

The last form follows from changing the order of differentiation in $\partial/\partial \Psi|_t, \partial \Psi/\partial t|_R$ which makes this vanish. We now substitute the expression for q , given in (11.5.6) into the right side of (11.5.7). When we take into account that

$$\frac{\partial}{\partial t} \Big|_R \frac{\delta V}{\delta \Psi} = \frac{\delta}{\delta \Psi} \frac{\delta V}{\partial t} \Big|_R \equiv 0, \quad (11.5.8)$$

because the volume of a flux surface through a fixed point R does not change, we obtain:

$$\frac{\partial q}{\partial t} \Big|_{\Psi} = \frac{1}{4\pi^2} \left\{ \frac{\delta V}{\delta \Psi} \left\langle \frac{\partial IR^{-2}}{\partial t} \Big|_R \right\rangle - \frac{\partial}{\partial \Psi} \Big|_t \left[\frac{\delta V}{\delta \Psi} \left\langle \frac{I}{R^2} \frac{\partial \Psi}{\partial t} \Big|_R \right\rangle \right] \right\} \quad (11.5.9)$$

For the second term, we shall substitute (11.4.14). The first term can be obtained from the toroidal component of Faraday's law as follows:

$$\left\langle \nabla \zeta_o \cdot \frac{\partial \mathbf{B}}{\partial t} \Big|_R \right\rangle = - \langle \nabla \times \mathbf{E} \cdot \nabla \zeta_o \rangle, \quad (11.5.10)$$

and thus

$$\left\langle \frac{\partial}{\partial t} \Big|_R \nabla \zeta_o \cdot \mathbf{B} \right\rangle = - \langle \nabla \cdot (\mathbf{E} \times \nabla \zeta_o) \rangle. \quad (11.5.11)$$

In Sect. 4.9.3. we prove the relationship:

$$\langle \nabla \cdot \mathbf{W} \rangle = \frac{1}{\delta V / \delta \Psi} \frac{\partial}{\partial \Psi} \Big|_t \left(\frac{\delta V}{\delta \Psi} \langle \mathbf{W} \cdot \nabla \Psi \rangle \right). \quad (11.5.12)$$

Therefore, (11.5.11) can be rewritten as:

$$\begin{aligned}\left\langle \frac{\partial}{\partial t} \Big|_R (IR^{-2}) \right\rangle &= - \frac{1}{\delta V / \delta \Psi} \frac{\partial}{\partial \Psi} \Big|_t \left[\frac{\delta V}{\delta \Psi} \langle \nabla \Psi \cdot \mathbf{E} \times \nabla \zeta_o \rangle \right] \\ &= - \frac{1}{\delta V / \delta \Psi} \frac{\partial}{\partial \Psi} \Big|_t \left[\frac{\delta V}{\delta \Psi} \langle \mathbf{E} \cdot \nabla \zeta_o \times \nabla \Psi \rangle \right]\end{aligned}\quad (11.5.13)$$

The quantity $\mathbf{E} \cdot \nabla \zeta_o \times \nabla \Psi$ is proportional to the magnitude of the poloidal electric field component (in Fig. 11.3). With $|\nabla \Psi| = RB_p$ and $|\nabla \zeta_o| = 1/R$, we have that $\mathbf{E} \cdot \nabla \zeta_o \times \nabla \Psi = B_p E_p^{(4)} \equiv B_p E_p^{(4), \text{react}}$. (Recall that $E_p^{(4)} = -|E_p^{(4)}|$.) After substitution in (11.5.9), we obtain

$$\begin{aligned}\frac{\partial q}{\partial t} \Big|_{\Psi} &= \frac{1}{4\pi^2} \left\{ \frac{\partial}{\partial \Psi} \Big|_t \left[- \frac{\delta V}{\delta \Psi} \langle B_p E_p^{(4), \text{react}} \rangle \right] \right. \\ &\quad \left. - \frac{\partial}{\partial \Psi} \Big|_t \left[\frac{\delta V}{\delta \Psi} \left\langle \frac{I}{R} (E_T^{(4), \text{tr}} + E_T^{(4), \text{react}}) \right\rangle \right] \right\}.\end{aligned}\quad (11.5.14)$$

Below (11.4.15) we saw that $E_T^{(4), \text{tr}} + E_T^{(4), \text{react}} = E_T^{(4)}$ (see also Fig. 11.3), and below (11.4.21), we stated that $E_T^{(4)} = -(B_p/B_T) E_p^{(4)}$. Taking into account that $I = RB_T$ then leads to

$$\frac{\partial q}{\partial t} \Big|_{\Psi} = \frac{1}{4\pi^2} \left\{ \frac{\partial}{\partial \Psi} \Big|_t \left[\frac{\delta V}{\delta \Psi} \langle -B_p E_p^{(4), \text{react}} \rangle \right] + \frac{\partial}{\partial \Psi} \Big|_t \left[\frac{\delta V}{\delta \Psi} \langle B_p E_p^{(4)} \rangle \right] \right\}. \quad (11.5.15)$$

Or, since $E_p^{(A)} \equiv E_p^{(A),\text{react}}$,

$$\frac{\partial q}{\partial t} \Big|_{\Psi} \equiv 0 . \quad (11.5.16)$$

as argued intuitively above.

The reader might get a false impression when consulting the literature. In most cases, the above analysis is performed for a finite-conductivity plasma (see, e.g., Stacey (1981) or Hirshman and Sigmar (1981)). Instead of identifying $\mathbf{E} \cdot \nabla \zeta_o \times \nabla \Psi$ in (11.5.13) with $B_p E_p^{(A)}$, we could have replaced $\nabla \zeta_o \times \nabla \Psi$ by $(\mathbf{B} - I \nabla \zeta_o)$. Then (11.5.14) would have been replaced by

$$\frac{\partial q}{\partial t} \Big|_{\Psi} = \frac{1}{4\pi^2} \left\{ - \frac{\partial}{\partial \Psi} \left[\frac{\delta V}{\delta \Psi} \langle -E_{||}^{(A),\text{react}} B \rangle + \frac{\delta V}{\delta \Psi} \left\langle IR^{-2} \frac{e(t)}{2\pi} \right\rangle \right] \right\} . \quad (11.5.17)$$

Now, for instance, Stacey (1981) does not clarify the term $E_{||}^{(A)}$ further, although it is understood that it is our $E_{||}^{(A),\text{react}}$. One might get the impression that only his $\langle E_{||}^{(A)} B \rangle$ term must be set to zero as $\sigma \rightarrow \infty$, while the “ubiquitous” $e(t)$ survives, erroneously leading to a nonzero $\partial q/\partial t|_{\Psi}$. When interpreting (11.5.17), we must realize that the *total* parallel electric field must vanish. Since we have that

$$\begin{aligned} \frac{e(t)}{2\pi} &= RE_T^{(A),\text{tr}} = RE_T^{(A),\text{tr}} \cdot \hat{\zeta}_o = R^2 E_T^{(A),\text{tr}} \cdot \nabla \zeta_o \\ &= R^2 E_T^{(A),\text{tr}} \cdot (\mathbf{B} - \nabla \zeta_o \times \nabla \Psi)/I = (R^2/I) E_{||}^{(A),\text{tr}} B , \end{aligned}$$

the parallel E -field contribution of the transformer voltage must vanish as well in our ideal plasma.

11.6 Remarks on the Evolution of a Finite-Resistivity Plasma

When a *finite-resistivity plasma* ($\sigma \neq \infty$) is considered, the frozen-flux theorem no longer holds, and the plasma and flux surfaces “leak through” each other: the flux surfaces move inward faster than the plasma (the flux surfaces still obey (11.4.13), whereas the plasma follows (11.4.3)). Note, however, that in the tokamak neoclassical banana regime, the radial inward pinch velocity, now called the neoclassical or Ware pinch, approximately equals the radially inward velocity of the flux surfaces. Thus even in the presence of collisions and therefore resistivity, the plasma and flux surfaces can remain tied together. For details, see Hinton and Hazeltine (1976, Chap. 3). Also, in a finite resistivity plasma, the parallel current does not keep increasing but saturates at a value equal to $\sigma E_{||}^{(A)}$. This can be understood qualitatively from an analogy with a circular conductor as in Fig. 11.2. When $\sigma \rightarrow \infty$, the current follows the flux change $|d\Psi/dt| = L|di/dt|$, where L is the self-inductance. When σ is finite, the current levels off to the above-mentioned value as $(1 - e^{-t/\tau})$ with the time constant $\tau = L/R$.

A more detailed discussion of the plasma dynamics in resistive plasmas is beyond the scope of this book. To do a proper quantitative analysis, the equilibrium and transport equations must be considered simultaneously, and the problem must be solved self-consistently. Interesting considerations concerning classical and neoclassical transport are discussed in Hinton and Hazeltine (1976) and Hirshman and Sigmar (1981); see also, the other references given in Chap. 4, above (4.7.7).

11.7 The Relationship Between Poloidal Disk and Ribbon Fluxes

Having considered the motion of an ideal plasma and flux surfaces embedded in it, we may say a few final words about the relationship between the poloidal disk and ribbon fluxes. In principle, the expression for the magnetic field given in (11.2.1), $\mathbf{B} = I \nabla \zeta_o + \nabla \zeta_o \times \nabla \Psi$, makes most physical sense for Ψ proportional to the poloidal flux inside the plasma:

$$\Psi \equiv \frac{\Psi_{\text{pol}}^r}{2\pi} .$$

From Fig. 11.3, we see that the poloidal disk flux is related to the ribbon flux as follows:

$$\Psi_{\text{pol}}^d = \Psi_{\text{pol}}^{\text{total}} - \Psi_{\text{pol}}^r . \quad (11.7.1)$$

The total poloidal flux, represented by $\Psi_{\text{pol}}^{\text{total}}$ is the flux through a disk bounded by the magnetic axis. It contains the transformer flux and the flux generated by the plasma currents.

Above we showed that the time rate of change of the poloidal ribbon flux and the poloidal disk flux through the ribbon and disk, respectively, constructed through a *fixed* point A (on the line $O-MA$, where MA stands for magnetic axis) are exactly opposite. Compare (11.4.12) and (11.4.14). This implies that the total poloidal flux remains constant. From the *moving* flux surface picture, we can draw the same conclusions: poloidal disk and ribbon fluxes are each conserved through their enlarging and shrinking areas, respectively, again establishing that the poloidal flux is constant. The above points out that

$$\frac{\partial \Psi_{\text{pol}}^{\text{total}}}{\partial t} \equiv 0 . \quad (11.7.2)$$

Because $\Psi_{\text{pol}}^{\text{total}}$ is merely a constant, it was not mentioned previously. Indeed, all we are concerned with are *derivatives* of fluxes. It did not appear in the time derivatives because of (11.7.2), nor in the gradient expressions, since $\nabla \Psi_{\text{pol}}^{\text{total}} = 0$. For this reason, we could identify $\Psi \equiv \Psi_{\text{pol}}^r/2\pi$ in (11.2.1) with $-\Psi_{\text{pol}}^d/2\pi$ and not as in (11.7.1). We must have $\nabla \Psi = \nabla \Psi_{\text{pol}}^r/2\pi = -\nabla \Psi_{\text{pol}}^d/2\pi$, and so we may ignore the constant $\Psi_{\text{pol}}^{\text{total}}$.

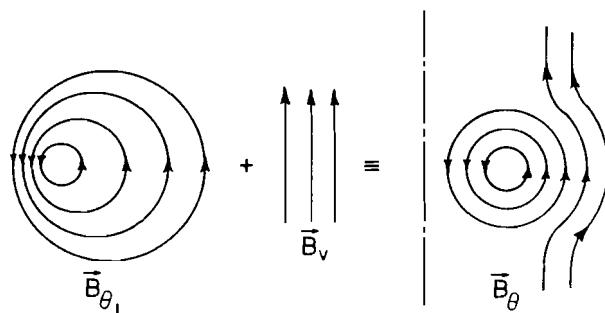


Fig. 11.5. The net poloidal magnetic field \vec{B}_{θ} is the sum of the vertical field \vec{B}_v and the field \vec{B}_{θ_1} .

In our discussion above, the vertical magnetic-field component, which is needed for equilibrium with respect to horizontal motion, has not been mentioned. This is because the vertical field is merely part of the net poloidal field in the sense that it reduces the original poloidal field (due to the toroidal plasma current component) on the inside of the torus and strengthens it on the outside (see Fig. 11.5 for an idealized case). It is the *net* poloidal field including the vertical field that determines the position and the shape of the flux surfaces, which are “centered” – not necessarily concentrically – about the magnetic axis. This is in fact the position where the *net* vertical field vanishes.

The vertical field is varied in order to achieve a desired position of the plasma column. The required field has been calculated for tokamaks by Shafranov (1966), Shafranov and Zakharov (1972), and Greene, Johnson and Weimer (1971). See also Zakharov and Shafranov (1986). Formulae analogous to those of Greene, Johnson and Weimer (1971) were obtained for a stellarator by Fielding and Hitchon (1980). For elementary textbook discussions on the necessity of a vertical or “maintaining” field, see e.g., Bateman (1978); Golant, Zhilinsky and Sakharov (1980) or Stacey (1981). From the point of view of this discussion, the vertical field is not distinguishable from the other poloidal fluxes; the total poloidal field is still “frozen in”.

In this Chapter we have idealized and neglected many influences concerning the vertical or “maintaining” field. For an overview of some of these potentially important effects, see the review papers by Mukhovatov and Shafranov (1971) and Zakharov and Shafranov (1986).

In most discussions concerning flux coordinates, this dynamic analysis is of secondary importance. One pretends that the flux surfaces are stationary and does not mention the time dependence. When we are interested in only the spatial dependences, time is merely a parameter, and we can establish the geometrical relationships at any time t .

12. The Relationship Between $\int dl/B$ and $dV/d\Psi_{\text{tor}}$

In many elementary discussions it is “proved” in an intuitive manner that the specific volume $\dot{V}(\Psi_{\text{tor}}) \equiv dV/d\Psi_{\text{tor}}$ equals the closed line integral $\oint dl/B$. That this is true only in an approximate sense, and that a distinction is to be made between rational and irrational surfaces is discussed in the authoritative review paper by Solov'ev and Shafranov (1970). For completeness, we reproduce their arguments in this Chapter albeit converted to our notation.

To evaluate $\oint dl/B$, we make use of the equation of a field line in flux coordinates

$$\frac{dl}{B} = \frac{d\zeta_f}{\mathbf{B} \cdot \nabla \zeta_f} = \frac{d\zeta_f}{B^{\zeta_f}} = d\zeta_f \left(\frac{2\pi\sqrt{g_f}}{\dot{\Psi}_{\text{tor}}} \right). \quad (12.1.1)$$

Recall that $\dot{\Psi}_{\text{tor}} = d\Psi_{\text{tor}}/d\varrho$. After multiplication in the numerator and the denominator by $dV/d\varrho$, we have

$$\int \frac{dl}{B} = \frac{dV}{d\Psi_{\text{tor}}} 2\pi \frac{\int \sqrt{g_f} d\zeta_f}{dV/d\varrho}. \quad (12.1.2)$$

In Chap. 4, we explained that

$$\frac{dV}{d\varrho} = \int_0^{2\pi} \int_0^{2\pi} \sqrt{g_f} d\theta_f d\zeta_f. \quad (12.1.3)$$

If $\sqrt{g_f}$ is expanded in a double Fourier series in θ_f and ζ_f as

$$\sqrt{g_f} = \sum_{m,n} g_{mn}^f(\varrho) e^{i(m\theta_f - n\zeta_f)}, \quad (12.1.4)$$

we obtain after substitution and integration in (12.1.3)

$$\frac{dV}{d\varrho} = 4\pi^2 g_{00}^f(\varrho). \quad (12.1.5)$$

The integral of (12.1.2) becomes, upon Fourier expansion,

$$\begin{aligned} \int \sqrt{g_f} d\zeta_f &= \int \sum_{m,n} g_{mn}^f(\varrho) e^{i(m\theta_f - n\zeta_f)} d\zeta_f \\ &= g_{00}^f(\varrho) \zeta_f + \int \sum'_{m,n} g_{mn}^f(\varrho) e^{i(m\theta_f - n\zeta_f)} d\zeta_f \\ &= \frac{1}{4\pi^2} \frac{dV}{d\Psi_{\text{tor}}} \zeta_f + \int \sum'_{m,n} g_{mn}^f(\varrho) e^{i(m\theta_f - n\zeta_f)} d\zeta_f . \end{aligned} \quad (12.1.6)$$

The prime on the summation sign indicates that the term $m = n = 0$ must be omitted. Combination of (12.1.2), (12.1.5) and (12.1.6) results in

$$\int \frac{dl}{B} = \frac{dV}{d\Psi_{\text{tor}}} \left\{ \frac{1}{2\pi} \zeta_f + \frac{1}{2\pi g_{00}^f(\varrho)} \int \sum'_{m,n} g_{mn}^f(\varrho) e^{i(m\theta_f - n\zeta_f)} d\zeta_f \right\} . \quad (12.1.7)$$

We now define what is called the “proper length” of a field line:

$$U(\varrho, \theta_{f0}) \equiv \frac{1}{N} \int_0^{2\pi N} \frac{d\zeta_f}{B^{\zeta_f}} . \quad (12.1.8)$$

This integral is taken along a field line $\theta_{f0} = \theta_f - t\zeta_f$, from $\zeta_f = 0$, to $\zeta_f = 2\pi N$. After performing the integral in (12.1.7), we have thus

$$U(\varrho, \theta_{f0}) = \frac{dV}{d\Psi_{\text{tor}}} \left\{ 1 + \frac{1}{2\pi N g_{00}^f(\varrho)} \sum'_{m,n} g_{mn}^f(\varrho) e^{im\theta_{f0}} \left[\frac{e^{i(mt-n)2\pi N} - 1}{i(mt-n)} \right] \right\} . \quad (12.1.9)$$

When the field line lies on a rational surface, it closes upon itself after, say, n_o circuits of the minor axis and m_o toroidal circuits. For that field line,

$$t(\varrho) = \frac{n_o}{m_o} . \quad (12.1.10)$$

For this type of field line, $N \equiv m_o$. Making use of the fact that

$$\lim_{x \rightarrow 0} \frac{e^{ixa} - 1}{ix} = a , \quad (12.1.11)$$

we have then with $a = 2\pi N$:

$$\begin{aligned} U(\varrho, \theta_{f0}) &= \frac{1}{m_o} \oint \frac{dl}{B} \\ &= \frac{dV}{d\Psi_{\text{tor}}} \left\{ 1 + \frac{1}{g_{00}^f(\varrho)} [g_{m_o n_o} e^{im_o \theta_{f0}} + g_{-m_o, -n_o} e^{-im_o \theta_{f0}}] \right\} . \end{aligned} \quad (12.1.12)$$

From this expression, it is clear that the integral $\oint dl/B$ taken along different field lines on the same rational magnetic surface will, in general, assume different values, depending on the value of the field-line label θ_{f0} . (Recall that θ_{f0} is equivalent to v or β in the Clebsch representation). In fact, depending on the

particular field-line label, U will oscillate about $dV/d\Psi_{\text{tor}}$. The fact that $\oint dl/B$ is not always constant on a rational surface has serious consequences as was mentioned above.

For the case of irrational surfaces, we observe that $\oint dl/B$ integrated from $\zeta_f = 0$ to $\zeta_f = 2\pi$ should not be very different from the value of $U(\varrho, \theta_{f0})$ on the nearby rational surface characterized by $t = n_o/m_o$. For a large number of circuits, $N \rightarrow \infty$, the function $U(\varrho, \theta_{f0})$ is averaged, with the result:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \int \frac{dl}{B} = \frac{dV}{d\Psi_{\text{tor}}} . \quad (12.1.13)$$

13. Transformation Properties of Vector and Tensor Components

In this Chapter we wish to investigate how the *components* of vectors and tensors change when another coordinate system is chosen. We want to find out what happens to the components under a general transformation of the form $u^i \equiv u^i(u^1, u^2, u^3)$ for $i = 1, 2, 3$. We find it convenient to use primes “on” the indices for quantities related to the new system; here, $u^{i'}$, rather than putting a prime or a bar on the u symbol, such as $u^{i'}$ or \bar{u}^i . This is for notational purposes, since the indices then will match on the RHS and LHS of the equations. This will become clear below.

13.1 Transformation of the Basis Vectors

The first task is to find expressions for the basis vectors in the new coordinate system in terms of the basis vectors in the old one. Let there be two curvilinear coordinate systems (u^1, u^2, u^3) and $(u^{1'}, u^{2'}, u^{3'})$. Both systems are defined by transformation equations of the form

$$u^j = u^j(x, y, z) \quad (13.1.1a)$$

$$u^{i'} = u^{i'}(x, y, z) . \quad (13.1.1b)$$

There then exists a transformation directly from u^j to $u^{i'}$

$$u^{1'} = u^{1'}(u^1, u^2, u^3)$$

$$u^{2'} = u^{2'}(u^1, u^2, u^3)$$

$$u^{3'} = u^{3'}(u^1, u^2, u^3) \quad (13.1.2)$$

or

$$u^{i'} = u^{i'}(u^1, u^2, u^3) , \quad \text{or} \quad u^{i'} = u^{i'}(u^j) \quad (13.1.3)$$

and vice versa:

$$u^j = u^j(u^{1'}, u^{2'}, u^{3'}) , \quad \text{or} \quad u^j = u^j(u^{i'}) . \quad (13.1.4)$$

(The only limitation we impose on these transformations is that the Jacobian is nonzero.) We define the local basis vectors in each system as usual:

$$\mathbf{e}_j = \frac{\partial \mathbf{R}(u^1, u^2, u^3)}{\partial u^j} \quad (13.1.5)$$

and

$$\mathbf{e}_{i'} = \frac{\partial \mathbf{R}(u^{1'}, u^{2'}, u^{3'})}{\partial u^{i'}} . \quad (13.1.6)$$

Since we can think of the dependence of \mathbf{R} on $u^{i'}$ (as in $\mathbf{R}(u^{1'}, u^{2'}, u^{3'})$) as being established via the u^j , i.e., $\mathbf{R}(u^j(u^{i'}))$, we can apply the chain rule of calculus to (13.1.6) as follows:

$$\mathbf{e}_{i'} = \frac{\partial \mathbf{R}(u^{i'})}{\partial u^{i'}} = \frac{\partial \mathbf{R}(u^j(u^{i'}))}{\partial u^{i'}} = \frac{\partial \mathbf{R}}{\partial u^j} \frac{\partial u^j}{\partial u^{i'}} \quad (13.1.7)$$

so that we have with definition (13.1.5)

$$\mathbf{e}_{i'} = \mathbf{e}_j \frac{\partial u^j}{\partial u^{i'}} , \quad (13.1.8)$$

where a summation over j is understood. Observe that the indices on both sides of the equation “line up” as usual. Equation (13.1.8) is called the “direct” transformation (from old to new). Application of the chain rule to the RHS of (13.1.5) leads to the coefficients for the “inverse” transformation

$$\mathbf{e}_j = \mathbf{e}_{i'} \frac{\partial u^{i'}}{\partial u^j} . \quad (13.1.9)$$

The summation is over i , not over i' (i' does not exist by itself); the prime is simply attached to i to denote transformed quantities. If we would sum over i' , this would be a dummy index, and we could replace it by another, e.g., an unprimed index.

Analogous operations can be applied to the reciprocal-basis vectors $\mathbf{e}^i = \nabla u^i$ and $\mathbf{e}^{i'} = \nabla u^{i'} . \nabla u^{i'}$. Via the chain rule, we obtain

$$\mathbf{e}^{i'} = \mathbf{e}^j \frac{\partial u^{i'}}{\partial u^j} \quad (13.1.10)$$

and

$$\mathbf{e}^j = \mathbf{e}^{i'} \frac{\partial u^j}{\partial u^{i'}} . \quad (13.1.11)$$

13.2 Transformation of Vector Components

In order to find the relationship between the contravariant components of a vector $\mathbf{A} = A^j \mathbf{e}_j = A^{i'} \mathbf{e}_{i'}$ in different coordinate systems, we substitute (13.1.9) into $A^j \mathbf{e}_j$ and equate it to $A^{i'} \mathbf{e}_{i'}$. This gives

$$A^{i'} = A^j \frac{\partial u^{i'}}{\partial u^j} . \quad (13.2.1)$$

The inverse expression follows from substituting (13.1.8) into $A^i e_i$ and equating it to $A^i e_i$,

$$A^j = A^{i'} \frac{\partial u^j}{\partial u^{i'}} . \quad (13.2.2)$$

A summation over repeated indices is implied, and the remaining index on the RHS is carried over to the LHS.

Objects that transform in the same fashion (having primed and non-primed indices in the same place) are said to transform *cogrediently* (Mathews and Walker 1970; Springer 1962; Behnke 1974; Menzel 1961). Objects that transform the other way around transform *contragrediently*. It is now *convention* to name the transformation properties after the similarities with those of the *tangent-basis vectors* e_i . Objects that transform cogrediently to the tangent-basis vectors (i.e., similarly to the e_i) are called *covariant* objects. Those that transform contragrediently to the tangent-basis vectors (i.e., differently from the e_i) are *contravariant* objects.

A comparison of (13.2.1) and (13.2.2) with (13.1.8) and (13.1.9) respectively, shows that the primes in the partial derivates are in the “wrong” places. Therefore, the components A^i transform contragrediently to e_i and are therefore called *contravariant components*.

To find the transformation properties of the components of A which have a subscript, we start again from $A = A_j e^j$ with $A_j = A \cdot e_j$. We now dot multiply this expression for A by $e_{i'}$ in order to obtain

$$A_{i'} = A_j e^j \cdot e_{i'} = A_j e^j \cdot e_k \frac{\partial u^k}{\partial u^{i'}} . \quad (13.2.3)$$

The scalar product of the basis vectors is a Kronecker delta δ_k^i so that

$$A_{i'} = A_j \frac{\partial u^j}{\partial u^{i'}} . \quad (13.2.4)$$

Similarly, the inverse transformation is

$$A_j = A_{i'} \frac{\partial u^{i'}}{\partial u^j} . \quad (13.2.5)$$

Equations (13.2.4) and (13.2.5) use primes in the same places as the tangent-basis vectors of (13.1.8) and (13.1.9), explaining the name *covariant* components for $A_{i'}$.

These transformation properties can be used as a *definition* of “vectors”. A “vector” could be defined as an object whose components transform in the way explained above. Depending on which transformation property applies, we talk about its co- or contravariant components. This definition is probably the most familiar to physicists and engineers.

13.3 Transformation of Tensor Components

The transformation rules given below are limited to second-order tensors. Generalization to higher order is straightforward. We shall only state the transformations because their derivation is completely analogous to that for vectors. Even better, we can simply write down the answer if we look carefully at the indices.

The covariant components F_{ij} of a tensor $F_{ij} e^i e^j$ transform as:

$$F_{i'j'} = F_{mn} \frac{\partial u^m}{\partial u^{i'}} \frac{\partial u^n}{\partial u^{j'}} \quad (13.3.1)$$

$$F_{ij} = F_{m'n'} \frac{\partial u^m}{\partial u^i} \frac{\partial u^{n'}}{\partial u^j} . \quad (13.3.2)$$

For the contravariant components, we have

$$F^{i'j'} = F^{mn} \frac{\partial u^{i'}}{\partial u^m} \frac{\partial u^{j'}}{\partial u^n} \quad (13.3.3)$$

$$F^{ij} = F^{m'n'} \frac{\partial u^i}{\partial u^{m'}} \frac{\partial u^j}{\partial u^{n'}} . \quad (13.3.4)$$

There is no distinction between the transformation properties of the two sorts of *mixed* components. We just need mixed components of the same sort on both sides of the equation:

$$F_{i'j'}^i = F_{j'}^i \frac{\partial u^{i'}}{\partial u^i} \frac{\partial u^j}{\partial u^{j'}} \quad \text{or} \quad F_j^{i'i} = F_j^{i'} \frac{\partial u^j}{\partial u^{j'}} \frac{\partial u^{i'}}{\partial u^i} \quad (13.3.5)$$

$$F_{i'j'}^i = F_{i'}^j \frac{\partial u^i}{\partial u^{i'}} \frac{\partial u^j}{\partial u^j} \quad \text{or} \quad F_j^{i'i} = F_{j'}^{i'} \frac{\partial u^j}{\partial u^{j'}} \frac{\partial u^{i'}}{\partial u^i} . \quad (13.3.6)$$

13.4 Transformation of Components of Special Tensors and Symbols

13.4.1 The Kronecker Delta

We show now that the Kronecker delta written as δ_j^i represents the mixed components of the identity tensor. We have, according to (13.3.5),

$$\delta_j^{i'} = \delta_n^m \frac{\partial u^{i'}}{\partial u^m} \frac{\partial u^n}{\partial u^{j'}} = \frac{\partial u^{i'}}{\partial u^m} \frac{\partial u^m}{\partial u^{j'}} = \frac{\partial u^{i'}}{\partial u^{j'}} = \delta_j^{i'} . \quad (13.4.1)$$

Here we have used the chain rule. This shows that the Kronecker delta, having the values of either 1 or 0, remains unchanged in another coordinate system.

The symbols δ_{ij} and δ^{ij} being equal to 1 if the indices coincide and zero otherwise can likewise be shown *not* to qualify as components of a tensor. As shown in Chap. 3, the covariant and contravariant counterparts of δ_j^i are the metric coefficients g_{ij} and g^{ij} , all being components of one and the same “fundamental tensor”. The transformation properties of the metric coefficients are shown momentarily.

13.4.2 Metric Coefficients

To show that the g_{ij} are legitimate components of a tensor, we make use of the property that the arc length is an invariant quantity not influenced by a change in coordinates. We must have that

$$(dl)^2 = g_{ij} du^i du^j = g_{m'n'} du^{m'} du^{n'} . \quad (13.4.2)$$

We apply the chain rule to the primed differentials as follows:

$$du^{m'} = \frac{\partial u^{m'}}{\partial u^i} du^i \quad \text{and} \quad du^{n'} = \frac{\partial u^{n'}}{\partial u^j} du^j . \quad (13.4.3)$$

This gives for (13.4.2)

$$(dl)^2 = g_{ij} du^i du^j = g_{m'n'} \frac{\partial u^{m'}}{\partial u^i} \frac{\partial u^{n'}}{\partial u^j} du^i du^j \quad (13.4.4)$$

or

$$g_{ij} = g_{m'n'} \frac{\partial u^{m'}}{\partial u^i} \frac{\partial u^{n'}}{\partial u^j} , \quad (13.4.5)$$

which is exactly the transformation law for covariant components.

13.4.3 Levi-Civita Symbols

To show that the Levi-Civita symbols ϵ_{ijk} and ϵ^{ijk} , which have a value +1, -1 or 0, are *not* tensor components, we assume for a moment that they are legitimate components and apply the tensor transformation rules. We will then find a contradiction. By way of example, we treat only ϵ_{ijk} . Its transformation properties would be

$$\epsilon_{l'm'n'} = \epsilon_{ijk} \frac{\partial u^i}{\partial u^{l'}} \frac{\partial u^j}{\partial u^{m'}} \frac{\partial u^k}{\partial u^{n'}} . \quad (13.4.6)$$

Since the RHS is summed over i, j and k , and because ϵ_{ijk} assumes only values +1, -1 and zero, there are nine terms. If we look carefully at that sum, we observe that it equals the Jacobian of the transformation for l, m and n equal to

any even permutation of 1, 2, 3; it equals minus the Jacobian for l, m, n an odd permutation and is zero otherwise. Therefore, we have

$$\epsilon_{l'm'n'} = \epsilon_{lmn} \frac{\partial(u^1, u^2, u^3)}{\partial(u^{l'}, u^{m'}, u^{n'})} \quad (13.4.7)$$

in the notation of Chap. 2. There is a rule similar to the chain rule for Jacobians:

$$\frac{\partial(u^1, u^2, u^3)}{\partial(u^{l'}, u^{m'}, u^{n'})} = \frac{\partial(u^1, u^2, u^3)}{\partial(x, y, z)} \frac{\partial(x, y, z)}{\partial(u^{l'}, u^{m'}, u^{n'})} = \frac{\sqrt{g'}}{\sqrt{g}} . \quad (13.4.8)$$

Hence we obtain for (13.4.7)

$$\epsilon_{l'm'n'} = \epsilon_{lmn} \frac{\sqrt{g'}}{\sqrt{g}} . \quad (13.4.9)$$

This shows that for $\epsilon_{lmn} = +1, -1$ or 0, the transformed component would be equal to $\pm\sqrt{g'}/\sqrt{g}$ and zero instead of $\pm 1, 0$. Therefore, the ϵ_{ijk} are not components of a tensor, as they do not transform properly. The argument for ϵ^{ijk} is analogous. Note that in (13.4.9), our rule of lining up indices no longer works. The reason is that we introduced a new symbol $\sqrt{g'}$ and we attached the prime to that.

The reader can check easily that the components of the “permutation tensor”, $E_{ijk} = \epsilon_{ijk}\sqrt{g}$ and $E^{ijk} = \epsilon^{ijk}/\sqrt{g}$ do transform as required.

13.4.4 Christoffel Symbols

The Christoffel symbols were defined in Sect. 2.6.1c as

$$\left\{ \begin{matrix} j \\ i \ k \end{matrix} \right\} = \frac{1}{2} g^{jn} \left[\frac{\partial g_{ni}}{\partial u^k} + \frac{\partial g_{nk}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^n} \right] \quad (13.4.10)$$

and

$$[j, ik] = \frac{1}{2} \left[\frac{\partial g_{ji}}{\partial u^k} + \frac{\partial g_{jk}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^j} \right] . \quad (13.4.11)$$

These symbols are *not* tensor components in general. We leave its proof as an exercise, because it is found in many textbooks on mathematical physics. We do give a few hints though as well as the answer. Compute $\partial g_{m'l}/\partial u^{n'}$ where $g_{m'l}$ is given by the inverse relationship of (13.4.5). The product rule of differentiation is applied, and the term involving $\partial g_{ji}/\partial u^{n'}$ is expanded via the chain rule as $(\partial g_{ji}/\partial u^k)(\partial u^k/\partial u^{n'})$. Doing this for all the terms in (13.4.11) leads to

$$[m', l'n'] = \frac{\partial u^j}{\partial u^{m'}} \frac{\partial u^i}{\partial u^{l'}} \frac{\partial u^k}{\partial u^{n'}} [j, ik] + \frac{\partial^2 u^i}{\partial u^{l'} \partial u^{n'}} \frac{\partial u^k}{\partial u^{m'}} g_{ik} . \quad (13.4.12)$$

For the other symbol, one also uses the transformed equation for the g^{jn} . This gives after use of the Kronecker delta:

$$\left\{ \begin{array}{c} s' \\ r' \end{array} \right\} = \frac{\partial u^{s'}}{\partial u^j} \frac{\partial u^i}{\partial u^{r'}} \frac{\partial u^k}{\partial u^i} \left\{ \begin{array}{c} j \\ i \end{array} \right\} + \frac{\partial^2 u^i}{\partial u^{r'} \partial u^{i'}} \frac{\partial u^{s'}}{\partial u^i}. \quad (13.4.13)$$

It is because of the second terms, which are, in general, nonzero that the symbols are not tensor components.

14. Alternative Derivations of the Divergence Formula

In this chapter, we discuss two alternative derivations for the divergence formula of (2.6.39). The one found most frequently in the applied mathematics literature involves Christoffel symbols. This derivation is rather remote from the usual plasma physics calculations, however. Nevertheless, for completeness and cohesiveness of our treatment, we include it as a last case. The method given first is the most natural.

14.1 A Straightforward Derivation of the Divergence Formula

Use the product rule of differentiation applied to $A = A^i e_i$ as follows

$$\begin{aligned} \operatorname{div} A \equiv \nabla \cdot A &= \nabla \cdot (A^i e_i) = e_i \cdot \nabla A^i + A^i \nabla \cdot e_i \\ &= e_i \cdot \nabla u^k \frac{\partial A^i}{\partial u^k} + A^i \nabla \cdot e_i. \end{aligned} \quad (14.1.1)$$

Here, we have applied the definition of the ∇ operator to the ∇A^i part. Thus we can write for the divergence

$$\operatorname{div} A \equiv \nabla \cdot A = \frac{\partial A^k}{\partial u^k} + \{ \nabla \cdot e_i \} A^i. \quad (14.1.2)$$

We now find the divergences of the basis vectors e_i . Applying the definition of the ∇ operator gives:

$$\{ \nabla \cdot e_i \} = e^k \cdot \frac{\partial e_i}{\partial u^k} = e^1 \cdot \frac{\partial e_i}{\partial u^1} + e^2 \cdot \frac{\partial e_i}{\partial u^2} + e^3 \cdot \frac{\partial e_i}{\partial u^3}. \quad (14.1.3)$$

(We could have done this in (14.1.1) but we want to display (14.1.2) the way it stands for future reference.) After replacing the contravariant basis vectors e^1 , e^2 and e^3 by the cross products of the covariant-basis vectors, we have:

$$\begin{aligned} \{ \nabla \cdot e_i \} &= \frac{1}{J} \left[\frac{\partial e_i}{\partial u^1} \cdot e_2 \times e_3 + \frac{\partial e_i}{\partial u^2} \cdot e_3 \times e_1 + \frac{\partial e_i}{\partial u^3} \cdot e_1 \times e_2 \right] \\ &= \frac{1}{J} \left[\frac{\partial e_1}{\partial u^i} \cdot e_2 \times e_3 + \frac{\partial e_2}{\partial u^i} \cdot e_3 \times e_1 + \frac{\partial e_3}{\partial u^i} \cdot e_1 \times e_2 \right]. \end{aligned}$$

In the first step, we made use of the commutativity of the dot product, while in the last line, we used (2.6.20). Using the product rule of differentiation in reverse, we have

$$\{\nabla \cdot \mathbf{e}_i\} = \frac{1}{J} \frac{\partial}{\partial u^i} (\mathbf{e}_1 \cdot \mathbf{e}_2 \times \mathbf{e}_3) , \quad (14.1.4)$$

which in turn, because of (2.5.22a) is

$$\begin{aligned} \{\nabla \cdot \mathbf{e}_i\} &= \frac{1}{J} \frac{\partial J}{\partial u^i} = \frac{\partial}{\partial u^i} \ln J = \frac{\partial}{\partial u^i} \ln \sqrt{g} = \frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}}{\partial u^i} \\ &= \frac{1}{2} \frac{\partial \ln g}{\partial u^i} = \frac{1}{2g} \frac{\partial g}{\partial u^i} . \end{aligned} \quad (14.1.5)$$

Substitution into the expression for $\nabla \cdot \mathbf{A}$, (14.1.2), leads to the “standard” divergence expression

$$\text{div } \mathbf{A} \equiv \nabla \cdot \mathbf{A} = \frac{\partial A^k}{\partial u^k} + \frac{A^i}{\sqrt{g}} \frac{\partial}{\partial u^i} (\sqrt{g}) \quad (14.1.6)$$

or because of the product-differentiation rule

$$\text{div } \mathbf{A} \equiv \nabla \cdot \mathbf{A} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial u^k} (\sqrt{g} A^k) = \frac{1}{J} \frac{\partial}{\partial u^k} (J A^k) , \quad (14.1.7)$$

since i and k are dummy indices. In the above derivations, we did not make explicit use of Christoffel symbols or covariant derivatives.

14.2 Divergence-Formula Derivation Employing Christoffel Symbols

Next, we give the derivation of (14.1.7) with Christoffel symbols and the covariant derivative, since it is this derivation that is found in most reference works. Rather than using the rule $\nabla \cdot$ (product) as in (14.1.1), we apply the definition of the ∇ -operator “acting on” the vector \mathbf{A} , and subsequently invoke the covariant derivative formalism developed in Sect. 2.6. Thus,

$$\text{div } \mathbf{A} \equiv \nabla \cdot \mathbf{A} = \nabla u^k \frac{\partial}{\partial u^k} \cdot \mathbf{A} = \nabla u^k \cdot \frac{\partial \mathbf{A}}{\partial u^k} = \mathbf{e}^k \cdot \mathbf{e}_j \left(\frac{\partial \mathbf{A}}{\partial u^k} \right)^j = \left(\frac{\partial \mathbf{A}}{\partial u^k} \right)^k , \quad (14.2.1)$$

where we have used the notation of (2.6.6). Because of (2.6.15) and (2.6.31), this becomes

$$\text{div } \mathbf{A} \equiv \nabla \cdot \mathbf{A} = A^k_k \quad (14.2.2)$$

or

$$\text{div } \mathbf{A} \equiv \nabla \cdot \mathbf{A} = \frac{\partial A^k}{\partial u^k} + \left\{ \begin{matrix} k \\ i \ k \end{matrix} \right\} A^i . \quad (14.2.3)$$

A summation is implied over all repeated indices, i.e., k and i . If we compare this last result with (14.1.2), the curly brackets anticipated a similarity between $\nabla \cdot \mathbf{e}_i$ and the Christoffel symbol in (14.2.3):

$$\nabla \cdot \mathbf{e}_i \equiv \left\{ \begin{matrix} k \\ i \ k \end{matrix} \right\} . \quad (14.2.4)$$

In (14.2.4) a summation over k is implied. Although we kept k fixed, above, the formulae can be used for k variable.

This identity also follows from (14.1.3) and the definition of the Christoffel symbol of the second kind, (2.6.25).

Because of the identity in (14.2.4), we can substitute the result for $\nabla \cdot \mathbf{e}_i$ obtained in (14.1.5) into (14.2.3) to retrieve the expression for the divergence in (14.1.7). However, there is an alternative derivation that makes no explicit use of the basis vectors. This derivation has the advantage (for theoretical physics) that, once the divergence is defined, as in (14.2.2), no reference to basis vectors is needed, but only to metric coefficients. A determinant can be computed by means of the Laplace development. If we call $g = \det[g_{ij}]$, we have that

$$g = \sum_j g_{ij} G^{ij} , \quad i = 1, 2 \text{ or } 3 \quad (14.2.5)$$

where G^{ij} is the cofactor of the element g_{ij} . The summation is only over j . Since G^{ij} does not contain the element g_{ij} explicitly, we can write the derivative of g with respect to g_{im} (i is the fixed index in (14.2.5), and m is one of the j 's)

$$\frac{\partial g}{\partial g_{im}} = G^{im} . \quad (14.2.6)$$

Application of chain-rule differentiation to the expression $\partial g / \partial u^k$ gives

$$\frac{\partial g}{\partial u^k} = \sum_{i,m} \frac{\partial g}{\partial g_{im}} \frac{\partial g_{im}}{\partial u^k} = \sum_{i,m} G^{im} \frac{\partial g_{im}}{\partial u^k} = \sum_{i,m} g g^{im} \frac{\partial g_{im}}{\partial u^k} . \quad (14.2.7)$$

In the last equality, we have applied a relationship discussed underneath (2.5.14)

From the definition of the Christoffel symbol of the first kind, (2.6.26), we observe that

$$[m, ik] + [i, mk] = \frac{1}{2} \left[\frac{\partial g_{mi}}{\partial u^k} + \frac{\partial g_{mk}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^m} \right] + \frac{1}{2} \left[\frac{\partial g_{im}}{\partial u^k} + \frac{\partial g_{ik}}{\partial u^m} - \frac{\partial g_{mk}}{\partial u^i} \right]$$

or because of the symmetry of g_{im} :

$$[m, ik] + [i, mk] = \frac{\partial g_{im}}{\partial u^k} . \quad (14.2.8)$$

Then (14.2.7) becomes

$$\frac{\partial g}{\partial u^k} = \sum_{i,m} gg^{im}([m, ik] + [i, mk]). \quad (14.2.9)$$

This can be rewritten with Christoffel symbols of the second kind according to (2.6.30), recalling the symmetry of g^{im} ,

$$\frac{\partial g}{\partial u^k} = \sum_{i,m} g \left(\begin{Bmatrix} i \\ i & k \end{Bmatrix} + \begin{Bmatrix} m \\ m & k \end{Bmatrix} \right), \quad (14.2.10)$$

which, because i and m are dummy (summation) indices, becomes

$$\frac{\partial g}{\partial u^k} = g \sum_i 2 \begin{Bmatrix} i \\ i & k \end{Bmatrix}. \quad (14.2.11)$$

Application of (2.6.27), omitting the summation sign, and interchanging the dummy indices i and k , finally leads to

$$\begin{Bmatrix} k \\ i & k \end{Bmatrix} = \frac{1}{2g} \frac{\partial g}{\partial u^i}$$

which is the same as (14.1.5).

References

- Abraham, R., Marsden J.E. (1978) *Foundations of Mechanics*, 2nd edn. Benjamin/Cummings, Reading, MA
- Akivis, M.A., Goldberg V.V. (1977) *An Introduction to Linear Algebra and Tensors*. Dover, New York
- Anania, G., Johnson J.L., Weimer K.E. (1983) *Phys. Fluids* **26**, 2210
- Anania, G., Johnson J.L. (1983) *Phys. Fluids* **26**, 3070
- Arnold, V.I. (1978) *Mathematical Methods of Classical Mechanics*. Springer, New York
- Baldwin, D.E. (1977) *Reviews of Modern Physics* **49**, 317
- Balescu, R. (1975) *Equilibrium and Nonequilibrium Statistical Mechanics*. Wiley, New York
- Balescu, R. (1988) *Transport Processes in Plasmas*, 2 vol. North-Holland, Amsterdam
- Bateman, G. (1978) *MHD Instabilities*. MIT Press, Cambridge, MA
- Bauer, F., Betancourt O., Garabedian P. (1978) *A Computational Method in Plasma Physics*. Springer, New York
- Bauer, F., Betancourt O., Garabedian P. (1984) *Magneto Hydrodynamic Equilibrium and Stability in Stellarators*. Springer, New York
- Beasley, C.O., Rome J.A., Attenberger S.E., Hirshman S.P. (1987) Oak Ridge Report, ORNL/TM-10222, Oak Ridge, TN 37831
- Bedeaux, D. (1980) In: Cohen, E.G.D. (ed) *Fundamental Problems in Statistical Mechanics*. North-Holland, New York, Vol. 5, 313
- Behnke, H., Bachmann F., Fladt K., Süss W. (eds) (1974) *Fundamentals of Mathematics*. MIT Press, Cambridge, MA
- Bergland, G.D. (1969) *IEEE Spectrum*, **6**, 41 (July)
- Bernstein, I.B. (1971) In: Simon A., Thompson W.B., (eds) *Advances in Plasma Physics*. Wiley, New York, Vol. 4, 311
- Bernstein, I.B. (1974) *Phys. Fluids* **17**, 547
- Betancourt, O., Garabedian P. (1985) *Phys. Fluids* **28**, 912
- Bhattacharjee, A., Wiley J.C., Dewar R.L. (1984) *Computational Phys. Commun.* **31**, 213
- Bird, R.B., Stewart W.E., Lightfoot E.N. (1960) *Transport Phenomena*, Wiley, New York
- Blum, J., Le Foll J. (1984) *Computer Physics Reports* **1**, 465
- Blum, J. (1989) *Numerical Simulation and Optimal Control in Plasma Physics*. Wiley (Gauthier-Villars), New York
- Boozer, A.H. (1980) *Phys. Fluids* **23**, 904
- Boozer, A.H. (1981) *Phys. Fluids* **24**, 1999
- Boozer, A.H. (1982) *Phys. Fluids* **25**, 520
- Boozer, A.H. (1983) *Phys. Fluids* **26**, 1288
- Boozer, A.H. (1984) Princeton report, PPPL-2094, Princeton, NJ 08544
- Boozer, A.H. (1984) *Phys. Fluids* **27**, 2055
- Boozer, A.H. (1984) *Phys. Fluids* **27**, 2110
- Boozer, A.H. (1984) Princeton report, PPPL-2154, Princeton, NJ 08544
- Boozer, A.H. (1985) Princeton report, PPPL-2094R, Princeton, NJ 08544
- Borisenko, A.I., Tarapov I.E. (1968) *Vector and Tensor Analysis with Applications*. Prentice-Hall, Englewood Cliffs, NJ, re-edited by Dover
- Braginskii, S.I. (1965) In: Leontovich, M.A. (ed) *Reviews of Plasma Physics*. Consultants Bureau, New York, Vol. 1, 205

- Buck, R.C. (1965) *Advanced Calculus*, 2nd edn. McGraw-Hill, New York
- Bulmer, R.H., Kaiser T.B., Nevins W.M., Newcomb W.A., Pearlstein L.D., Strauss H.R. Wollman S., Wakatani M. (1983) *Plasma Physics and Controlled Nuclear Fusion Research* (Baltimore, 1982). IAEA, Vienna, Vol. 1, 531
- Burke, W.L. (1985) *Applied Differential Geometry*. Cambridge University Press, Cambridge, England
- Butkov, E. (1968) *Mathematical Physics*. Addison-Wesley, Reading, MA
- Callen, J.D., Dory R.A. (1972) *Phys. Fluids* **15**, 1523
- Cary, J.R., Littlejohn R.G. (1983) *Annals of Physics* **151**, 1
- Catto, P.J., Hazeltine R.D. (1981) *Phys. Fluids* **24**, 1663
- Catto, P.J., Myra J.R. (1986) *Plasma Phys. and Contr. Fusion* **28**, 959
- Chapman, S., Cowling T.G. (1973) *The Mathematical Theory of Non-uniform Gases*, 3rd edn. Cambridge University Press, Cambridge, England
- Chodura, R., Schlüter A. (1981) *J. Computational Physics* **41**, 681
- Choquet-Bruhat, Y., DeWitt-Morette C., Dillard-Bleick M. (1982) *Analysis and Physics*, revised edn. North-Holland, New York
- Cordey, J.G., Watson C.J.H. (1969) *Proc. of Conf. on Nuclear Fusion Reactors*. UKAEA, Culham Laboratory, England, 122
- Cross, R.C. (1989) *Am. J. Phys.* **57**, 722
- Davis, H.F., Snider A.D. (1975) *Introduction to Vector Analysis*, 3rd edn. Allyn and Bacon, Boston, MA
- Deschamps, G.A., (1970) In: Roubine, E. (ed) *Mathematics Applied to Physics*. Springer Berlin, Heidelberg, Chap. 3, 111
- Dobrott, D., Frieman E.A. (1971) *Phys. Fluids* **14**, 349
- Drew, T.B. (1961) *Handbook of Vector and Polyadic Analysis*. Reinhold, New York
- Ferziger, J.H., Kaper H.G. (1972) *Mathematical Theory of Transport Processes in Gases*. North-Holland, New York
- Feynman, R.P., Leighton R.B., Sands M. (1963) *The Feynman Lectures on Physics*, Addison-Wesley, Reading, MA (3 vols)
- Fielding, P.J., Hitchon W.N.G. (1980) *J. Plasma Physics* **24**, 453
- Filonenko, N.N., Sagdeev R.Z., Zaslavsky G.M. (1967) *Nuclear Fusion* **7**, 253
- Flanders, H. (1963) *Differential Forms with Applications to the Physical Sciences*. Academic, New York
- Freidberg, J.P. (1982) *Rev. Mod. Phys.* **54**, 801
- Freidberg, J.P. (1987) *Ideal Magnetohydrodynamics*. Plenum, New York
- Furth, H.P., Rosenbluth M.N. (1964) *Phys. Fluids* **7**, 764
- Gelfand, I.M. (1961) *Lectures on Linear Algebra*. Wiley, New York
- Goetz, A. (1970) *Introduction to Differential Geometry*. Addison-Wesley, Reading, MA
- Golant, V.E., Zhilinsky A.P., Sakharov I.E. (1980) *Fundamentals of Plasma Physics*. Wiley, New York
- Goldstein, H. (1981) *Classical Mechanics*, 2nd edn. Addison-Wesley, Reading, MA
- Grad, H., Rubin H. (1958) *Proc. of the Second United Nations Conference on the Peaceful Uses of Atomic Energy*, United Nations, Geneva, Vol. 31, 190
- Grad, H. (1967) *Phys. Fluids*, **10**, 137
- Grad, H., Hogan J. (1970) *Phys. Rev. Lett.* **24**, 1337
- Grad, H. (1971) *Plasma Physics and Controlled Nuclear Fusion Research* (Madison, 1971). IAEA, Vienna, Vol. 3, 229
- Grad, H. (1974) *Advances of Plasma Physics*. Wiley, New York, Vol. 5, 103
- Grad, H., Hu P.N., Stevens D.C. (1975) *Proc. Nat. Acad. Sci. USA*, **72**, 3789
- Greene, J.M., Johnson J.L. (1961) *Phys. Fluids* **4**, 875
- Greene, J.M., Johnson J.L. (1962) *Phys. Fluids*, **5**, 510
- Greene, J.M., Johnson J.L., Weimer K.E. (1971) *Phys. Fluids*, **14**, 671
- Guggenheimer, H.W. (1977) *Differential Geometry*. Dover, New York

- Hall, L.S., McNamara B. (1975) *Phys. Fluids* **18**, 552
- Hamada, S. (1962) *Nuclear Fusion* **2**, 23
- Harris, F.J. (1978) *Proceedings of the IEEE*, **66**, 51
- Hazeltine, R.D., Meiss J.D. (1985) *Physics Reports* **121**, 1
- Hinton, F.L., Hazeltine R.D. (1976) *Reviews of Modern Physics* **48**, 239
- Hirshman, S.P., Jardin S.C. (1979) *Phys. Fluids* **22**, 731
- Hirshman, S.P., Sigmar D.J. (1981) *Nuclear Fusion* **21**, 1079
- Hirshman, S.P., (1982) Oak Ridge National Laboratory Report, ORNL/TM-8393, unpublished
- Hirshman, S.P., Whitson J.C. (1983) *Phys. Fluids* **26**, 3553
- Hirshman, S.P., Lee D.K. (1986) *Computational Phys. Commun.* **39**, 161
- Hitchon, W.N.G., Fielding P.J. (1981) *Nuclear Fusion* **21**, 775
- Hoffman, K., Kunze P. (1971) *Linear Algebra*. Prentice-Hall, Englewood Cliffs, NJ
- Hoffmann, B. (1977) *About Vectors*. Dover, New York
- Jackson, J.D. (1975) *Classical Electrodynamics*. 2nd edn. Wiley, New York
- Jancel, R., Kahan T. (1966) *Electrodynamics of Plasmas*. Wiley, New York
- Johnson, J.L., Oberman C.R., Kulsrud R.M., Frieman E.A. (1958) *Phys. Fluids*, **1**, 281
- Johnson, J.L., Boozer A.H., Koniges A.E., Kuo-Petravic G., Manickam J., Monticello D.A., Mynick H., Park W., Reiman A., Rewoldt G. 1984, Princeton report, PPPL-2143, Princeton, NJ 08544
- Johnson, J.L. (1986) *Computational Phys. Rep.* **4**, 37
- Kerst, D.W. (1962) *Plasma Physics (J. of Nuclear Energy, Part C)* **4**, 253
- Kieras, C.E. (1982) Ph.D. Thesis, University of Wisconsin, Physics, Appendix A
- Killeen, J. (ed) (1976) *Methods in Computational Physics*, Vol. 16. Academic, New York
- Kovrzhnykh, L.M. (1984) *Nuclear Fusion* **24**, 851
- Krall, N.A., Trivelpiece A.W. (1973) *Principles of Plasma Physics*. McGraw-Hill, New York
- Kruskal, M.D., Kulsrud R.M. (1958) *Phys. Fluids* **1**, 265
- Kulsrud, R.M. (1983) In: Rosenbluth, M.N., Sagdeev, R.Z. (eds) *Handbook of Plasma Physics*. Vol. 1: *Basic Plasma Physics I*. North-Holland, New York, 115
- Kuo-Petravic, G., Boozer A.H., Rome J.A., Fowler R.H. (1983) *J. of Computational Physics* **51**, 261
- Kuo-Petravic, G. (1984) *Computer Physics Communications* **33**, 353
- Kuo-Petravic, G., Boozer A.H. (1986) Princeton Report, PPPL-2320, Princeton, NJ 08544 (Also, in *J. Computational Physics* **73**, 107 (1987))
- Lackner, K. (1976) *Computational Phys. Rep.* **12**, 33
- Lamb, H. (1945) *Hydrodynamics*, 6th edn. Dover, New York
- Lao, L.L., Hirshman S.P., Wieland R.M. (1981) *Phys. Fluids* **24**, 1431
- Lao, L.L. (1984) *Computational Phys. Rep.* **31**, 201
- Lee, D.K., Hedrick C.L., Callen J.D. (1986) *Phys. Rev. A* **33**, 2675
- Lichnerowicz, A. (1962) *Elements of Tensor Calculus*. Wiley, New York
- Lichtenberg, A.J., Lieberman M.A. (1983) *Regular and Stochastic Motion*. Springer, New York
- Littlejohn, R.G. (1984) In: Horton C.W., Jr., Reichl L.E., (eds) *Statistical Physics and Chaos in Fusion Plasmas*. Wiley, New York, 43
- Lüst, R., Schlüter A. (1957) *Z. Naturforschung* **12a**, 850 (in German)
- Margenau, H., Murphy G.M. (1976) *The Mathematics of Physics and Chemistry*. Krieger, Huntington, NY
- Mathews, J., Walker R.L. (1970) *Mathematical Methods of Physics*. Benjamin/Cummings, Menlo Park, CA
- Menzel, D.H. (1961) *Mathematical Physics*. Dover, New York
- Milne-Thomson, L.M. (1968) *Theoretical Hydrodynamics*, 5th edn. Macmillan, London
- Misner, C.W., Thorne K.S., Wheeler J.A. (1973) *Gravitation*. W.H. Freeman, San Francisco, CA
- Miyamoto, K. (1980) *Plasma Physics for Nuclear Fusion*. MIT Press, Cambridge, MA

- Morozov, A.I., Solov'ev L.S. (1966) In: Leontovich, M.A. (ed) *Reviews of Plasma Physics*. Consultants Bureau, New York, Vol. 2, 1
- Morse, P.M., Feshbach H. (1953) *Methods of Theoretical Physics*. McGraw-Hill, New York
- Mukhovatov, V.S., Shafranov V.D. (1971) Nuclear Fusion **11**, 605
- Myra, J.R., Catto P.J. (1982) Phys. Rev. Lett. **48**, 620
- Myra, J.R., Catto P.J., Hazeltine R.D. (1982) Phys. Fluids **25**, 2028
- Newcomb, W.A. (1958) Annals of Physics **3**, 347
- Newcomb, W.A. (1959) Phys. Fluids **2**, 362
- Newcomb, W.A. (1981) J. Plasma Physics **26**, 529
- Panofsky, W.K.H., Phillips M. (1962) *Classical Electricity and Magnetism*. Addison-Wesley, Reading, MA
- Pao, Y.-P. (1976) Phys. Fluids **19**, 1177
- Pao, Y.-P. (1978) Phys. Fluids **21**, 1120
- Pearlstein, L.D., Kaiser T.B., Newcomb W.A. (1981) Phys. Fluids **24** 1326
- Peversev, G.V., Shafranov V.D., Zakharov L.E. (1978) In: *Theoretical and Computational Plasma Physics* (Proc. of the 3rd Kiev Conference on Plasma Physics Theory, Trieste, 1977) IAEA, Vienna, 469
- Pytte, A., Boozer A.H. (1981) Phys. Fluids **24**, 88
- Reiman, A.H., Boozer A.H. (1984) Phys. Fluids **27**, 2446
- Reiman, A.H., Greenside H.S. (1988) J. Computational Phys. **75**, 423
- Reiman, A.H., Pompfrey N. (1989) Princeton Report, PPPL-2652, to be published
- Rome, J.A. (1989) J. Computational Phys. **82**, 348
- Rosenbluth, M.N., Sagdeev R.Z., Taylor J.B. (1966) Nuclear Fusion **6**, 297
- Schmidt, G. (1979) *Physics of High Temperature Plasmas*. Academic, New York
- Schouten, J.A. (1951) *Tensor Analysis for Physicists*. Clarendon, Oxford
- Schutz, B.F. (1980) *Geometrical Methods of Mathematical Physics*. Cambridge University Press, Cambridge, UK
- Schwenn, U. (1984) Computational Phys. Commun. **31**, 167
- Seshadri, S.R. (1973) *Fundamentals of Plasma Physics*. Elsevier, Amsterdam
- Shafranov, V.D. (1958) Soviet Physics JETP **6**, 545 [J. Exptl. Phys. (USSR) **33**, 710 (1957)]
- Shafranov, V.D. (1960) Soviet-Physics JETP **37**, 775 [J. Exptl. Phys. (USSR) **37**, 1088 (1959)]
- Shafranov, V.D. (1966) In: Leontovich M.A. (ed) *Reviews of Plasma Physics*. Consultants Bureau, New York, Vol. 2, 103
- Shafranov, V.D. (1968) Nuclear Fusion **8**, 253
- Shafranov, V.D., Zakharov L.E. (1972) Nuclear Fusion **12**, 599
- Shkarofsky, I.P., Johnston, T.W., Bachynski M.P. (1966) *The Particle Kinetics of Plasmas*. Addison-Wesley, Reading, MA
- Solov'ev, L.S. (1968) Soviet Physics JETP **26**, 400 [Zh. Eksp. Teor. Fiz. **53**, 626 (1967)]
- Solov'ev, L.S. (1968) Soviet Physics JETP **26**, 1167 [Zh. Eksp. Teor. Fiz. **53**, 2063 (1967)]
- Solov'ev, L.S., Shafranov V.D. (1970) In: Leontovich, M.A. (ed) *Reviews of Plasma Physics*. Consultants Bureau, New York, Vol. 5, 1
- Solov'ev, L.S. (1975) In: Leontovich, M.A. (ed) *Reviews of Plasma Physics*. Consultants Bureau, New York, Vol. 6, 239
- Sowerby, L. (1974) *Vector Field Theory with Applications*. Longman, London
- Spiegel, M.R. (1959) *Vector Analysis and an Introduction to Tensor Analysis*. McGraw-Hill, New York
- Spies, G.O., Nelson D.B. (1974) Phys. Fluids **17**, 1879
- Spies, G.O. (1978) Nuclear Fusion **18**, 1671
- Spitzer, L., Jr. (1958) Phys. Fluids **1**, 253; also reprinted in IEEE PS **9**, 130 (1981)

- Spivak, M. (1970 and 1975) *A Comprehensive Introduction to Differential Geometry*. Perish Publ., Boston, MA, Vols. 1–2, 1970, and Vols. 3–5, 1975
- Springer, C.E. (1962) *Tensor and Vector Analysis*. Ronald, New York
- Stacey, W.M., Jr. (1981) *Fusion Plasma Physics*. Wiley, New York
- Stoker, J.J. (1969) *Differential Geometry*. Wiley, New York
- Stratton, J.A. (1941) *Electromagnetic Theory*. McGraw-Hill, New York
- Symon, K.R. (1971) *Mechanics*, 3rd edn. Addison-Wesley, Reading, MA
- Van Bladel, J. (1964) *Electromagnetic Fields*. McGraw-Hill, New York
- White, R.B., Boozer A.H., Goldston R., Hay R., Albert J., Karney C.F.F. (1983) *Plasma Physics and Controlled Nuclear Fusion Research* (Baltimore 1982), IAEA, Vienna, Vol. 3, 391
- White, R.B., Chance M.S. (1984) Phys. Fluids **27**, 2455
- White, R.B. (1989) *Theory of Tokamak Plasmas*. North-Holland, Amsterdam
- Whiteman, K.J. (1977) Rep. Prog. Phys. **40**, 1033
- Wimmel, H.K. (1982) European J. of Physics **3**, 223
- Woodson, H.H., Melcher J.R. (1968) *Electromechanical Dynamics*, Part III. Wiley, New York
- Wrede, R.C. (1963) *Introduction to Vector and Tensor Analysis*. Wiley, New York; re-edited by Dover
- Zakharov, L.E., Shafranov V.D. (1981) In: Kadomtsev, B.B. (ed) *Plasma Physics*. MIR, Moscow, 13
- Zakharov, L.E., Shafranov V.D. (1986) In: Leontovich, M.A. (ed) *Reviews of Plasma Physics*. Consultants Bureau, New York, Vol. 11, 153

Subject Index

- Abraham* 54
Absolute derivative 32
Absolute differential 31
Abstract vector analysis
– and the covariant derivative 31
Action-angle form
– of the magnetic-field Hamiltonian 173
Algebra
–, abstract considerations 19
–, Grassmann 41
Ampère's law
–, differential form 56, 111, 124, 133, 135
–, integral form 125, 134, 136, 138, 158, 192
Anania 163
Angle 67
–, cylindrical 68–69
–, poloidal
– –, generalized 71–72
– – – like variable 71
– – –, usual (elementary) definition 69
–, toroidal
– –, generalized 71–72
– – – like variable 71
– – –, usual (elementary) definition 69
Annihilator
–, the flux-surface average as 92–93
Arc length, differential 20
Arc length along a coordinate curve
–, differential 28, 136
– – in the toroidal direction 125
Arc length along B
– and the Boozer-Grad function χ 109, 110
– as variable in a Clebsch system 103, 104, 156
– in terms of poloidal and toroidal flux coordinates 156–157
Area element in a coordinate surface, differential 28, 29
Arnold 54
Arnold-Kolmogorov-Moser theorem see KAM theorem
Attenberger 161, 168

Average
–, flux-surface ~ 85–93
– – and the dl/B integral in toroidal systems 87–89
– – and the dl/B integral in mirrors 89
– – of the inverse Jacobian 149
– – of the Jacobian according to Shafranov 149
– rotational-transformation angle 82
Axial magnetic flux, in mirrors 77, 82
Axis
–, magnetic 66, 70, 77, 81–82
–, major 66, 69
–, minor 66, 69
Axisymmetric
– device 66
– tokamak 65
Azimuthal magnetic flux, in mirrors 77, 82

 β (or v) coordinate surfaces
– in toroidal systems 75, 87–88, 113
“Bac-cab” rule
–, application of 9, 114, 150, 186
– identity 38 (V. 2)
Bachynski 46, 52, 54
“Back-and-forth” relations
– between-basis vectors
– –, as reciprocal sets 15, 24, 167, 174
– –, via metric coefficients 21
– between components of a vector, via metric coefficients 21
– between reciprocal sets of vectors 8
Baldwin 164
Balescu 1, 166
Banana regime, neoclassical pinch velocity 198
Basis 11, 13
–, local 13
“Basis-dyads” 41
– constructed from basis vectors 41
–, expansion of a dyad along 41
“Basis-tensors” 41
– constructed from basis vectors 41
–, expansion of a dyad along 41
–, expansion of a tensor along 42

Basis vectors 8, 11, 15
 —, “back-and-forth” relations
 —, as reciprocal sets 15, 24, 167, 174
 —, via metric coefficients 21
 —, differentiation of 30–32
 —, related to the Christoffel symbols 34–35
 — for the Boozer-Grad system 110
 — for the Cartesian system 9
 — for the generic Clebsch system 104
 — for the Hamada system 154
 —, Jacobian as a triple product of 23, 24
 —, reciprocal ~ 14–15
 —, reciprocal sets acting as 8, 14
 —, tangent ~ 11–13
 —, transformation of 204–205
 —, unit 13
Bateman 1, 19, 84, 101, 122, 163, 200
Bauer 164
Beasley 161, 168
Bedeaux 100
Behnke 2, 54, 206
Bergland 161
Bernstein 101
Betancourt 108, 164
Bhattacharjee 164
Binormal 62, 63
Bipolar coordinates 68, 179–180, 183–185
 —, constant coordinate curves 179, 180, 184
 —, relationship to “proper” toroidal coordinates 182–183
Bird 7, 46, 52, 54
Blum 79
Book 38
Boozer 1, 19, 107, 108, 113, 122, 136, 142, 143, 149, 151, 161, 162, 163, 164, 166, 168, 169, 171, 174, 175
Boozer (toroidal flux) coordinates 2, 98, 121, 142–146, 161
 —, mathematical definition of 143
Boozer-Grad
 — covariant form of \mathbf{B} 110, 112, 141, 157
 —, and the covariant $(\varrho, \theta_f, \zeta_f)$ components of \mathbf{B} 140–142
 — form of \mathbf{B} in vacuum 114, 145
Boozer-Grad coordinate system 2, 98, 107–115
 —, Jacobian of 109
Boozer-Grad coordinates 107–115
 —, conversion to toroidal flux coordinates 157–159
 —, covariant form of \mathbf{B} 109, 110, 111, 112
 —, logic behind 108
Boozer-Grad function χ 107–115
 — as the magnetic scalar potential in vacuum 108, 114
Borisenko 7, 17, 54
Bracket notation
 — for dyad components 41
 — for vector components 16
Braginskii 190
Buck 10
Bulmer 164
Burke 54
Butkov 7, 13, 32, 54
Callen 163
Canonical
 — coordinates 3, 99, 165–175
 — form of \mathbf{B} 171
 — perturbation theory 174
 — variables, for magnetic-field description 171, 172
Cartesian
 — components of a position vector 10
 — components of a vector 9
 — coordinate system 9, 16
 —, irrelevance of co- and contravariant designation 17
 —, irrelevance of sub- and superscript notation 17
 — coordinates 11
Cary 166
Catto 91, 92, 108
Chain rule of differentiation (of calculus)
 — applied to the derivative of metric coefficients 213
 — for Jacobians 209
 — for safety-factor evolution 196
 — for the transformation properties of basis vectors 205
Chance 163
Chapman 46, 54
Chodura 164
Choquet-Bruhat 54
Christoffel symbol(s) 34–35, 52, 178, 209–210
 — as defined by Stacey (1981) 34
 — in a rectilinear system 35
 — of the first kind 34
 — of the second kind 34
 —, transformation properties of 209–210
 — used in the divergence-formula derivation 212–214

Clebsch coordinate system(s) 2, 67, 100–115, 156–159
 —, Boozer-Grad (ϱ, β or v, χ) 107–115, 157–159
 —, Jacobian for 109
 —, generic (ϱ, β or v, l) 100–107, 156–157
 —, Jacobian for 104, 105
Clebsch coordinates
 —, Boozer-Grad variant 107–115, 157–159
 —, conversion to toroidal flux coordinates 156–159
 —, generic 100–107, 156–157
 — in toroidal devices 73, 75–76
 — in open-ended devices (mirrors) 76
 — relationship to co- and contravariant formalism 103–107
Clebsch form of \mathbf{B}
 — as contravariant representation 105, 108
Clebsch representation of \mathbf{B} 100, 103, 105, 108, 113, 117, 156, 157
 — in Hamada coordinates 152
 — in tokamaks 126
 — in toroidal flux coordinates 121, 156, 166, 167
 —, in general geometry 121
 —, in tokamaks 121, 123
Clebsch representation of \mathbf{J} 112, 129, 130, 141
 — in Hamada coordinates 152
Clebsch stream function 100, 102, 107
 —, field line-labeling function as 103
 —, flux-surface-labeling function as 103
 — for the current density (λ or η) 112, 129, 141, 150
 — for the magnetic field (β or v)
 —, in terms of arbitrary angle variables 117–119
 —, in terms of Boozer coordinates 144
 —, in terms of flux-coordinate angles 120, 131, 156
Clebsch transformation 100
Clebsch-type coordinate systems 67, 100–115
Closed-loop integral involving $d\mathbf{l}/B$
 — and the current-closure condition 65, 148
 —, for the Hamada system 151
 — and the flux-surface average 87–89
 — and the specific flux volume 201–203
Coefficients, metric 20–22
Cofactor of a matrix element 22, 213
Cogredient(ly) 206
Coil current(s) 129–130, 145
 — and the magnetic scalar potential 134, 138–139
 — in a stellarator/torsatron 145
 — in a tokamak 145
Complementary solution, of a differential equation 133
Component form of the equation of a magnetic-field line 55
Components
 —, curl in terms of derivatives of covariant ~ 37
 —, divergence in terms of derivatives of contravariant ~ 36, 212
 —, mixed, of a dyad 42
 — of a dyad 41, 42
 — of a second-order tensor 47
 — of a third-order tensor 50
 — of the covariant derivative 31
 — of the fundamental tensor 48–50
 — of the identity tensor 48–50
 — of the metric tensor 48–50
 — of the permutation tensor 52
 —, parallel, of a vector 37–38
 —, perpendicular, of a vector 37–38
 —, physical 126
 —, relationship between tensor ~ 47–48
 —, transformation properties
 —— of tensor ~ 207
 —— of vector ~ 205–206
 —— properties of vector and tensor ~ 178, 204–210
Components, contravariant
 — of a dyad 42
 — of a vector 16, 17, 21
 — of \mathbf{B}
 — in Hamada coordinates 152
 — in toroidal flux coordinates 120, 131, 167
 — of \mathbf{J}
 — and the covariant \mathbf{B} components 135, 136
 — in Hamada coordinates 152
 — in current coordinates 130
Components, covariant
 — of a dyad 42
 — of a vector 16, 17, 21
 — of \mathbf{B} 122
 — in toroidal flux coordinates 133–140, 136, 138, 139, 140
 — of the gradient vector 36
 — of the magnetic vector potential 133
 —, related to the Boozer-Grad form 131–142
Conditions
 —, solvability, for a magnetic differential equation 94, 147
Conductivity
 —, electrical 56
 —, finite, in a plasma 198–199

Conductivity (cont.)
 --, infinite, in a plasma 187–190
 --, perpendicular 190
Confined plasmas
 --, curvilinear coordinates for the description of 67
 --, ideal MHD equilibrium conditions for toroidally ~ 146–150
Confinement systems
 --, curvilinear coordinates in 67–76
Constraints
 --, solvability, for a magnetic differential equation 94, 147
Contraction, of a tensor 45
Contragredient(ly) 206
Contravariant 3, 16, 206
 -- basis vector(s) 19, 21
 --, for the generic Clebsch system 104
 --, transformation properties 204–205
 -- components
 -- of a dyad 42
 -- of a second-order tensor 47
 -- of a third-order tensor 50
 -- of a vector 10, 16, 17, 21
 -- of \mathbf{B}
 --, in Hamada coordinates 152
 --, in toroidal flux coordinates 120, 131, 167
 -- of \mathbf{J}
 --, and the covariant \mathbf{B} components 135, 136
 --, in current coordinates 130
 --, in Hamada coordinates 152
 -- of the basis vectors 19
 -- of the covariant derivative 31
 -- of the fundamental tensor 48–50
 -- of the identity tensor 48–50
 -- of the metric tensor 48–50
 -- of the permutation tensor 52
 --, relationship to other tensor components 47–48
 --, transformation of 205–206, 207
 --, mnemonic aid 16
 -- object 16
 -- representation of \mathbf{B}
 -- as a Clebsch form 105, 108
 --, generic 104, 116, 174
 -- in a Clebsch coordinate system 104, 105, 108, 113
 -- in Boozer-Grad coordinates 113
 -- in toroidal flux coordinates 142, 150, 166, 167
 -- “vectors” 17
Convective derivative 191, 195
Convention for the double-dot product 46

Conversion
 -- from Boozer-Grad coordinates
 -- to toroidal flux coordinates 157–159
 -- from Clebsch coordinates
 -- to toroidal flux coordinates 156–159
Coordinate curve(s) 10, 11, 67
 --, constant ~ in 2D 179, 182, 184
 --, differential arc length along 28, 125, 136
 -- for the Clebsch system 103
 -- for the cylindrical-toroidal (hybrid) system 69
 -- for the “elementary” toroidal system 69
 -- for the generalized cylindrical-toroidal system 71
 -- in axisymmetric systems 123
 -- in stellarators versus tokamaks 145, 146
 --, magnetic-field lines as 60, 61
Coordinate surface(s) 10, 11, 67
 --, flux surfaces as 70, 116
 -- for the Boozer-Grad function χ 113
 -- for the Clebsch system 103
 -- for the coordinate (ν or β) labeling field lines 75, 87, 88, 113
 -- for the cylindrical-toroidal (hybrid) system 69
 -- for the “elementary” toroidal system 69
 -- for the generalized cylindrical-toroidal system 71
 -- for the “proper” toroidal system 183
 -- for the spherical system 182
 -- in stellarators versus tokamaks 145, 146
 --, magnetic surfaces as 70, 116
Coordinate system(s)
 --, bipolar 68, 179–180, 183–185
 --, Boozer-Grad 98, 107–115
 --, Cartesian 9
 --, Clebsch 67, 100–115
 --, curvilinear, in confinement systems 67–76
 --, cylindrical 11, 68, 69
 --, cylindrical-toroidal (hybrid) 67, 68–70
 --, “elementary” toroidal 67, 68–70
 --, flux 97–164, 165–170
 --, generic Clebsch 100–103
 --, “natural” 72
 --, orthogonal 11
 --, polar 179, 181–182
 --, “proper” toroidal 68, 179–185
 --, right-handed 68
 --, spherical 11, 179, 181–182
 --, straight- \mathbf{B} 1, 2
see flux-coordinate systems
Coordinates
 --, bipolar 68, 179–180, 183–185
 --, Boozer 142–146
 --, Boozer-Grad 98, 107–115

--, conversion to toroidal flux coordinates 157–159
 --, Boozer’s toroidal flux 98, 142–146
 --, canonical 3, 99, 165–175
 --, Cartesian 11
 --, Clebsch 98, 100–115
 --, generic 100–103
 --, in open-ended systems 76
 --, in toroidal systems 73, 75–76
 --, current 129
see also straight current-density coordinates
 --, curvilinear 1, 11, 10–17
 --, in confinement systems 67–76
 --, cylindrical 11, 68, 69
 --, flux 1, 10, 70–73, 97–164
 --, Clebsch-type 100–115
 --, toroidal 116–155, 165–170
 --, generalized
 --, cylindrical-toroidal 70–73
 --, magnetic 3, 165–175
 --, Hamada 98, 150–155
 --, illustrative example of flux ~ in a tokamak 73
 --, magnetic-flux 1, 10, 70–73, 97–164, 165–170
 --, magnetic, generalized 3, 165–175
 --, non-flux ~ in a tokamak 126–128
 --, orthogonal 11
 --, polar 181–182
 --, “proper” toroidal 68, 178, 179–185
 --, rectilinear 11
 --, spherical 11, 179, 181–182
 --, straight \mathbf{B} 98, 116–122
see also flux coordinates
 --, straight current-density 98, 129–131
 --, relationship to flux coordinates 130
 --, versus Hamada coordinates 130–131
 --, symmetry flux ~ in a tokamak 122–126
 --, toroidal-flux 98, 116–155, 165–170
 --, vector differentiation in curvilinear 30–37
Cordey 164
Correction term for vector differentiation 31, 32, 33
Covariant 3, 16, 206
 -- basis vectors 19, 21
 --, for the generic Clebsch system 104
 --, for the Hamada system 154
 --, transformation properties 204–205
 --, Boozer-Grad form of \mathbf{B} 112, 141
 -- components
 -- of a dyad 42
 -- of a second-order tensor 47
 -- of a third-order tensor 50
 -- of a vector 10, 16, 17, 21
 -- of \mathbf{B} 122
 -- in toroidal flux coordinates 133–140, 136, 138, 139, 140
 -- related to the Boozer-Grad form of \mathbf{B} 141
 -- of the basis vectors 19
 -- of the covariant derivative 31
 -- of the curvature vector 61
 -- of the fundamental tensor 48–50
 -- of the identity tensor 48–50
 -- of the magnetic vector potential 133
 -- of the metric tensor 48–50
 -- of the permutation tensor 52
 --, relationship to other tensor components 47–48
 --, transformation of 206, 207
 -- derivative 30–35
 -- in terms of the metric coefficients 32–34
 -- used in the divergence-formula derivation 212
 -- using Christoffel symbols 35
 -- differentiation 30–35
 -- mnemonic aid 16
 -- objects 16
 -- representation of \mathbf{B}
 --, generic 104, 105, 140
 --, in “final” Boozer-Grad form 110, 112, 141, 157
 --, in “intermediate” Boozer-Grad form 109, 111, 112
 -- in (ρ, β, l) coordinates 105, 109
 -- in toroidal flux coordinates 142
 -- representation of the magnetic vector potential, generic 132, 133
 -- “vectors” 17
Cowling 46, 54
Cross 63
Cross product
 --, divergence of a ~ (vector identity) 39 (V. 9)
 -- identities 38–39
 -- of vectors in curvilinear coordinates 27
Curl
 --, divergence of a ~ (vector identity) 39 (V. 16)
 -- in terms of the covariant derivative 37
 -- of a cross product (vector identity) 39 (V. 10)
 -- of a gradient (vector identity) 39 (V. 15)
 -- of a vector expressed in curvilinear coordinates 36–37
Current
 -- behavior in an ideal tokamak plasma 187–190

Current-closure condition 65, 148
 — and the Hamada system 151
 Current coordinates 129
 —, mathematical definition of 130
 —, notation for 129
 —, relation to flux coordinates 130
 — versus Hamada coordinates 130–131
 Current density
 —, Clebsch representation of 112, 129, 141
 — (Clebsch) stream function 112, 129, 150
 —, boundary (“initial”) conditions for 114
 — related to the Pfirsch-Schlüter current 113–114
 — related to the pressure gradient 113–114
 —, contravariant components of the 135, 136
 — flux (==“current”)
 —, poloidal 129
 —, toroidal 130
 — lines
 —, equation (integrated form)
 —— in current coordinates 130
 —— in toroidal flux coordinates 129
 —, straight ~ coordinates 129–130
 —, mathematical definition of 130
 —, relation to flux coordinates 130
 — versus Hamada coordinates 130–131
 Current, poloidal plasma ~ 140
 Current, total poloidal ~ 140
 Currents, coil 129, 130
 — and the magnetic scalar potential 134, 138–139
 Curvature
 —, components with respect to a surface 60
 —, covariant components of the ~ vector 61
 —, geodesic 60, 61, 62
 —, local radius of 58
 —, magnetic-field line ~ 58–62, 63
 —, normal 60, 61, 62
 —, radius of ~ vector 58
 — vector 58, 59, 60, 61
 Curve(s)
 —, constant coordinate ~ in 2D 179, 182, 184
 —, coordinate 10, 11, 67 (*see also* coordinate curves)
 —, magnetic-field lines as coordinate ~ 60, 61
 Curvilinear
 — coordinate system(s)
 ——, cylindrical-toroidal hybrid 67, 68–70
 — for plasma-confinement systems 67–76

—, generalized cylindrical toroidal 70–73
 —, Jacobian of a 23
 — coordinates 1, 10–17, 11
 — for plasma-confinement systems 67–76
 ——, generalized poloidal and toroidal angles as 71
 ——, vector differentiation in 30–37
 — transformation 10
 “cyc” (definition of “ijk cyc 123”) 15
 Cyclic permutation 15
 Cylindrical angle, related to the toroidal angle 68, 69, 123
 Cylindrical (coordinate) system 11, 68
 —, zero off-diagonal metric coefficients in 20
 Cylindrical-toroidal hybrid coordinate system 67, 68–70
 Davis 15
 Degenerate magnetic surface (magnetic axis) 66
 del operator 35–37
 —, definition of 35
 —, parallel component 59
 —, perpendicular component 59
 Delta, Kronecker 14, 22, 48–50
 Derivative
 —, absolute 32
 —, convective 191, 195
 —, covariant 30–35, 212
 —, directional 59
 —, intrinsic 32
 — of a vector in curvilinear coordinates 30–37
 Deschamps 54
 Determinant
 —, Jacobian expressed as a 23
 — of the metric tensor 51
 — of the product of two matrices 25
 Differential
 — absolute 31
 — arc length 20
 —— along a coordinate curve 28, 136
 —— along a toroidal coordinate curve 125
 —— along B and the Boozer-Grad function χ 109, 110
 — area element in a coordinate surface 28, 29
 — elements 28–29
 — equation
 ——, homogeneous solution of 133
 ——, magnetic 93–96, 144, 147, 149, 153, 154
 —— constraints (solvability conditions) 94, 147

—, particular solution of 133
 — forms, references on 54
 — geometry, references on 10, 54
 —, intrinsic 31
 — position vector as contravariant “vector” 17
 — volume elements 29
 Differentiation
 —, covariant 30–35
 — of basis vectors 30–32
 — of vector components 30–32
 — of vectors 30–37
 —— and Christoffel symbols 30–35
 Diffusion equation for the magnetic field 56
 Directional derivative 59
 Disk
 —, magnetic flux through a 77, 78, 79, 81
 —, poloidal ~ flux 77, 78, 79, 81
 Divergence
 — formula for vectors 36
 —— alternative derivations 178, 211–214
 — of a basis vector 211–212
 —— in terms of a Christoffel symbol 213
 —— in terms of the Jacobian 212
 — of a cross product (vector identity) 39 (V. 9)
 — of a curl (vector identity) 39 (V. 16)
 — of a dyad (identity) 53 (T. 10)
 — of a position vector (vector identity) 39 (V. 17)
 — of a vector expressed in curvilinear coordinates 36
 dl/B integral
 — and flux-surface averaging 87–89
 — and the specific flux volume 201–203
 Dobrott 163
 Dory 163
 Dot product 26
 —, double ~ 46
 —, gradient of a ~ (vector identity) 39 (V. 12)
 — in a general curvilinear system 26
 — in an orthogonal system 26
 — in an orthonormal system 26
 — of a dyad with a vector 40–41
 — of a vector with a tensor 44
 Double-dot product 46, 52–53 (T. 4–T. 8)
 Drew 43
 Dual
 — basis (vectors) 15
see also reciprocal basis (vectors)
 — relations between basis vectors 15, 24
see also “back-and-forth” relations
 Dummy index 14

Dyad
 —, components of a 41
 —, divergence of a (identity) 53 (T. 10)
 —, expansion of a, along “basic-dyads” 41
 — identities 52–54
 —, nonion form of a 43
 Dyadic(s) 43
 — and second-order tensors 43
 — literature on 54
 —, nonion form of a 43
 — notation 3
 —, references on 54
 Dyads 40–44, 46–47, 52–54, 55
 —, definition of 40
 —, double-dot product between 46, 52 (T. 4)
 —, examples of, in plasma physics 40
 — literature on 54
 — notation of 40–41
 —, references on 54
 Dynamic equilibrium of a tokamak plasma 79, 178, 186–200
 $E \times B$ motion 58, 190–191, 192, 193
 Einstein convention *see* Summation convention
 Electric field
 — in an infinitely conducting tokamak plasma 187–190
 —, motion of plasma and flux surfaces due to an 190–195
 Electrical conductivity 56
 —, finite, in a plasma 198–199
 —, infinite, in a plasma 187–190
 —, perpendicular 190
 “elementary” toroidal (coordinate) system 67, 68–70, 128
 Elmo Bumpy Torus 84
 EMF 192
 End losses (contributions) in mirrors 91
 Equation
 —, magnetic differential 93–96
 — of a current-density line
 —— in arbitrary toroidal coordinates (integrated form) 129
 —— in current coordinates (integrated form) 130
 — of a magnetic-field line 55–56
 —, differential form 55, 56, 157, 169, 172, 201
 ——, in canonical coordinates 172
 ——, in Clebsch-type coordinates 107
 ——, in flux coordinates 84, 120, 201
 ——, integrated form 117, 168, 173
 ——, in Clebsch-type coordinates 107
 ——, in flux coordinates 87, 119, 120

Equations, Hamilton's, for a magnetic-field-line 172
Equilibrium
 -, dynamic ~ in a tokamak plasma 79, 178, 186–200
 -, flux-surface determination and 162–164
 -- references on 162–164
 -, ideal MHD ~ conditions for toroidally confined plasmas 146–150
 -- and the Hamada system 151, 152
 -, MHD ~ equation 146, 150
 -, plasma ~ time evolution, references 79
Euclidean vector space 17, 19, 48
Euler formula 184
Evaluation of the magnetic-field-line Hamiltonian 174–175
Even permutation 15
Evolution in time of an ideal tokamak plasma equilibrium 186–200
Existence
 - of Hamada coordinates 152
 - of magnetic surfaces 65, 148, 165–166
Expansion
 - coefficients of a vector along a basis 9
 -, Fourier 95–96, 118, 150, 153, 161, 168–170, 175, 201–202
 - of a dyad along "basis dyads" 41
 - of a vector 8
 -- along reciprocal sets of vectors 8
 -- along the reciprocal basis set in curvilinear coordinates 15
 -- along the tangent basis set in curvilinear coordinates 15–16
 -, stellarator ~ 163
Extended phase space 173
Exterior product 41

Faraday (-Lenz) law 56, 187, 189, 191, 192, 197
Ferziger 46, 54
Feshbach 2, 7, 34, 43, 46, 54, 179
Feynman 132
Field line, magnetic
 - as coordinate curve 60, 61
 - curvature 58–62, 63
 - definition 55
 - equation of a 55–56
 - differential form 55, 56, 157, 169, 172, 201
 --, in canonical coordinates 172
 --, in Clebsch-type coordinates 107
 --, in flux coordinates 84, 120, 201
 --, integrated form 117, 168, 173
 --, in Clebsch-type coordinates 107
 --, in flux coordinates 87, 119, 120
 - Hamiltonian 65, 170–175, 172

--, evaluation of 174–175
 - integration, method for establishing flux coordinates 161, 162, 168–170
 - label, notations for 108
 - principal normal to a 58, 61
 -, "proper length" of a 148, 178
 -- and the specific flux volume 201–203
 -, straight ~ coordinates (in toroidal systems) 116–122
see also magnetic-field line
Field lines, magnetic
 - and magnetic tension 63–64
 - as elastic cords 63
 -, motion of 56–58
Fielding 68, 163, 179, 200
Filonenko 170
Finite-resistivity, evolution of a ~ plasma 198–199
Flanders 54
Fluid velocity of the plasma 56
Flux
 - frozen ~ theorem 56–58
 - function 81
 -, transformer ~ 186, 187, 188, 189, 190, 199
Flux-coordinate system
 - references 122, 161, 162
see also flux coordinates
Flux-coordinate transformation 98–99, 120
 - references 161
 - summary 160–164
 -- for open-ended systems 164
 -- for toroidal systems 160–164
Flux coordinates
 - Boozer's toroidal 142–146, 161
 --, mathematical definition of 143
 - Clebsch (ϱ, β, I)
 -- in open ended systems (qualitative discussion) 76
 -- in toroidal systems (qualitative discussion) 73, 75–76
 - Clebsch-type 100–115, 161
see also Clebsch coordinates
 - conversion from Clebsch coordinates to toroidal ~ 156–159
 - definition of 1, 98
 - establishment of 160–164, 166–170
 - Hamada 150–155, 161
 --, references on 151
 - illustrative example 73
 - "justification" for the use of 160
 - magnetic *see* magnetic-flux coordinates
 - mathematical definition of toroidal ~ 120
 - non- ~ in tokamaks 126–128
 - non-uniqueness of 121

- notation
 -- for Boozer's toroidal ~ 143
 -- for Hamada ~ 151
 -- for symmetry ~ in tokamaks 123
 -- for toroidal ~ 120
 -, overview of 98–99
 -, qualitative description 70–73
 -, references on 1, 122, 161, 162
 -, symmetry ~ in a tokamak 122–126
 -, toroidal 98, 116–155
Flux, magnetic
 - axial 77, 82
 - azimuthal 77, 82
 - form-invariant expression 79–80
 -, poloidal 77, 78, 79, 81
 -- as flux-surface label 81, 116, 191–193, 199–200
 -, toroidal 77, 80, 81
 -- as flux-surface label 81, 116, 193–195
 -, total time derivative of the ~ through a surface 57
Flux surface
 - average 85–93, 196
 -- as annihilator of $B \cdot \nabla$ 92–93
 --, definition of 85
 -- expressed as a surface integral 85, 86
 -- expressed as a volume integral 85
 -- in an axisymmetric device 87–89
 -- in open-ended devices 89
 -- in terms of toroidal and poloidal angles 86
 -- of a divergence 90–92
 -- of the inverse Jacobian 149
 -- of the Jacobian in a Hamada system 151
 --, properties 90–93
 -- related to the line integral of dl/B 87–89
 -, definition
 -- mathematical 81
 -- physical 64, 65, 77
 -, determination/computation 162–164
 --, numerical treatments 163–164
 --, references 162–164
 - function 81
 - label 70
 --, Clebsch stream function as 103
 --, notations for 108
 --, radially increasing 116
 - labeling 77–82
 - motion in an ideal tokamak plasma 190–195
 - motion of 79
 -, poloidal disk flux as label for 78–79, 81, 191–192, 199–200

-, poloidal ribbon flux as label for 78–79, 81, 192–193, 199–200
 - quantity 81
 -, toroidal flux as label for 77, 81, 193–195
 - velocity 192, 193, 195, 198
Flux surfaces, (magnetic) 61, 64–67
 - as constant pressure surfaces 65
 - as coordinate surfaces 70
 - definition 64, 65, 77, 81
 - degenerate 66
 - existence of 65, 165–166
 - in open-ended systems 67
 - in stellarators 66
 - in tokamaks 66
 - in toroidal systems 65–66
 - irrational 65
 - rational 65
 - vacuum 65
see also magnetic (-flux) surfaces
Force balance (equation) 65, 111, 146, 150
Form-invariant expression
 - for the poloidal and toroidal currents 129, 130
 - for the poloidal (magnetic) flux 80, 119
 - for the toroidal (magnetic) flux 80, 118, 196
Formulae
 - Frenet-Serret ~ 62
 - tensor ~ 52–54
 - vector ~ 38–39
Fourier expansion
 - for evaluating $\int dl/B$ 201–202
 - for evaluating the field-line Hamiltonian 175
 - for solving a magnetic differential equation 95–96, 150, 153
 - method for establishing flux coordinates 161, 168–170
 - of the Jacobian 201
 - of the magnetic-field components 118
Fowler 161, 168
Frame, Frenet-Serret 62
Freidberg 163
Frenet-Serret
 - formulae 62
 - frame 62
 - trihedron 63
Frieman 163
Frozen-flux theorem 56–58, 79, 192, 193, 198
Function(s)
 - flux 81
 - multiple-valued 70, 71
 - single-valued 72

Fundamental tensor 48–50, 208
 – in relation to the identity tensor 48–50, 208
 – in relation to the metric tensor 48–50, 208
 –, sum of dyad form of the 49
Fruth 164

Garabedian 108, 164
Gauss' theorem, application of 80, 90

Gaussian window function 161
 Generalized
 – angle variables 71–72
 – cylindrical-toroidal coordinates 70–73
 – magnetic coordinates 3, 165–175
 – poloidal angle 71–72
 – toroidal angle 71–72

Generic
 – Clebsch coordinates 100–103
 – toroidal flux coordinates 116–122

Geodesic curvature 60, 61, 62
 Geometrical object 43
Goetz 10, 60
Golant 56, 190, 200
Goldstein 46, 54, 166, 173
Grad 79, 100, 101, 108, 163, 164, 166, 170
Grad
 – -Schlüter-Shafranov equation 162–163
 – -Shafranov equation *see Grad-Schlüter-Shafranov equation*

see also Boozer-Grad
 gradient 35–36
 – as a covariant “vector” 18
 –, curl of a (vector identity) 39 (V. 15)
 – of a dot product (vector identity) 39 (V. 12)

– of a function, definition 14
 – of the coordinate surfaces as basis vectors 10
 – vector, covariant components of the 36

Grassmann algebra 41
 –, references on 54

Greene 1, 122, 151, 162, 163, 200

Greenside 96

Grid of coordinate curves
 – for the cylindrical-toroidal (hybrid) system 70
 – for the “elementary” toroidal system 70
 – for the generalized cylindrical-toroidal system 71

Guggenheim 10, 54, 60

Hall 164

Hamada 81, 151

Hamada coordinate system 2, 150–155
 –, contravariant components of B in a 152

–, contravariant components of J in a 152
 –, Jacobian of a 151
 Hamada coordinates 98, 121, 150–155, 161
 – and the equilibrium conditions 151, 152
 –, mathematical definition 152, 153
 –, notation of 151
 –, references on 151
 Hamilton's equations for magnetic-field description 172
 Hamiltonian
 –, evaluation of the magnetic-field line ~ 174–175
 –, integrable (separable) 173
 –, magnetic-field line ~ 99, 170–174, 172
 –, theory, references on 166
Harris 161
Hazeltine 1, 79, 84, 92, 93, 108, 122, 163, 190, 198
 Helically symmetric stellarator 65
Hinton 79, 84, 190, 198
Hirshman 1, 36, 79, 119, 140, 161, 163, 164, 168, 198
Hitchon 68, 163, 179, 200
Hoffmann 54
Hogan 79
 Homogeneous solution
 – of a differential equation 133
 – of a magnetic differential equation 95
 Hybrid cylindrical-toroidal coordinate system 68

Ideal

–, constancy of the rotational transform in an ~ tokamak plasma 195–199
 –, dynamical equilibrium of an ~ tokamak plasma 186–200
 – MHD
 – – equilibrium conditions for toroidally confined plasmas 146–150
 – – – and the Hamada system 151, 152
 – – equilibrium equation (force balance) 65, 111, 146, 150
 – – plasma 56, 57, 79, 186, 187, 188, 189, 190, 199

Idempotent *see identity tensor*

Identities

–, dyadic 52–54
 –, tensor 52–54
 –, vector 38–39

Identity for tokamak geometry 186–187

Identity tensor 48–50

–, components of the 48
 –, relationship to the metric tensor 22, 48–50
 –, transformation properties of the 207–208

Independent vectors, linearly 8

Index

– as subscript 16
 – as superscript 16
 –, bracketed 9, 41
 – convention 16
 –, dummy 14
 –, repeated 14
 –, saturation over an 44, 45
 Inductive electric field in a tokamak 187
 Infinitesimal
 – element *see* differential element
 – volume element d^3R 85, 86
 – volume element dV 86
 Inner product *see* dot product
 Integrable Hamiltonian 173
 Integral, $\int dl/B$ 87–89, 201–203
 Intrinsic
 – derivative 32
 – differential 31
 Inverse matrix 32
 Irrational
 – number 65, 83
 – flux surfaces 65
 – –, relationship $\int dl/B$ and $dV/d\Psi$ on 202
 – –, rotational transform of 83
 Islands, magnetic 66, 148, 165, 166

Jackson 56

Jacobi's identity for the triple cross product 38 (V. 3)

Jacobian 10, 23–25
 – and the determinant of the metric tensor 25
 –, chain rule of differentiation for ~s 209
 –, flux-surface average of 151
 – for a flux-coordinate system in tokamaks 122, 123–124, 128
 – for the generic Clebsch system 104, 105
 – in a (general) toroidal flux coordinate system 122, 142, 143, 167
 – in an orthogonal system 25
 –, inverse 24
 – of a curvilinear system 23
 – of an inverse transformation 23
 – of a transformation 23
 – of the Boozer flux-coordinate system 143
 – of the Boozer-Grad coordinate system 109
 – of the Hamada coordinate system 151
 – written as a triple product 23, 24

Jancel 46, 54

Jardin 79

Johnson 1, 46, 52, 54, 122, 151, 162, 163, 164, 200

Kahan 46, 54

Kaiser 164

KAM theorem 165, 166

Kaper 46, 54

Kerst 170

Kieras 1, 122

Killeen 164

Kolmogorov-Arnold-Mozer (KAM) theorem 165, 166

Kovrizhnykh 13

Krall 46, 54

Kramer 2

Kronecker delta 22, 48–50

– as mixed component of the identity tensor 48
 –, transformation properties of 207–208

Kruskal 79, 93, 95, 100, 122, 164, 170

Kulsrud 79, 93, 95, 100, 101, 122, 163, 164, 170

Kuo-Petravic 161, 162, 166, 168, 169, 174, 175

1 = constant surfaces in toroidal systems 75

Labeling

– of magnetic-field lines 73, 75, 76
 – of magnetic-flux surfaces 77–82, 191–195, 199–200

Labels(s)

–, Clebsch stream function as
 – – field-line ~ 103
 – – flux-surface ~ 103
 – for magnetic-flux surfaces 70, 77–82, 116
 – –, poloidal magnetic flux as 81, 116
 – –, pressure as 81, 116
 – –, radially decreasing 81
 – –, radially increasing 81, 116
 – –, rotational transform as 116
 – –, safety factor as 116
 – –, toroidal magnetic flux as 81, 116
 – –, volume inside a flux surface as 81, 116
 –, magnetic-field line 73, 75, 76
 –, notations for
 – – field-line ~ 108
 – – flux-surface ~ 108
Lackner 163
 Lamé factors 13
Lao 163, 164
 Laplace development of a determinant 213
 Laplacian
 – in relation to the GSS magnetic differential operator 162
 – of a function (“definition”) 39 (V. 13)
 – of a vector (vector identity) 39 (V. 14)

Le Foll 79

Lee 105, 164
Leighton 132
Length
 – along a field line in terms of poloidal and toroidal flux-coordinate angles 156–157
 – “proper”, of a field line 148, 178, 201–203
see also arc length
Levi-Civita symbol(s) 50–52, 178
 – and the permutation tensor 50
 –, definition of the 27, 50
 –, transformation properties of 208–209
Levitated octupole 139
Lichtenberg 166, 170, 173
Lieberman 166, 170, 173
Lightfoot 7, 46, 52, 54
Line, see field line or magnetic-field line
 – of force *see* magnetic-field line
Linear combination 8
Linearly independent vectors 8
Literature *see* references
Littlejohn 166
Local basis 13
Local radius of curvature 58, 63
Long-thin-cylinder approximation in mirrors 164
Loop, closed $dI/B \sim$ integral 65
 – and flux-surface averaging 87–89
 – and the specific flux volume 201–203
 – as the “proper” length along a field line 148, 178, 201–203
Loop voltage 187, 193
Lorentz force 63
Low β plasma 59
Low pressure plasma 60
Lüst 163

Magnetic axis 66, 70, 77, 199, 200
 – and the vertical magnetic field 78, 200
 –, mathematical definition equation 81–82
Magnetic coordinates 73
 –, Clebsch-type 100–115
 –, generalized 165–175
see also magnetic-flux coordinates and flux coordinates
Magnetic differential equation 93–96
 –, constraints for the solution of 94, 147
 – for the Boozer coordinate transformation 144
 – for the Hamada coordinate transformation 153
 – for the metric coefficients of the Hamada system 154–155
 – for the periodic part of the current-density stream function 147, 149
 –, solvability conditions 94, 147

Magnetic differential operator for the GSS
 – equation 162
Magnetic-field line
 – as coordinate curve 60, 61
 – as elastic cord 63
 – curvature 58–62, 63
 –, definition 55
 –, diffusion equation for 56
 –, equation of a 55–56
 –, differential form 55, 56, 157, 169, 172, 201
 –, in canonical coordinates 172
 –, in Clebsch-type coordinates 107
 –, in flux coordinates 84, 120, 201
 –, integrated form 117, 168, 173
 –, in Clebsch-type coordinates 107
 –, in flux coordinates 87, 119, 120
 – Hamiltonian 99, 170–174, 172
 –, evaluation of 174–175
 – label 73, 75, 76
 –, Clebsch stream function as 103
 –, motion of 56–58
 –, principal normal to a 58, 61
 –, “proper length” of 148, 178
 – and the specific flux volume 201–203
 –, straight ~ coordinates in toroidal systems 116–122
see also flux coordinates
 – tension 63–64
see also field line
Magnetic-field structure related concepts 55–96
Magnetic field (vector)
 – as the gradient of the scalar magnetic potential 114, 158
 – components
 –, contravariant, in toroidal flux coordinates 120, 131, 167
 –, covariant 122
 –, in toroidal flux coordinates 133–140
 –, poloidal 122, 200
 –, in tokamaks 127, 186, 189
 – representation
 –, Boozer-Grad 110, 112, 141
 –, Clebsch 100, 103, 105, 108, 113, 117, 121, 123, 126, 152, 156, 157
 –, contravariant 104, 105, 108, 113, 116, 142, 150, 167, 174
 –, covariant 104, 105, 109, 110, 111, 112, 140, 141, 142
 –, “usual” ~ in tokamaks 124, 125, 126, 127, 128, 186
 –, toroidal 122
 –, in tokamaks 124, 127, 186, 189
 –, vertical 78, 200

Magnetic flux
 –, axial 77, 82
 –, azimuthal 77, 82
 – density 119
 –, form-invariant expression for 79–80, 118, 119
 –, frozen ~ theorem 56–58
 –, poloidal 77, 78, 79, 80, 81
 –, as label for flux surfaces 81, 116, 191–193, 199–200
 –, toroidal 77, 80, 81
 –, as label for flux surfaces 81, 116, 193–195
 –, transformer ~ 186, 187, 188, 189, 190, 199
Magnetic-flux coordinates
 –, definition of 1, 98
 –, establishment of 160–164, 166–170
 –, illustrative example 73
 –, overview of 98–99
see also flux coordinates
Magnetic (-flux) surfaces 61, 64–67, 77
 – as constant pressure surfaces 65
 – as coordinate surfaces 70
 –, definition of 64, 65, 77, 81
 –, degenerate 66
 –, existence of 65, 165–166
 – in open-ended systems 67
 – in stellarators 66
 – in tokamaks 66
 – in toroidal systems 65–66
 –, irrational 65
 – labeling 70, 77–82
 –, motion of 79, 190–195
 –, rational 65
 –, vacuum 65
see also flux surface and flux surfaces
Magnetic induction 119
Magnetic islands 66, 148
Magnetic mirror 67
see also open-ended systems
Magnetic pressure 63–64
Magnetic scalar potential 108, 133, 138–139, 158
 – and coil currents 134, 138–139
 –, B as the gradient of 114, 158
 –, mathematical 135
 –, multiple-valued 134, 139
 –, periodic part of 134, 139, 142
 –, physically acceptable 134, 139
Magnetic tension 63–64
Magnetic vector potential 170–171
 – and the canonical form of B 171
 – as stream function for the poloidal field 127
 – for (general) toroidal flux coordinate systems 131–133
 –, generic covariant form 170
 – in tokamaks 126, 127
 –, physical versus mathematical 132
 –, single-valued (physical) 132, 133
Magnetically confined plasmas, curvilinear coordinates for the description of 67–76
Magnitude of a vector 26
Major
 – axis 66, 69
 – radius 69
Margenau 7, 54, 183
Marsden 54
Mathews 7, 10, 32, 54, 206
Matrix
 – “belonging” to a dyad 42
 – cofactor of a 22
 – constructed from the components of a dyad 42
 –, inverse 22
 – product 22, 25
McNamara 164
Meiss 1, 93, 122, 163
Melcher 54
Menzel 54, 206
Metric coefficients 13, 20–22
 – conjugate 20
 –, covariant derivative in terms of 32–34
 –, derivative of, and Christoffel symbols 213–214
 – for the “proper” toroidal system (scale factors) 183
 – in orthogonal systems 20
 –, reciprocal 20
 – related to the scale factors h_i 20, 28
 –, transformation properties of 208
Metric tensor 48–50
 –, Jacobian related to the determinant of the 25
 –, relationship to the fundamental tensor 48–50, 208
 –, relationship to the identity tensor 48–50, 208
 –, transformation properties of the components of 208
MHD, ideal 56–57, 79
 – equilibrium conditions for toroidally confined plasmas 146–150
 –, and the Hamada system 151, 152
 – equilibrium equation (force balance) 65, 111, 146, 150
 – plasma 56, 57, 79, 186, 187, 188, 189, 190, 199
Milne-Thomson 7, 46, 54, 100

Minor
 – axis 66, 69
 – radius 69
Mirror-type devices 67
 –, flux surface average in 89, 91–93
 –, magnetic-surface labeling 82
 –, magnetic surfaces 67
see also open-ended devices
Misner 54
Mixed components
 – of a dyad 42
 – of a second-order tensor 47
 – of the fundamental tensor 48–50
 – of the identity tensor 48–50
 – of the metric tensor 48–50
 –, relationship to other tensor components 47–48
 –, transformation properties of 207
Miyamoto 1, 122
Momentum, canonical, for magnetic-field description 172
Morozov 1, 65, 82, 84, 162
Morse 2, 7, 34, 43, 46, 54, 179
Motion
 – of flux surfaces
 – defined by the poloidal disk flux 191–192
 – defined by the poloidal ribbon flux 192–193
 – defined by the toroidal flux 193–195
 – relative to an ideal MHD plasma 79
 – relative to a non-ideal MHD plasma 79
 – of magnetic field lines 56–58
 – of plasma and flux surfaces 190–195
 – of the plasma due to $E \times B$ 190–191, 192, 193
Mukhovatov 200
Multiple-valued
 – angles 70, 71, 116
 – functions 70, 71, 116
Murphy 7, 54, 183
Myra 91, 92

Nabla operator *see* del operator
Natural coordinates 72, 151
see also Hamada coordinates
Near-integrable systems 173
Nelson 164
Neoclassical
 – banana regime 198
 – (Ware) pinch 198
Newcomb 57, 58, 93, 95, 164
Nonion form
 – of a dyad 43
 – of a dyadic 43

– of a second-order tensor 43
Non-uniqueness of flux coordinates
 121
Normal
 – curvature 60, 61, 62
 – curvature vector 61
 –, principal 58, 61, 62, 63
Notation
 –, bracketed, for indices 9, 41
 –, consistency of 16
 – for Christoffel symbols 34
 – for current coordinates 129
 – for dyads/dyadics 3, 41
 – for field-line labels 108
 – for flux-coordinate angles 73, 120
 – for flux-surface labels 108
 – for Hamada coordinates 151
 – for tensors 3, 41
 – for the contravariant components of a dyad/dyadic/tensor 42
 – for the contravariant components of a vector 16
 – for the covariant components of a dyad/dyadic/tensor 42
 – for the covariant components of a vector 16
 – for the covariant derivative 32
 – for the mixed components of a dyad/dyadic/tensor 42
 – for the poloidal angle in the “elementary” system 70
 – for the toroidal angle in the “elementary” system 70
 – for the unit tangent vector to a magnetic-field line 55, 58, 104
 – for symmetry coordinates in a tokamak 123
 – for unit vectors 9
 – for vectors 3
 –, importance of 7
 –, Russian, for the dot product 41
 –, Russian, for the vector product 41
 –, subscript 16
 –, summary table for Boozer’s, Solov’ev and Shafranov’s and our 142
 –, superscript 16
NRL plasma formulary 38
numerical
 – evaluation of the field-line Hamiltonian 174–175
 – field-line integration 168–170
 – methods
 – for flux-surface computation (references) 163–164
 – for plasma-equilibrium computation (references) 163–164

Oberman 163
Object(s)
 – contravariant 16, 206
 – covariant 16, 206
Octupole, levitated 139
Odd permutation 15
Ohm’s law 56
Open-ended devices (mirrors)
 – Clebsch coordinates in 76
 – curvilinear coordinates for 67, 76
 –, flux-surface average in 89, 91–93
 –, magnetic-surface labeling 82
 –, magnetic surfaces in 67
Orthogonal system(s) 11, 13
 –, differential area element in a coordinate surface 29
 –, partially, in tokamaks 123
Outer product 41
Overview 2
 – of Chapter 2 7
 – of part I 6
 – of part II 98–99
 – of part III 178
Panofsky 139
Pao 79
Parallel
 – component of a vector 37–38
 – del operator 59
 – electric field
 – due to a tokamak transformer 187, 198
 – in an ideal plasma 189
Parallelepiped, volume of 8
Paraxial approximation in mirrors 164
Particular solution
 – of a differential equation 133
 – of a magnetic differential equation 95
Pearlstein 164
Pereversev 79
Perfect tori 67, 69
Perfectly conducting plasma 56, 57
Periodic
 – magnetic scalar potential 134, 139, 142
 – magnetic vector potential 132, 133
Permutation
 –, cyclic 15
 –, even 15
 –, odd 15
Permutation symbol *see* Levi-Civita symbol
Permutation tensor 51
 – and the Levi-Civita symbol 51
 –, components of 52
 –, definition 51
Perpendicular
 – component of a vector 37–38
 – conductivity 190
 – del operator 59
 – electric field in an ideal tokamak plasma 189–190
Phase space, extended 173
Phillips 139
Physical component 126
Pinch velocity
 –, neoclassical (Ware) 198
 –, radial 190, 192, 194, 198
Plasma
 –, dynamic equilibrium of a tokamak ~ 178, 186–200
 – equilibrium time evolution, references 79
 – fluid velocity 56
 –, ideal MHD ~ 56, 57, 79, 186, 187, 188, 189, 190, 199
 –, low β 59
 –, low pressure 60
 – motion due to $E \times B$ 58, 190–191, 192, 193
 –, perfectly conducting 56, 57
Polar coordinates 179, 181–182
 –, constant coordinate curves 179, 182
Poloidal
 – angle 69, 71
 – angle, generalized 71–72
 – angle-like variable 70
 – cross section 71
 – current
 – density flux (=current) 129
 –, plasma 140
 –, total 140
 – direction 69
 – disk flux for labeling flux surfaces 77, 78, 79, 81, 116, 191–192, 199–200
 – electric field in an ideal tokamak plasma 189, 190
 – flux-coordinate angle in tokamaks 128
 – magnetic field 122, 200
 – in tokamaks 127, 186, 189
 –, stream function for 127
 – magnetic flux 77, 78, 79, 81
 – as flux-surface label 81, 116, 191–193, 199–200
 –, form-invariant expression for 80, 119
 – plane 69
 – ribbon flux for labeling flux surfaces 77, 78, 79, 80, 81, 192–193, 199–200
 – transit 65
Polyad 44
Polyadic 44
Pomphrey 161, 168
Position vector 10, 11, 14, 16, 17
 –, divergence of 39 (V. 17)

Potential
 - magnetic scalar 108, 133, 138–139, 158
 -- and coil currents 134, 138–139
 --, B as the gradient of 114, 158
 --, mathematical 135
 --, multiple-valued 134, 139
 --, periodic part of 134, 139, 142
 --, physically acceptable 134, 139
 -, magnetic vector 170–171
 -- and the canonical form of B 171
 -- as stream function for the poloidal field 127
 -- for (general) toroidal flux-coordinate systems 131–133
 --, generic covariant form 170
 -- in tokamaks 126, 127
 --, physical versus mathematical 132
 --, single-valued (physical) 132, 133
 Pressure
 - as flux-surface label 77, 81, 82, 116
 -, constant- \sim surfaces 65
 -, magnetic 63–64
 Principal normal to a field line 58, 61, 62, 63
 Product
 -, cross 27
 -, dot 26
 -- of a vector and a tensor 44
 -, double-dot 46
 -, exterior 41
 -, outer 41
 -, scalar, *see* dot product
 -, tensor 41
 -, vector *see* cross product
 "Proper length" of a field line 148, 178
 - and the specific flux volume 201–203
 "Proper" toroidal coordinates 3, 68, 178, 179–185
 - and bipolar coordinates 182–183
 -, scale factors h_i 183
 Pseudo toroidal coordinate system 68
Pytte 163
 Quadrupole symmetry and end losses in mirrors 92
 Quaternions 2
 Radial pinch velocity 190, 192, 194, 198
 Radially increasing/decreasing flux-surface labels 81, 116
 Radius
 -, major 69
 -, minor 69
 - of curvature 58, 63
 - of curvature vector 58
 Rank of a tensor 45

Rational
 - flux surfaces 65
 --, current-closure condition on 65
 --, relationship of $\oint dl/B$ and $dV/d\Psi$ on 202
 --, rotational transform on 83
 - number 65, 83
 Reaction current in an ideal tokamak plasma 189, 191, 194
 "Real" toroidal coordinates *see* "proper" toroidal coordinates
 Reciprocal
 - basis vectors 14–15
 --, transformation properties of 204–205
 - sets of vectors 8–9
 --, "back-and-forth" relationships of 8
 Reference list 215–219
 References
 - on differential forms and Grassmann algebra 54
 - on differential geometry 10
 -- and tensor calculus 54
 - on flux coordinates 122, 161, 162
 -- and magnetic geometry 1
 - on flux-surface computation 162–164
 - on Hamada coordinates 151
 - on Hamiltonian theory 166
 - on plasma-equilibrium time evolution 79
 - on tensors, dyads and dyadics 54
 - on the existence of flux surfaces 166
 - on the Grad-Schlüter-Shafranov equation 163
 - on the Hamiltonian character of B 170
Reiman 96, 161, 164, 166, 168
 Repeated index 14
 Representation of A (magnetic vector potential)
 - covariant
 --, generic 132, 170
 -- in flux coordinates 133
 Representation of B
 - Boozer-Grad
 --, "final" form 110, 112, 141, 157
 --, "intermediate" form 109, 111, 112
 -, Clebsch 100, 103, 105, 108, 113, 117, 156, 157
 -- in Hamada coordinates 152
 -- in tokamaks 121, 123, 126
 -- in toroidal flux coordinates 121, 156, 166, 167
 - contravariant
 -- as a Clebsch form 105, 108
 --, generic 104, 116, 174
 -- in a Clebsch coordinate system 104, 105, 108, 113
 -- in Boozer-Grad coordinates 113

-- in toroidal flux coordinates 142, 150, 166, 167
 -, covariant
 --, generic 104, 105, 140
 -- in "final" Boozer-Grad form 110, 112, 141, 157
 -- in "intermediate" Boozer-Grad form 109, 111, 112
 -- in (ρ, β, l) coordinates 105, 109
 -- in toroidal flux coordinates 142
 -, "usual", in tokamaks 124, 125, 126, 127, 128, 186
 Representation of J , Clebsch
 - generic 112, 129, 141
 - in current coordinates 130
 - in Hamada coordinates 152
 Resistivity, evolution in a finite plasma 198–199
 Review papers
 - on flux coordinates 122, 161, 162
 -- and magnetic geometry 1
 - on Hamada coordinates 151
 - on plasma equilibrium evolution 79
 - treating the Grad-Schlüter-Shafranov equation 163
see also references
 Ribbon
 - as shape for a $\beta =$ constant coordinate surface 75
 -, magnetic flux through (poloidal ribbon flux) 77, 78, 79, 80, 81
 Right-handed coordinate systems 68
Rome 161, 168
Rosenbluth 164, 170
 Rot *see* curl
 Rotational transform 81, 82–84
 - as flux-surface label 116
 - as "frequency" in the action-angle Hamiltonian picture 173
 - as the slope of a field line in flux coordinates 120
 -, constancy of, in an ideal tokamak plasma 195–198
 - definition of, for time-dependent problems 83, 195
 Rotational-transformation angle 82
Rubin 100, 101, 163
 Ruled surface 87, 113
 Safety factor 81, 83, 84, 196
 - as flux-surface label 116
 - time evolution of, in an ideal tokamak plasma 196–198
Sagdeev 170
Sakharov 56, 190, 200
Sands 132

Saturation over an index 22, 44, 45
 Scalar
 - as a zeroth-order tensor 44
 - product *see* dot product
 - triple product 38 (V. 1)
 Scalar potential, magnetic 108, 133, 138–139, 158
 - and coil currents 134, 138–139
 -, B as the gradient of 114, 158
 -, mathematical 135
 -, multiple-valued 134, 139
 -, periodic part of 134, 139, 142
 -, physically acceptable 134, 139
 Scale factors h_i 13, 20, 28
 - for the "proper" toroidal system 183
Schlüter 163, 164
Schlüter-Grad-Shafranov-equation *see* Grad-Schlüter-Shafranov equation
Schmidt 56, 57, 63, 64
Schutz 54
Schwenn 164
 Second-order tensor(s) 40–50
 - and dyadics 43
 - identities 52–54
 -, nonion form of 43
 -, "simple" (dyads) 40–44, 46–47, 52–54, 55
 Secular term(s) 111, 112, 118, 128, 132, 133–135, 139
 Separable Hamiltonian 173
Seshadri 54
Shafranov 1, 62, 66, 68, 83, 93, 108, 122, 140, 149, 151, 153, 161, 162, 163, 164, 179, 200, 201
Shafranov-Grad equation *see* Grad-Schlüter-Shafranov equation
Shafranov shift 73, 123
 shear 84
Shkarofsky 46, 52, 54
Sigmar 79, 119, 198
 Single valued
 - functions 72
 - magnetic scalar potential 134–135, 139–140
 - magnetic vector potential 132
Snider 15
Solov'ev 1, 62, 65, 66, 82, 83, 84, 93, 108, 122, 133, 151, 153, 161, 162, 163, 201
 Solution of a magnetic differential equation 95
 Solvability conditions for a magnetic differential equation 94, 147
 Space
 - 3-D ~ 17
 - dual vector ~ 17, 19
 - Euclidean 17, 19, 48

- Space (cont.)
 - extended phase ~ 173
 - of contravariant vectors 19
 - of covariant vectors 19
 - (ordinary) 3-D configuration ~ 2
 Specific flux volume 178
 - and the "proper length" of a field line 201–203
 - related to the line integral of dl/B 148, 201–203
 Spherical coordinate system 11, 181–182
 - zero off-diagonal metric coefficients in 20
 Spherical coordinates 179, 181–182
 - coordinate surfaces 182
Spiegel 7, 10, 34, 54
Spies 164
Spitzer 170
Spivak 10, 54
Springer 206
Stacey 7, 34, 163, 198, 200
 Stellarator 65
 - Clebsch representation of B in 121
 - coil currents in 139
 - expansion 163
 - magnetic surfaces in a 66
Stewart 7, 46, 52, 54
 Stochastic behavior of field lines 165–166
Stokes' law 57
Stoker 10, 15, 54
 Straight
 - B systems 1, 2
 - current-density-line coordinates 73, 98, 129–131
 - magnetic-field-line coordinates 72–73, 98, 116–122
Stratton 7, 54
 Stream function, Clebsch 100, 102, 107
 - field-line labeling function as 103
 - flux-surface labeling function as 103
 - for the current density, (λ or η) 112, 129, 141, 150
 - differential equation for the periodic part of 146–149
 - for the magnetic field, (β or v)
 -- in terms of angular variables in a tokamak 128
 -- in terms of arbitrary angle variables 117–119
 -- in terms of Boozer coordinates 144
 -- in terms of flux-coordinate angles 120, 131, 156
 Stream function, magnetic vector potential as, for the poloidal field 127
 Stream line 100
 Subscript (notation) 9, 11, 16

- Sum of dyad form
 - of the fundamental tensor 49
 - of the identity tensor 49
 - of the metric tensor 49
 Summation convention 14
 Superscript (notation) 9, 11, 16, 17–18
 Surfaces
 - constant pressure 65
 - coordinate 10, 11, 67
 -- for the "proper" toroidal system 183
 -- for the spherical system 182
 - cylindrical 64
 - flux 61, 64–67
 - irrational 65
 - magnetic 61, 64–67
 - rational 65
 - ruled 87
 - toroidal 64
 Symbol(s)
 - Christoffel 34–35, 52, 178, 209–210
 - Kronecker-delta 22, 48–50
 - Levi-Civita 50–52, 178, 208–209
 Symmetry
 - and the existence of flux surfaces 65, 165–166
 - angle as toroidal angle in tokamaks 120, 121, 122, 186
 - of Christoffel symbols 34
 - of metric coefficients 20
 - quadrupole, and end losses in mirrors 92
Symon 7, 54, 64
 Systems, coordinate *see* coordinate systems
- Tangent basis vectors 10, 11–13
 - subscript notation for 13
 - transformation of 204–205
 - unit ~ 13
 Tangent vector to a field line 55, 58
Tarapov 7, 17, 54
Taylor 170
 Tension
 - in elasticity theory 64
 - magnetic 63–64
 Tensor
 - analysis 40
 - components
 -- of a second-order 47
 -- of a third-order 50
 -- relationship between the various 47–48
 - transformation properties of 178, 204–210, 207
 -- for the identity tensor 207–208
 -- for the metric tensor 208
 -- for the permutation tensor 209
 - contraction 45

- , fundamental 48–50, 208
 - identities 52–54
 - identity 48–50, 207
 - metric 48–50, 208
 - product 41
 - rank of a 45
 Tensors 40–54
 - "basis ~" 41
 - double-dot product of 46
 - dyadics as 43
 - dyads as "simple" ~ 40–44, 46–47, 52–54, 55
 - first-order 44
 - fourth-order 44
 - literature on 54
 - n-th order 44
 - nonion form of 43
 - polyadic as 44
 - polyads as "simple" ~ 44
 - references on 54
 - second-order 40–50
 -- "simple" ~ (dyads) 40–44, 46–47, 52–54, 55
 -- written as a sum of at least three dyads 42
 - tetradiques as 44
 - tetrads as "simple" ~ 44
 - third-order 44, 50
 -- "simple" ~ (triads) 44
 - triadic as 44
 - triads as "simple" ~ 44
 - zeroth-order 44
see also dyadics and dyads
 Tensorial objects 40–54
 Tetrad 44
 Tetric 44
 Textbook treatments
 - of differential forms and Grassmann algebra 54
 - of differential geometry and curvilinear coordinates 10, 54
 - of flux coordinates and magnetic geometry 1
 - of plasma-equilibrium evolution 79
 - of quaternions 2
 - of tensors, dyadics and dyads 54
 - of the Grad-Schlüter-Shafranov equation 163
 - of vectors and tensors (in curvilinear coordinate systems) 7, 54
 Third-order tensor 44, 50
Thorne 54
 Time evolution of an ideal tokamak plasma equilibrium 186–200
 Tokamak
 - dynamic equilibrium 79, 178, 186–200

- , geometrical identity in a 186–187
 - magnetic-flux surfaces in a 66
 - non-flux-coordinates in a 126–128
 - representation of B in a
 -- Clebsch 121, 123, 126
 -- flux-coordinate 121, 123, 126
 --, "usual" tokamak form 124, 125, 127, 186
 - simple example of flux coordinates in a 73
 - symmetry flux coordinates in a 122–126
 - the electric field in an ideal ~ plasma 187–190
 Tori, ideal or perfect 67, 69
 Toroidal
 - angle 69, 71
 - angle, generalized 71–72
 - angle-like variable 71
 - direction 66, 69, 146
 - electric field in an ideal tokamak plasma 189, 190
 - transit 65
 - winding number 83
 Toroidal Boozer-flux-coordinate system 2, 142–146
 Toroidal confinement systems
 - canonical coordinates in 165–175
 - curvilinear coordinates in 67–76
 - flux coordinates in 116–155
 - flux surface average in 86–89, 90–93
 - ideal MHD equilibrium conditions in 146–150
 - magnetic-surface labeling in 77–82
 - magnetic surfaces in 65–66
 - rotational transform in 82–84
 Toroidal coordinate system(s)
 - "elementary" 67, 68–70, 128
 - for toroidal confinement systems 68–76
 - hybrid cylindrical- ~ 67, 68–70
 - "pseudo" 68
 - zero off-diagonal metric coefficients in 20
 Toroidal coordinates
 - "proper" 3, 68, 178, 179–185
 -- and bipolar coordinates 182–183
 --, scale factors h_i 183
 Toroidal current-density flux (or current) 129
 Toroidal flux coordinates 98, 116–155
 - Boozer coordinates 98, 142–146
 - conversion from Clebsch coordinates to 156–159
 - Hamada coordinates 150–155
 - mathematical definition of 120
 - symmetry ~ in tokamaks 122–126

- Toroidal magnetic field 122
 - in tokamaks 124, 127, 186, 189
 Toroidal magnetic flux 77
 - as flux-surface label 77, 80, 81, 116, 193–195
 - form-invariant expression for 79–80, 118, 196
 Toroidal straight B coordinates 2, 116–155
 Toroids 67
 Torsatron, coil currents in 139
 Torsion 62
 Transform, rotational 81, 82–84
 - as flux-surface label 116
 - as “frequency” in the action-angle Hamiltonian picture 173
 - as the slope of a field line in flux coordinates 120
 - constancy of, in an ideal tokamak plasma 195–198
 - definition of, for time-dependent problems 83, 195
 Transformation
 - angle, rotational 82
 - Clebsch 100
 - Curvilinear 10
 - invertible 10
 - Jacobian of a 23
 - one-to-one 10
 Transformation equation
 - for alternative flux coordinates 121
 - for current coordinates 130
 - for flux coordinates 120
 - for Hamada coordinates 152–153
 - for toroidal Boozer coordinates 143–144
 Transformation, flux-coordinate
 - establishment of 98–99, 168–170
 - mathematical definition of 120
 - summary of 99, 160–164
 Transformation properties
 - of basis vectors 19, 204–205
 - of tensor components 178, 207
 - of the Christoffel symbols 209–210
 - of the Kronecker delta 207–208
 - of the Levi-Civita symbols 208–209
 - of the metric coefficients 208
 - of the vector derivative 31
 - of vector components 178, 205–206
 Transformer (in a tokamak) 186, 187, 188, 189, 190, 191, 193, 198, 199
 Triad
 - as synonym for trihedron 62
 - as third-order “simple” tensor 44
 Triadic 44
 Trihedral, Frenet-Serret 62
 Trihedron 8
 - Frenet-Serret 62, 63

- Triple product
 -, Jacobian written as a 23, 24
 - of vectors
 - -, scalar ~ 8, 38 (V. 1)
 - -, vector ~ (bac-cab rule) 38 (V. 2)
 - written as the determinant of a matrix 23
Trivelpiece 46, 54
- Unit tangent vector 13
 - to a magnetic-field line 55, 58
 - -, notation for 55, 58, 104
 Unit vector
 - along B 55, 58
 - -, notation for 55, 58, 104
 - in a Cartesian coordinate system 9
 - -, notation for 9
- Vacuum magnetic surfaces 65
 Van Bladel 46, 54
 Vector
 - contravariant components of a 16
 - covariant components of a 16
 - curl of a 36–37
 - curvature ~ 58, 59, 60, 61
 - derivative and its transformation properties 31
 - differentiation
 - - and Christoffel symbols 34–35
 - - in curvilinear coordinates 30–37
 - divergence of a 36
 - gradient ~ 35–36
 - identities 38–39
 - Laplacian of a (vector identity) 39 (V. 14)
 - magnitude of a 26
 - notation 3
 - parallel component of a 37–38
 - perpendicular component of a 37–38
 - position ~ 10, 11, 14, 17
 - product see cross product
 - radius of curvature ~ 58
 - transformation of ~ components 178, 205–207
 - triple product (bac-cab rule) 38 (V. 2)
- Vector potential, magnetic 170–171
 - and the canonical form of B 171
 - as stream function for the poloidal field 127
 - for (general) toroidal flux-coordinate systems 131–133
 - generic covariant form 170
 - in tokamaks 126, 127
 - physical versus mathematical 132
 - single-valued (physical) 132, 133

- Vector space
 -, dual 17, 19
 -, Euclidean 17, 19
 - of contravariant vectors 19
 - of covariant vectors 19
 Vectors 7–39
 - as first-order tensors 40, 44
 - basis ~ 8, 10, 15
 - -, transformation properties 204–205
 - contravariant ~ 17
 - contravariant basis ~ 19
 - covariant ~ 17
 - covariant basis ~ 19
 - cross product between 27
 - dot product between 26
 - identities involving tensors and 52–54
 - reciprocal basis ~ 14–15
 - scalar triple product of 8, 23
 - tangent basis ~ 11
 Velocity
 - fluid ~ of the plasma 56
 - flux-surface ~ 192, 193, 195, 198
 - radial inward pinch ~ 190, 192, 194, 198
 Vertical magnetic field 78, 200
 Voltage, loop 187, 193
 Volume element
 - differential ~ 29
 - infinitesimal ~ d^3R 85, 86
 - infinitesimal ~ dV 86

- Volume inside a flux surface
 - as flux-surface label 77, 81, 82, 116
 - in the Hamada system 151
 - mathematical expression for 81
 Volume, specific flux 178
 - and the “proper length” of a field line 201–203
 - related to the line integral of dl/B 148, 201–203
- Walker* 7, 10, 32, 54, 206
 Ware pinch 198
 Watson 164
 Weimer 163, 200
 Wheeler 54
 White 1, 108, 163, 166
 Whiteman 166
 Whitson 164
 Wieland 163
 Wimmel 52
 Winding number, toroidal 83
 Window function 161
 Woodson 54
 Wrede 7, 19, 54
- Zakharov 163, 164, 200
 Zaslavski 170
 zeroth-order tensor 44
 Zhilinsky 56, 190, 200