Pavel Grinfeld

Introduction to Tensor Analysis and the Calculus of Moving Surfaces

Springer

Introduction to Tensor Analysis and the Calculus of Moving Surfaces

Pavel Grinfeld

Introduction to Tensor Analysis and the Calculus of Moving Surfaces

Springer

Pavel Grinfeld
Department of Mathematics
Drexel University
Philadelphia, PA, USA

ISBN 978-1-4614-7866-9 ISBN 978-1-4614-7867-6 (eBook)
DOI 10.1007/978-1-4614-7867-6
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013947474

Mathematics Subject Classifications (2010): 4901, 11C20, 15A69, 35R37, 58A05, 51N20, 51M05, 53A05, 53A04

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The purpose of this book is to empower the reader with a magnificent new perspective on a wide range of fundamental topics in mathematics. Tensor calculus is a language with a unique ability to express mathematical ideas with utmost utility, transparency, and elegance. It can help students from all technical fields see their respective fields in a new and exciting way. If calculus and linear algebra are central to the reader's scientific endeavors, tensor calculus is indispensable. This particular textbook is meant for advanced undergraduate and graduate audiences. It envisions a time when tensor calculus, once championed by Einstein, is once again a common language among scientists.

A plethora of older textbooks exist on the subject. This book is distinguished from its peers by the thoroughness with which the underlying essential elements are treated. It focuses a great deal on the geometric fundamentals, the mechanics of change of variables, the proper use of the tensor notation, and the interplay between algebra and geometry. The early chapters have many words and few equations. The definition of a tensor comes only in Chap. 6—when the reader is ready for it.

Part III of this book is devoted to the calculus of moving surfaces (CMS). One of the central applications of tensor calculus is differential geometry, and there is probably not one book about tensors in which a major portion is not devoted to manifolds. The CMS extends tensor calculus to *moving manifolds*. Applications of the CMS are extraordinarily broad. The CMS extends the language of tensors to physical problems with moving interfaces. It is an effective tool for analyzing boundary variations of partial differential equations. It also enables us to bring the calculus of variations within the tensor framework.

While this book maintains a reasonable level of rigor, it takes great care to avoid a formalization of the subject. Topological spaces and groups are not mentioned. Instead, this book focuses on concrete objects and appeals to the reader's geometric intuition with respect to such fundamental concepts as the Euclidean space, surface, length, area, and volume. A few other books do a good job in this regard, including [2, 8, 31, 46]. The book [42] is particularly concise and offers the shortest path to the general relativity theory. Of course, for those interested in relativity, Hermann

Weyl's classic *Space, Time, Matter* [47] is without a rival. For an excellent book with an emphasis on elasticity, see [40].

Along with eschewing formalism, this book also strives to avoid vagueness associated with such notions as the infinitesimal differentials dx^i. While a number of fundamental concepts are accepted without definition, all subsequent elements of the calculus are derived in a consistent and rigorous way.

The description of Euclidean spaces centers on the basis vectors $\mathbf{Z}_i$. These important and geometrically intuitive objects are absent from many textbooks. Yet, their use greatly simplifies the introduction of a number of concepts, including the metric tensor $Z_{ij} = \mathbf{Z}_i \cdot \mathbf{Z}_j$ and Christoffel symbol $\Gamma^i_{jk} = \mathbf{Z}^i \cdot \partial \mathbf{Z}_j / \partial Z^k$. Furthermore, the use of vector quantities goes a long way towards helping the student see the world in a way that is independent of Cartesian coordinates.

The notation is of paramount importance in mastering the subject. To borrow a sentence from A.J. McConnell [31]: "The notation of the tensor calculus is so much an integral part of the calculus that once the student has become accustomed to its peculiarities he will have gone a long way towards solving the difficulties of the theory itself." The introduction of the tensor technique is woven into the presentation of the material in Chap. 4. As a result, the framework is described in a natural context that makes the effectiveness of the rules and conventions apparent. This is unlike most other textbooks which introduce the tensor notation in advance of the actual content.

In spirit and vision, this book is most similar to A.J. McConnell's classic *Applications of Tensor Calculus* [31]. His concrete no-frills approach is perfect for the subject and served as an inspiration for this book's style. Tullio Levi-Civita's own *The Absolute Differential Calculus* [28] is an indispensable source that reveals the motivations of the subject's co-founder.

Since a heavy emphasis in placed on vector-valued quantities, it is important to have good familiarity with geometric vectors viewed as objects on their own terms rather than elements in $\mathbb{R}^n$. A number of textbooks discuss the geometric nature of vectors in great depth. First and foremost is J.W. Gibbs' classic [14], which served as a prototype for later texts. Danielson [8] also gives a good introduction to geometric vectors and offers an excellent discussion on the subject of differentiation of vector fields.

The following books enjoy a good reputation in the modern differential geometry community: [3, 6, 23, 29, 32, 41]. Other popular textbooks, including [38, 43] are known for taking the formal approach to the subject.

Virtually all books on the subject focus on applications, with differential geometry front and center. Other common applications include analytical dynamics, continuum mechanics, and relativity theory. Some books focus on particular applications. A case in point is L.V. Bewley's *Tensor Analysis of Electric Circuits And Machines* [1]. Bewley envisioned that the tensor approach to electrical engineering would become a standard. Here is hoping his dream eventually comes true.

Philadelphia, PA Pavel Grinfeld

Contents

Chapter 1
Why Tensor Calculus?

"Mathematics *is* a language." Thus was the response of the great American scientist J. Willard Gibbs when asked at a Yale faculty meeting whether mathematics should really be as important a part of the undergraduate curriculum as classical languages.

Tensor calculus is a specific language within the general language of mathematics. It is used to express the concepts of multivariable calculus and its applications in disciplines as diverse as linear algebra, differential geometry, calculus of variations, continuum mechanics, and, perhaps tensors' most popular application, general relativity. Albert Einstein was an early proponent of tensor analysis and made a valuable contribution to the subject in the form of the *Einstein summation convention*. Furthermore, he lent the newly invented technique much clout and contributed greatly to its rapid adoption. In a letter to Tullio Levi-Civita, a co-inventor of tensor calculus, Einstein expressed his admiration for the subject in the following words: "I admire the elegance of your method of computation; it must be nice to ride through these fields upon the horse of true mathematics while the like of us have to make our way laboriously on foot."

Tensor calculus is not the only language for multivariable calculus and its applications. A popular alternative to tensors is the so-called *modern language of differential geometry*. Both languages aim at a geometric description independent of coordinate systems. Yet, the two languages are quite different and each offers its own set of relative strengths and weaknesses.

Two of the greatest geometers of the twentieth century, Elie Cartan and Hermann Weyl, advised against the extremes of both approaches. In his classic *Riemannian Geometry in an Orthogonal Frame [5]* (see also, [4]), Cartan recommended to"as far as possible avoid very formal computations in which an orgy of tensor indices hides a geometric picture which is often very simple." In response, Weyl (my personal scientific hero) saw it necessary to caution against the excessive adherence to the coordinate free approach [47]: "In trying to avoid continual reference to the components, we are obliged to adopt an endless profusion of names and symbols in addition to an intricate set of rules for carrying out calculations, so that the balance

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_1,

of advantage is considerably on the negative side. An emphatic protest must be entered against these orgies of formalism which are threatening the peace of even the technical scientist."

It is important to master both languages and to be aware of their relative strengths and weaknesses. The ultimate choice of which language to use must be dictated by the particular problem at hand. This book attempts to heed the advice of both Cartan and Weyl and to present a clear geometric picture along with an effective and elegant analytical technique that is tensor calculus. It is a by-product of the historical trends on what is in fashion that tensor calculus has presently lost much of its initial popularity. Perhaps this book will help this magnificent subject to make a comeback.

Tensor calculus seeks to take full advantage of the robustness of coordinate systems without falling subject to the artifacts of a particular coordinate system. The power of tensor calculus comes from this grand compromise: it approaches the world by introducing a coordinate system at the very start—however, it never specifies which coordinate system and never relies on any special features of the coordinate system. In adopting this philosophy, tensor calculus finds its golden mean.

Finding this golden mean is the primary achievement of tensor calculus. Also worthy of mention are some of the secondary benefits of tensor calculus:

A. The tensor notation, even detached from the powerful concept of a tensor, can often help systematize a calculation, particularly if differentiation is involved. The tensor notation is incredibly compact, especially with the help of the Einstein summation convention. Yet, despite its compactness, the notation is *utterly* robust and surprisingly explicit. It hides nothing, suggests correct moves, and translates to step-by-step recipes for calculation.

B. The concept of a tensor arises when one aims to preserve the geometric perspective and still take advantage of coordinate systems. A *tensor* is an encoded geometric object in a particular coordinate system. It is to be decoded at the right time: when the algebraic analysis is completed and we are ready for the answer. From this approach comes the true power of tensor calculus: it combines, with extraordinary success, the best of both geometric and algebraic worlds.

C. Tensor calculus is algorithmic. That is, tensor calculus expressions, such as $N_i \nabla^\alpha Z_\alpha^i$ for mean curvature of a surface, can be systematically translated into the lower-level combinations of familiar calculus operations. As a result, it is straightforward to implement tensor calculus symbolically. Various implementations are available in the most popular computer algebra systems and as stand-alone packages.

As you can see, the answer to the question *Why Tensor Calculus?* is quite multifaceted. The following four motivating examples continue answering this question in more specific terms. If at least one of these examples resonates with the reader and compels him or her to continue reading this textbook, then these examples will have accomplished their goal.

Motivating Example 1: The Gradient

What is the gradient of a function F and a point P? You are familiar with two definitions, one geometric and one analytical. According to the geometric definition, the gradient ∇F of F is the vector that points in the direction of the greatest increase of the function F, and its magnitude equals the greatest rate of increase. According to the analytical definition that requires the presence of a coordinate system, the gradient of F is the triplet of numbers

$$\nabla F = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y} \right). \tag{1.1}$$

Are the two definitions equivalent in some sense? If you believe that the connection is

$$\nabla F = \frac{\partial F}{\partial x}\mathbf{i} + \frac{\partial F}{\partial y}\mathbf{j}, \tag{1.2}$$

where $\mathbf{i}$ and $\mathbf{j}$ are the coordinate basis, you are in for a surprise! Equation (1.2) can only be considered valid if it produces the same vector in all coordinate systems. You may not be familiar with the definition of a coordinate basis in general curvilinear coordinates, such as spherical coordinates. The appropriate definition will be given in Chap. 5. However, equation (1.2) yields different answers even for the two coordinate systems in Fig. 1.1.

For a more specific example, consider a temperature distribution T in a two-dimensional rectangular room. Refer the interior of the room to a rectangular coordinate system x, y where the coordinate lines are *one* meter apart. This coordinate system is illustrated on the left of Fig. 1.1. Express the temperature field in terms of these coordinates and construct the vector gradient ∇T according to equation (1.2).

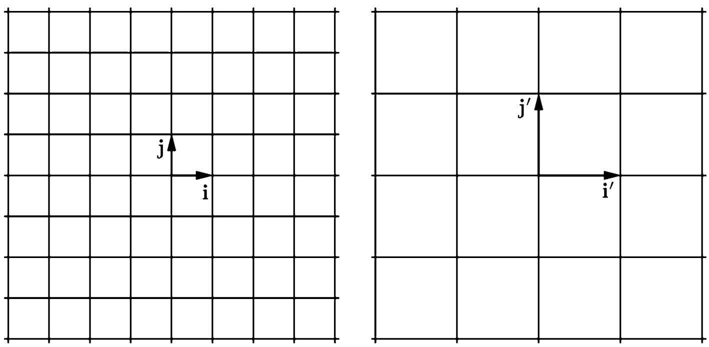


Fig. 1.1 When the expression in equation (1.2) is evaluated in these two coordinate systems, the results are not the same.

Alternatively, refer the interior of the room to another rectangular system x', y', illustrated on the right of Fig. 1.1), whose coordinate lines are *two* meters apart. For example, at a point where $x = 2$, the new coordinate x' equals 1. Therefore, the new coordinates and the old coordinates are related by the identities

$$x = 2x' \qquad y = 2y'. \tag{1.3}$$

Now repeat the construction of the gradient according to equation (1.2) in the new coordinate system: refer the temperature field to the new coordinates, resulting in the function $F'(x', y')$, calculate the partial derivatives and evaluate the expression in equation (1.2), except with "primed" elements:

$$(\nabla T)' = \frac{\partial F'}{\partial x'}\mathbf{i}' + \frac{\partial F'}{\partial y'}\mathbf{j}'. \tag{1.4}$$

How does ∇T compare to $(\nabla T)'$? The magnitudes of the new coordinate vectors $\mathbf{i}'$ and $\mathbf{j}'$ are double those of the old coordinate vectors $\mathbf{i}$ and $\mathbf{j}$. What happens to the partial derivatives? Do they halve (this would be good) or do they double (this would be trouble)?

They double! This is because in the new coordinates, quantities change twice as fast. In evaluating the rate of change with respect to, say, x, one increments x by a small amount Δx, such as $\Delta x = 10^{-3}$, and determines how much the function $F(x, y)$ changes in response to that small change in x. When one evaluates the partial derivative with respect to x' in the new coordinate system, the same increment in the new variable x' is, in physical terms, twice as large. It results in twice as large a change $\Delta F'$ in the function $F'(x', y')$. Therefore, $\Delta F'/\Delta x'$ is approximately twice as large as $\Delta F/\Delta x$ and we conclude that partial derivatives double:

$$\frac{\partial F'(x', y')}{\partial x'} = 2\frac{\partial F(x, y)}{\partial x}. \tag{1.5}$$

Thus, the relationship between $(\nabla T)'$ and ∇T reads $(\nabla T)' = 4\,\nabla T$ and the two results are different. Therefore, equation (1.2) cannot be used as the analytical definition of the gradient because it yields different results in different coordinate systems.

Tensor calculus offers a solution to this problem. Indeed, one of the central goals of tensor calculus is to construct expressions that evaluate to the same result in all coordinate systems. The fix to the gradient can be found in Sect. 6.8 of Chap. 6.

Exercise 1. Suppose that the temperature field T is given by the function $F(x, y) = x^2 e^y$ in coordinates x, y. Determine the function $F(x', y')$, which gives the temperature field T in coordinates x', y'.

Exercise 2. Confirm equation (1.5) for the functions $F(x, y)$ and $F'(x', y')$ derived in the preceding exercise.

Exercise 3. Show that the expression

$$\nabla T = \frac{1}{\sqrt{\mathbf{i}\cdot\mathbf{i}}}\frac{\partial F}{\partial x} + \frac{1}{\sqrt{\mathbf{j}\cdot\mathbf{j}}}\frac{\partial F}{\partial y} \tag{1.6}$$

yields the same result for all rescalings of Cartesian coordinates.

Exercise 4. Show that equation (1.2) yields the same expression in all Cartesian coordinates. The key to this exercise is the fact that any two Cartesian coordinate systems x, y and x', y' are related by the equation

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{1.7}$$

What is the physical interpretation of the numbers a, b, and α?

Exercise 5. Conclude that equation (1.6) is valid for all orthogonal affine coordinate systems. Affine coordinates are those with straight coordinate lines.

Motivating Example 2: The Determinant

In your linear algebra course, you studied determinants and you certainly learned the almost miraculous property that the determinant of a product is a product of determinants:

$$|AB| = |A|\,|B|\,. \tag{1.8}$$

This relationship can be explained by the deep geometric interpretation of the determinant as the signed volume of a parallelepiped. Some textbooks—in particular the textbook by Peter Lax [27]—take this approach. Other notable textbooks (e.g., Gilbert Strang's), derive (1.8) by row operations. Yet other classics, including Paul Halmos's [22] and Israel M. Gelfand's [13], prove identity (1.8) by analyzing the complicated algebraic definition of the determinant.

In tensor notation, the argument found in Halmos and Gelfand can be presented as an elegant and beautiful calculation. This calculation can be found in Chap. 9 and fits in a single line

$$C = \frac{1}{3!}\delta^{ijk}_{rst}c^r_i c^s_j c^t_k = \frac{1}{3!}e^{ijk}e_{rst}a^r_l b^l_i a^s_m b^m_j a^t_n b^n_k = \frac{1}{3!}ABe_{lmn}e^{lmn} = AB, \text{ Q.E.D.} \tag{1.9}$$

You will soon learn to carry out such chains with complete command. You will find that, in tensor notation, many previously complicated calculations become quite natural and simple.

Motivating Example 3: Vector Calculus

Hardly anyone can remember the key identities from vector calculus, let alone derive them. Do you remember this one?

$$\mathbf{A} \times (\nabla \times \mathbf{B}) = \nabla_{\mathbf{B}} (\mathbf{A} \cdot \mathbf{B}) - (\mathbf{A} \cdot \nabla) \mathbf{B}? \tag{1.10}$$

(And, more importantly, do you recall how to interpret the expressions $\nabla_{\mathbf{B}} (\mathbf{A} \cdot \mathbf{B})$ and $(\mathbf{A} \cdot \nabla) \mathbf{B}$?) I will admit that I do not have equation (1.10), either. However, when one knows tensor calculus, one need not memorize identities such as (1.10). Rather, one is able to derive and interpret them on the fly:

$$\varepsilon^{rsi} A_s \varepsilon_{ijk} \nabla^j B^k = \left(\delta^r_j \delta^s_k - \delta^s_j \delta^r_k\right) A_s \nabla^j B^k = A_s \nabla^r B^s - A_s \nabla^s B^r. \tag{1.11}$$

In equation (1.11) we see Cartan's orgy of formalism of equation (1.10) replaced by Weyl's orgy of indices. In this particular case, the benefits of the tensor approach are evident.

Motivating Example 4: Shape Optimization

The problem of finding a surface with the least surface area that encloses the domain of a given volume, the answer being a sphere, is one of the oldest problems in mathematics. It is now considered to be a classical problem of the calculus of variations. Yet, most textbooks on the calculus of variations deal only with the two-dimensional variant of this problem, often referred to as the *Dido Problem* or the *problem of Queen Dido*, which is finding a curve of least arc length that incloses a domain of a given area. The full three-dimensional problem is usually deemed to be too complex, while the treatment of the two-dimensional problem often takes two or more pages.

The calculus of a moving surface—an extension of tensor calculus to deforming manifolds, to which Part III of this textbook is devoted—solves this problem directly, naturally, and concisely. The following derivation is found in Chap. 17 where all the necessary details are given. Here, we give a general outline merely to showcase the conciseness and the elegance of the analysis.

The modified objective function E incorporating the isoperimetric constraint for the surface S enclosing the domain Ω reads

$$E = \int_S dS + \lambda \int_\Omega d\Omega, \tag{1.12}$$

where λ is a Lagrange multiplier. The variation δE with respect to variations C in shape is given by

$$\delta E = \int_S C\left(-B_\alpha^\alpha + \lambda\right) dS, \tag{1.13}$$

where B_α^α is *mean curvature*. Since C can be treated as an independent variation, the equilibrium equation states that the optimal shape has constant mean curvature

$$B_\alpha^\alpha = \lambda. \tag{1.14}$$

Naturally, a sphere satisfies this equation.

This derivation is one of the most vivid illustrations of the power of tensor calculus and the calculus of moving surfaces. It shows how much can be accomplished with the help of the grand compromise of tensor calculus, including preservation of geometric insight while enabling robust analysis.

Part I
Tensors in Euclidean Spaces

Chapter 2
Rules of the Game

2.1 Preview

According to the great German mathematician David Hilbert,"mathematics is a game played according to certain simple rules with meaningless marks on paper." The goal of this chapter is to lay down the simple rules by which the game of tensor calculus is played.

We take a relatively informal approach to the foundations of our subject. We do not mention $\mathbb{R}^n$, groups, isomorphisms, homeomorphisms and polylinear forms. Instead, we strive to build a common context by appealing to concepts that we find intuitively clear. Our goal is to establish an understanding of Euclidean spaces and of the fundamental operations that take place in a Euclidean space—most notably, operations on vectors including differentiation of vectors with respect to a parameter.

2.2 The Euclidean Space

All objects discussed in this book live in a *Euclidean space*.

What *is* a Euclidean space? There is a formal definition, but for us an informal definition will do: a Euclidean space corresponds to the physical space of our everyday experience. There are many features of this space that we take for granted. The foremost of these is its ability to accommodate straight-edged objects. Take a look at Rodin's "The Thinker" (*Le Penseur*) in Fig. 2.1. Of course, the sculpture itself does not have a single straight line in it. But the gallery in which it is housed is a perfect rectangular box.

Our physical space is called *Euclidean* after the ancient Greek mathematician Euclid of Alexandria who lived in the fourth century BC. He is the author of the *Elements* [11], a book in which he formulated and developed the subject of geometry. The *Elements* reigned as the supreme textbook on geometry for an

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_2,

Fig. 2.1 Rodin's "Thinker" finds himself in a Euclidean space characterized by the possibility of straightness

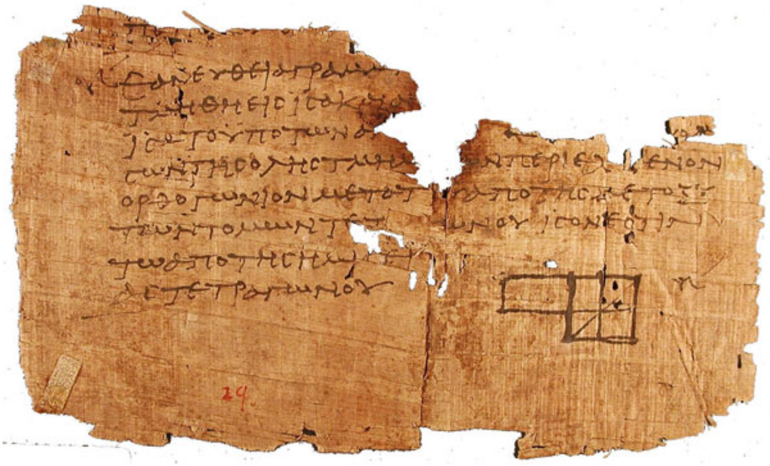


Fig. 2.2 A page from *The Elements*

astoundingly long period of time: since its creation and until the middle of the twentieth century. The *Elements* set the course for modern science and earned Euclid a reputation as the "Father of Geometry".

Euclid's geometry is based on the study of straight lines and planes. Figure 2.2 shows a page from one of the oldest (circa 100AD) surviving copies of *The Elements*. That page contains a drawing of a square and an adjacent rectangle. Such a drawing may well be found in a contemporary geometry textbook. The focus is on straight lines.

Not all spaces have the ability to accommodate straight lines. Only the straight line, the plane, and the full three-dimensional space are Euclidean. The rest, such as the surface of a sphere, are not. Some curved spaces display only certain features of Euclidean spaces. For example, the surface of a cylinder, while non-Euclidean, can be cut with a pair of scissors and unwrapped into a flat Euclidean space. The same can be said of the surface of a cone. We know, on an intuitive level, that these are *essentially flat* surfaces arranged, with no distortions, in the *ambient* three-dimensional space. We have not given the term "essentially flat" a precise meaning, but we will do so in Chap. 12.

2.3 Length, Area, and Volume

Are you willing to accept the concept of a Euclidean space without a formal definition? If so, you should similarly accept two additional geometric concepts: *length* of a segment and *angle* between two segments.

The concept of *area* can be built up from the concept of length. The area of a rectangle is the product ab of the lengths of its sides. A right triangle is half of a rectangle, so its area is the product ab of the lengths of its sides adjacent to the right angle. The area of an arbitrary triangle can be calculated by dividing it into two right triangles. The area of a polygon can be calculated by dividing it into triangles.

The concept of *volume* is similarly built up from the concept of length. The volume of a rectangular box is the product abc of the lengths of its sides. The volume of a *tetrahedron* is $Ah/3$ where h is the height and A is the area of the opposite base. Interestingly, the geometric proof of this formula is not simple. The volume of a polyhedron can be calculated by dividing it intro tetrahedra.

Our intuitive understanding of lengths, areas, and volumes also extends to curved geometries. We understand the meaning of the surface area of a sphere or a cylinder or any other curved surface. That is not to say that the *formalization* of the notion of area is completely straightforward. Even for the cylinder, an attempt to define area as the limit of surface areas of inscribed polyhedra was met with fundamental difficulties. An elementary example [39, 48], in which the inscribed areas approach infinity, was put forth by the German mathematician Hermann Schwarz.

However, difficulties in *formalizing* the concepts of surface area and volume did not prevent mathematicians from to calculating those quantities effectively. An early and fundamental breakthrough came from Archimedes who demonstrated that the volume of a sphere is $\frac{4}{3}\pi R^3$ and its surface area is $4\pi R^2$. The final word on the subject of quadrature (i.e., computing areas and volumes) came nearly 2,000 years later in the form of Isaac Newton's calculus. We reiterate that difficulties in formalizing the concepts of length, area, and volume and difficulties in obtaining analytical expressions for lengths, areas, and volumes should not prevent us from discussing and using these concepts.

In our description of space, *length* is a *primary concept*. That is, the concept of length comes first and other concepts are defined in terms of it. For example, below we define the dot product between two vectors as the product of their lengths and the cosine of the angle between them. Thus, the dot product a *secondary concept* defined in terms of length. Meanwhile, length is not defined in terms of anything else—it is accepted without a definition.

2.4 Scalars and Vectors

Scalar fields and *vector fields* are ubiquitous in our description of natural phenomena. The term *field* refers to a collection of *scalars* or *vectors* defined at each point of a Euclidean space or subdomain Ω. A *scalar* is a real number. Examples of scalar fields include temperature, mass density, and pressure. A *vector* is a directed segment. Vectors are denoted by boldface letters such as $\mathbf{V}$ and $\mathbf{R}$. Examples of vector fields include gravitational and electromagnetic forces, fluid flow velocity, and the vector gradient of a scalar field. This book introduces a new meaningful type of field—*tensor fields*—when the Euclidean space is referred to a coordinate system.

2.5 The Dot Product

The *dot product* is an operation of fantastic utility. For two vectors $\mathbf{U}$ and $\mathbf{V}$, the dot product $\mathbf{U}\cdot\mathbf{V}$ is defined as the product of their lengths $|\mathbf{U}|$ and $|\mathbf{V}|$ and the cosine of the angle between them:

$$\mathbf{U}\cdot\mathbf{V} = |\mathbf{U}|\,|\mathbf{V}|\cos\alpha. \tag{2.1}$$

This is the only sense in which the dot product is used in this book. The great utility of the dot product comes from the fact that most, if not all, geometric properties can be expressed in terms of the dot product.

The dot product has a number of fundamental properties. It is commutative

$$\mathbf{U}\cdot\mathbf{V} = \mathbf{V}\cdot\mathbf{U} \tag{2.2}$$

which is clear from the definition (2.1). It is also linear in each argument. With respect to the first argument, linearity means

$$(\mathbf{U}_1 + \mathbf{U}_2)\cdot\mathbf{V} = \mathbf{U}_1\cdot\mathbf{V} + \mathbf{U}_2\cdot\mathbf{V}. \tag{2.3}$$

This property is also known as *distributivity*. Unlike *commutativity*, which immediately follows from the definition (2.1), distributivity is not at all obvious and the reader is invited to demonstrate it geometrically.

Exercise 6. Demonstrate geometrically that the dot product is distributive.

As we mentioned above, most geometric quantities can be expressed in terms of the dot product. The length $|\mathbf{U}|$ of a vector $\mathbf{U}$ is given by

$$|\mathbf{U}| = \sqrt{\mathbf{U} \cdot \mathbf{U}}. \tag{2.4}$$

Similarly, the angle α between two vectors $\mathbf{U}$ and $\mathbf{V}$, is given by

$$\alpha = \arccos \frac{\mathbf{U} \cdot \mathbf{V}}{\sqrt{\mathbf{U} \cdot \mathbf{U}}\sqrt{\mathbf{V} \cdot \mathbf{V}}}. \tag{2.5}$$

The definition (2.1) of the dot product relies on the concept of the angle between two vectors. Therefore, angle may be considered a *primary* concept. It turns out, however, that the concept of angle can be derived from the concept of length. The dot product $\mathbf{U} \cdot \mathbf{V}$ can be defined in terms of lengths alone, without a reference to the angle α:

$$\mathbf{U} \cdot \mathbf{V} = \frac{|\mathbf{U} + \mathbf{V}|^2 - |\mathbf{U} - \mathbf{V}|^2}{4}. \tag{2.6}$$

Thus, the concept of angle is not needed for the definition of the dot product. Instead, equation (2.5) can be viewed as the definition of the angle between $\mathbf{U}$ and $\mathbf{V}$.

2.5.1 Inner Products and Lengths in Linear Algebra

In our approach, the concept of length comes first and the dot product is built upon it. In linear algebra, the dot product is often referred to as the *inner product* and the relationship is reversed: the inner product $(\mathbf{U}, \mathbf{V})$ is any operation that satisfies three properties (distributivity, symmetry, and positive definiteness), and length $|\mathbf{U}|$ is defined as the square root of the inner product of $\mathbf{U}$ with itself

$$|\mathbf{U}| = \sqrt{(\mathbf{U}, \mathbf{U})}. \tag{2.7}$$

Thus, in linear algebra, *the inner product is a primary concept and length is a secondary concept.*

2.6 The Directional Derivative

A directional derivative measures the rate of change of a scalar field F along a straight line. Let a straight ray l emanate from the point P. Suppose that P^* is a nearby point on l and let P^* approach P in the sense that the distance PP^*

approaches zero. Then the directional derivative dF/dl at the point P is defined as the limit

$$\frac{dF(P)}{dl} = \lim_{P^* \to P} \frac{F(P^*) - F(P)}{PP^*}. \tag{2.8}$$

Instead of a ray, we could consider an entire straight line, but we must still pick a direction along the line. Additionally, we must let the distance PP^* be signed. That is, we agree that PP^* is positive if the direction from P to P^* is the same as the chosen direction of the line l, and negative otherwise.

The beauty of the directional derivative is that it is an entirely geometric concept. Its definition requires two elements: straight lines and length. Both are present in Euclidean spaces, making definition (2.8) possible. Importantly, the definition of a directional derivative does not require a coordinate system.

The definition of the direction derivative can be given in terms of the ordinary derivative of a scalar function. This can help avoid using limits by hiding the concept of the limit at a lower level. Parameterize the points P^* along the straight line l by the signed distance s from P to P^*. Then the values of function of s in the sense of ordinary calculus. Let us denote that function by $f(s)$, using a lower case f to distinguish it from F, the scalar field defined in the Euclidean space. Then, as it is easy to see

$$\frac{dF(P)}{dl} = f'(0), \tag{2.9}$$

where the derivative of $f(s)$ is evaluated in the sense of ordinary calculus. Definition (2.9) assumes that the parameterization is such that $s = 0$ at P. More generally, if the point P is located at $s = s_0$, then the definition of dF/dl is

$$\frac{dF(P)}{dl} = f'(s_0). \tag{2.10}$$

The following exercises are meant reinforce the point that directional derivatives can be evaluated without referring the Euclidean space to a coordinate system.

Exercise 7. Evaluate dF/dl for $F(P) =$ "Distance from point P to point A" in a direction perpendicular to AP.

Exercise 8. Evaluate dF/dl for $F(P) =$ "1/(Distance from point P to point A)" in the direction from P to A.

Exercise 9. Evaluate dF/dl for $F(P) =$ "Angle between OA and OP", where O and A are two given points, in the direction from P to A.

Exercise 10. Evaluate dF/dl for $F(P) =$ "Distance from P to the straight line that passes through A and B", where A and B are given points in the direction parallel to AB. The distance between a point P and a straight line is defined as the shortest distance between P and any of the points on the straight line. The same definition applies to the distance between a point and a curve.

Exercise 11. Evaluate dF/dl for $F(P) =$ "Area of triangle PAB", where A and B are fixed points, in the direction parallel to AB.

Exercise 12. Evaluate dF/dl for $F(P) =$ "Area of triangle PAB", where A and B are fixed points, in the direction orthogonal to AB.

2.7 The Gradient

The concept of the directional derivative leads to the concept of the *gradient*. The gradient ∇F of F is defined as a *vector* that points in the direction of the greatest increase in F. That is, it points in the direction l along which dF/dl has the greatest value. The length of the gradient vector equals the rate of the increase, that is $|\nabla F| = dF/dl$. Note that the symbol ∇ in ∇F is **bold** indicating that the gradient is a vector quantity. Importantly, the gradient is a geometric concept. For a given scalar field, it can be evaluated, at least conceptually, by pure geometric means without a reference to a coordinate system.

Exercise 13. Describe the gradient for each of the functions from the preceding exercises.

The gradient ∇F and the directional derivative dF/dl along the ray l are related by the dot product:

$$\frac{dF}{dl} = \nabla F \cdot \mathbf{L}, \tag{2.11}$$

where $\mathbf{L}$ is the unit vector in the direction of the ray l. This is a powerful relationship indeed and perhaps somewhat unexpected. It shows that knowing the gradient is sufficient to determine the directional derivatives in all directions. In particular, the directional derivative is zero in any direction orthogonal to the gradient.

If we were to impose a coordinate system on the Euclidean space (doing so would be very much against the geometric spirit of this chapter), we would see equation (2.11) as nothing more than a special form of the chain rule. This approach to equation (2.11) is discussed in Chap. 6 after we have built a proper framework for operating safely in coordinate systems. Meanwhile, you should view equation (2.11) as a geometric identity

Exercise 14. Give the geometric intuition behind equation (2.11). In other words, explain geometrically, why knowing the gradient is sufficient to determine the directional derivatives in all possible directions.

2.8 Differentiation of Vector Fields

In order to be subject to differentiation, a field must consist of elements that posses three basic properties. First, the elements can be added together to produce another element of the same kind. Second, the elements can be multiplied by real numbers. Can geometric vectors be added together and multiplied by real numbers? Certainly, yes. So far, geometric vectors are good candidates for differentiation with respect to a parameter.

The third property is the ability to approach a limit. The natural definition of limits for vector quantities is based on distance. If vector $\mathbf{A}$ depends on a parameter h, we say that $\mathbf{A}(h) \to \mathbf{B}$ as $h \to 0$, if

$$\lim_{h\to 0} |\mathbf{A}(h) - \mathbf{B}| = 0. \tag{2.12}$$

In summary, geometric vectors posses the three required properties. Thus, geometric vectors can be differentiated with respect to a parameter.

Consider the radial segment on the unit circle and treat it as a vector function $\mathbf{R}(\alpha)$ of the angle α. In order to determine $\mathbf{R}'(\alpha)$, we appeal to the definition (2.8). We take ever smaller values of h and construct the ratio $(\mathbf{R}(\alpha + h) - \mathbf{R}(\alpha))/h$ in Fig. 2.3. The figure shows the vectors $\mathbf{R}(\alpha)$ and $\mathbf{R}(\alpha + h)$, their diminishing difference $\Delta\mathbf{R} = \mathbf{R}(\alpha + h) - \mathbf{R}(\alpha)$, as well as the quantity $\Delta\mathbf{R}/h$. It is apparent that the ratio $\Delta\mathbf{R}/h$ converges to a specific vector. That vector is $\mathbf{R}'(\alpha)$.

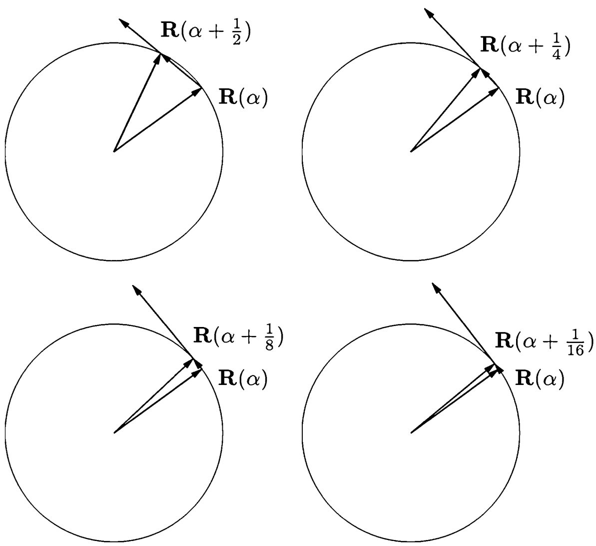


Fig. 2.3 A limiting process that constructs the derivative of the vector $\mathbf{R}(\alpha)$ with respect to α

In practice, it is rare that essentially geometric analysis can produce an answer. For example, from the example above, we may conjecture that $\mathbf{R}'(\alpha)$ is orthogonal to $\mathbf{R}'(\alpha)$ and of unit length, but it is nontrivial to demonstrate so convincingly in purely geometric terms. Certain elements of the answer can be demonstrated without resorting to algebraic analysis in a coordinate system, but typically not all.

In our current example $\mathbf{R}'(\alpha)$ can be determined completely without any coordinates. To show orthogonality, note that $\mathbf{R}(\alpha)$ is the unit length, which can be written as

$$\mathbf{R}(\alpha) \cdot \mathbf{R}(\alpha) = 1. \tag{2.13}$$

Differentiating both sides of this identity with respect to α, we have, by the product rule

$$\mathbf{R}'(\alpha) \cdot \mathbf{R}(\alpha) + \mathbf{R}(\alpha) \cdot \mathbf{R}'(\alpha) = 0, \tag{2.14}$$

where we have assumed, rather reasonably but without a formal justification, that the derivative of a scalar product satisfies the product rule. We therefore have

$$\mathbf{R}(\alpha) \cdot \mathbf{R}'(\alpha) = 0, \tag{2.15}$$

showing orthogonality.

To show that $\mathbf{R}'(\alpha)$ is unit length, note that the length of the vector $\mathbf{R}(\alpha + h) - \mathbf{R}(\alpha)$ is $2\sin\frac{h}{2}$. Thus the length of $\mathbf{R}'(\alpha)$ is given by the limit

$$\lim_{h\to 0} \frac{2\sin\frac{h}{2}}{h} = 1 \tag{2.16}$$

and we have confirmed that $\mathbf{R}'(\alpha)$ is orthogonal to $\mathbf{R}(\alpha)$ and is unit length. In the next chapter, we rederive this result with the help of a coordinate system as an illustration of the utility of coordinates.

Exercise 15. Show that the length of $\mathbf{R}(\alpha + h) - \mathbf{R}(\alpha)$ is $2\sin\frac{h}{2}$.

Exercise 16. Confirm that the limit in (2.16) is 1 by L'Hôpital's rule.

Exercise 17. Alternatively, recognize that the limit in (2.16) equals $\sin' 0 = \cos 0 = 1$.

Exercise 18. Determine $\mathbf{R}''(\alpha)$.

2.9 Summary

In this chapter, we introduced the fundamental concepts upon which the subject of tensor calculus is built. The primary elements defined in the Euclidean space are scalar and vector fields. Curve segments, two-dimensional surface patches, and three-dimensional domains are characterized lengths, areas, and volumes. This chapter dealt with geometric objects that can be discussed without a reference to a coordinate system. However, coordinate systems are necessary since the overwhelming majority of applied problems cannot be solved by geometric means alone. When coordinate systems are introduced, tensor calculus serves to preserve the geometric perspective by offering a strategy that leads to results that are independent of the choice of coordinates.

Chapter 3
Coordinate Systems and the Role of Tensor Calculus

3.1 Preview

Tensor calculus was invented in order to make geometric and analytical methods work together effectively. While geometry is one of the oldest and most developed branches of mathematics, coordinate systems are relatively new, dating back to the 1600s. The introduction of coordinate systems enabled the use of algebraic methods in geometry and eventually led to the development of calculus. However, along with their tremendous power, coordinate systems present a number of potential pitfalls, which soon became apparent. Tensor calculus arose as a mechanism for overcoming these pitfalls. In this chapter, we discuss coordinate systems, the advantages and disadvantages of their use, and explain the need for tensor calculus.

3.2 Why Coordinate Systems?

Coordinate systems make tasks easier! The invention of coordinates systems was a watershed event in seventeenth-century science. Joseph Lagrange (1736–1813) described the importance of this event in the following words: "As long as algebra and geometry have been separated, their progress has been slow and their uses limited, but when these two sciences have been united, they have lent each mutual forces, and have marched together towards perfection." Coordinate systems enabled the use of powerful algebraic methods in geometric problems. This in turn led to the development of new algebraic methods inspired by geometric insight. This unprecedented wave of change culminated in the invention of calculus. Prior to the invention of coordinate systems, mathematicians had developed extraordinary skill at solving problems by geometric means. The invention of coordinates and the subsequent invention of calculus opened up problem solving to the masses.

The individual whose name is most closely associated with the invention of coordinate systems is René Descartes (1596–1650), whose portrait is found in

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*, DOI 10.1007/978-1-4614-7867-6_3, © Springer Science+Business Media New York 2013

Fig. 3.1 Réne Descartes is credited with the invention of coordinate systems. His use of coordinates was brilliant

Fig. 3.1. We do not know if Descartes was the first to think of assigning numbers to points, but he was the first to use this idea to incredible effect. For example, the coordinate system figured prominently in Descartes' tangent line method, which he described as "Not only the most useful and most general problem in geometry that I know, but even that I have ever desired to know." The original description of the method may be found in Descartes' masterpiece *The Geometry* [9]. An excellent historical perspective can be found in [15].

An example from Chap. 2 provides a vivid illustration of the power of coordinates. In Sect. 2.8, we considered a vector $\mathbf{R}(\alpha)$ that traces out the unit as the parameter α changes from 0 to 2π. We concluded, by reasoning geometrically, that the derivative $\mathbf{R}'(\alpha)$ is unit length and orthogonal to $\mathbf{R}(\alpha)$. Our derivation took a certain degree of geometric insight. Even for a *slightly* harder problem—for example, $\mathbf{R}(\alpha)$ tracing out an ellipse instead of a circle—$\mathbf{R}'(\alpha)$ would be much harder to calculate.

With the help of coordinates, the same computation becomes elementary and even much harder problems can be handled with equal ease. Let us use Cartesian coordinates and denote the coordinate basis by $\{\mathbf{i},\mathbf{j}\}$. Then, $\mathbf{R}(\alpha)$ is given by

$$\mathbf{R}(\alpha) = \mathbf{i}\cos\alpha + \mathbf{j}\sin\alpha, \tag{3.1}$$

which yields $\mathbf{R}'(\alpha)$ by a simple differentiation:

$$\mathbf{R}'(\alpha) = -\mathbf{i}\sin\alpha + \mathbf{j}\cos\alpha. \tag{3.2}$$

That's it: we found $\mathbf{R}'(\alpha)$ in one simple step. To confirm that $\mathbf{R}'(\alpha)$ is orthogonal to $\mathbf{R}(\alpha)$, dot $\mathbf{R}'(\alpha)$ with $\mathbf{R}(\alpha)$:

$$\mathbf{R}'(\alpha)\cdot\mathbf{R}(\alpha) = -\sin\alpha\cos\alpha + \sin\alpha\cos\alpha = 0. \tag{3.3}$$

To confirm that $\mathbf{R}'(\alpha)$ is unit length, compute the dot product

$$\mathbf{R}'(\alpha)\cdot\mathbf{R}'(\alpha) = \sin^2\alpha + \cos^2\alpha = 1. \tag{3.4}$$

Even this simple example shows the tremendous utility of operating within a coordinate system. On the other hand, coordinates must be handled with skill. The use of coordinates comes with its own set of pitfalls. Overreliance of coordinates is often counterproductive. Tensor calculus was born out of the need for a systematic and disciplined strategy for using coordinates. In this chapter, we will touch upon the difficulties that arise when coordinates are used inappropriately.

3.3 What Is a Coordinate System?

A coordinate system assigns sets of numbers to points in space in a systematic fashion. The choice of the coordinate system is dictated by the problem. If a problem can be solved with the help of one coordinate system it may also be solved with another. However, the solution may be more complicated in one coordinate system than the other. For example, we can refer the surface of the Earth to spherical coordinates (described in Sect. 3.6.5). When choosing a spherical coordinate system on a sphere, one needs to make two decisions—where to place the poles and where the azimuth count starts. The usual choice is a good one: the coordinate poles coincide with the North and South poles of the Earth, and the main meridian passes through London.

This coordinate system is convenient in many respects. For example, the length of day can be easily determined from the latitude and the time of year. Climate is very strongly tied to the latitude as well. Time zones roughly follow meridians. Centripetal acceleration is strictly a function of the latitude. It is greatest at the equator ($\theta = \pi/2$).

Imagine what would happen if the coordinate poles were placed in Philadelphia ($\theta = 0$) and the opposite pole ($\theta = \pi$) in the middle of the ocean southwest of Australia? Some tasks would become easier, others harder. For example, calculating the distance to Philadelphia would become a very simple task: the point with coordinates (θ, ϕ) is $R\theta$ miles away, where R is the radius of the Earth. On the other hand, some of the more important tasks, like determining the time of day, would become substantially more complicated.

3.4 Perils of Coordinates

It may seem then, that the right strategy when solving a problem is to pick the right coordinate system. Not so! The conveniences of coordinate systems come with great costs including loss of generality and loss of geometric insight. This can quite often be the difference between succeeding and failing at solving the problem.

Loss of Geometric Insight

A famous historical episode can illustrate how an analytical calculation, albeit brilliant, can fail to identify a simple geometric picture. Consider the task of finding the least surface area that spans a three-dimensional contour U. Mathematically, the problem is to minimize the area integral

$$A = \int_S dS \tag{3.5}$$

over all possible surfaces S for which the contour boundary U is specified. Such a surface is said to be *minimal*. The problem of finding minimal surfaces is a classical problem in the *calculus of variations*. Leonhard Euler laid the foundation for this subject in 1744 in the celebrated work entitled *The method of finding plane curves that show some property of maximum and minimum [12]*. Joseph Lagrange advanced Euler's geometric ideas by formulating an analytical method in *Essai d'une nouvelle méthode pour déterminer les maxima et les minima des formules intégrales indéfinies [26]* published in 1762. Lagrange derived a relationship that a surface in Cartesian coordinates (x, y, z) represented by

$$z = F(x, y) \tag{3.6}$$

must satisfy. That relationship reads

$$F_{xx} + F_{yy} + F_{xx}F_y^2 + F_{yy}F_x^2 - 2F_xF_yF_{xy} = 0. \tag{3.7}$$

The symbols F_{xx} denotes $\partial^2 F/\partial x^2$ and the rest of the elements follow the same convention. In the context of calculus of variations, equation (3.7) is known as the *Euler-Lagrange equation* for the functional (3.5).

What is the geometric meaning of equation (3.7)? We have the answer now but, as hard as it is to believe, neither Lagrange nor Euler knew it! The geometric insight came 14 years later from a young French mathematician Jean Baptiste Meusnier. Meusnier realized that minimal surfaces are characterized by zero mean curvature. We denote mean curvature by B_α^α and discuss it thoroughly in Chap. 12. For a surface given by (3.6) in Cartesian coordinates, mean curvature B_α^α is given by

$$B_\alpha^\alpha = \frac{F_{xx} + F_{yy} + F_{xx}F_y^2 + F_{yy}F_x^2 - 2F_xF_yF_{xy}}{\left(1 + F_x^2 + F_y^2\right)^{3/2}}. \tag{3.8}$$

Therefore, as Meusnier discovered, minimal surfaces are characterized by zero mean curvature:

$$B_\alpha^\alpha = 0. \tag{3.9}$$

The history of calculus of variations is one of shining success and has bestowed upon its authors much deserved admiration. Furthermore, we can learn a very valuable lesson from the difficulties experienced even by the great mathematicians. By choosing to operate in a particular coordinate system, Lagrange purposefully sacrificed a great deal of geometric insight in exchange for the power of analytical methods. In doing so, Lagrange was able to solve a wider range of problems than the subject's founder Euler.

Analytic Complexity

As we can see, coordinate systems are indispensable for problem solving. However, the introduction of a *particular* coordinate system must take place at the right moment in the solution of a problem. Otherwise, the expressions that one encounters can become impenetrably complex. The complexity of the expression (3.8) for mean curvature speaks to this point. Lagrange was able to overcome the computational difficulties that arose in obtaining equation (3.7). Keep in mind, however, that equation (3.7) corresponds to one of the simplest problems of its kind. Imagine the complexity that would arise is the analysis of a more complicated problem. For example, consider the following problem that has important applications for biological membranes: to find the surface S with a given contour boundary U that minimizes the integral of the square of mean curvature:

$$E = \int_S \left(B_\alpha^\alpha\right)^2 dS. \tag{3.10}$$

In Cartesian coordinates, E is given by the expression

$$E = \int_S \frac{\left(F_{xx} + F_{yy} + F_{xx}F_y^2 + F_{yy}F_x^2 - 2F_xF_yF_{xy}\right)^2}{\left(1 + F_x^2 + F_y^2\right)^3} dxdy. \tag{3.11}$$

As it is shown in Chap. 17, the Euler–Lagrange equation corresponding to equation (3.10) reads

$$\nabla_\beta \nabla^\beta B_\alpha^\alpha - B_\alpha^\alpha B_\beta^\gamma B_\gamma^\beta - \frac{1}{2}\left(B_\alpha^\alpha\right)^3 = 0 \tag{3.12}$$

$$
\begin{aligned}
&\frac{1}{(1+F_x^2+F_y^2)^{9/2}} \\
&(-3 F_{xx}^3 + 15 F_x^2 F_{xx}^3 - 10 F_x F_{xx} F_{xxx} - 10 F_x^3 F_{xx} F_{xxx} + F_{xxxx} + 2 F_x^2 F_{xxxx} + F_x^4 F_{xxxx} + 2 F_{xxyy} + 6 F_x^2 F_{xxyy} + 6 F_x^4 F_{xxyy} + \\
&\quad 2 F_x^6 F_{xxyy} - 12 F_x F_{xxy} F_{xy} - 24 F_x^3 F_{xxy} F_{xy} - 12 F_x^5 F_{xxy} F_{xy} - 8 F_{xx} F_{xy}^2 + 16 F_x^2 F_{xx} F_{xy}^2 + 24 F_x^4 F_{xx} F_{xy}^2 - \\
&\quad 6 F_x F_{xx} F_{xyy} - 12 F_x^3 F_{xx} F_{xyy} - 6 F_x^5 F_{xx} F_{xyy} - 4 F_x F_{xxxy} F_y - 8 F_x^3 F_{xxxy} F_y - 4 F_x^5 F_{xxxy} F_y - 6 F_{xx} F_{xxy} F_y + \\
&\quad 18 F_x^2 F_{xx} F_{xxy} F_y + 24 F_x^4 F_{xx} F_{xxy} F_y + 48 F_x F_{xx}^2 F_{xy} F_y - 60 F_x^3 F_{xx}^2 F_{xy} F_y - 4 F_{xxx} F_{xy} F_y + 12 F_x^2 F_{xxx} F_{xy} F_y + \\
&\quad 16 F_x^4 F_{xxx} F_{xy} F_y + 40 F_x F_{xy}^3 F_y + 16 F_x^3 F_{xy}^3 F_y - 24 F_x^5 F_{xy}^3 F_y - 12 F_{xy} F_{xyy} F_y + 36 F_x^4 F_{xy} F_{xyy} F_y + 24 F_x^6 F_{xy} F_{xyy} F_y - \\
&\quad 4 F_x F_{xyyy} F_y - 12 F_x^3 F_{xyyy} F_y - 12 F_x^5 F_{xyyy} F_y - 4 F_x^7 F_{xyyy} F_y - 9 F_{xx}^3 F_y^2 + 30 F_x^2 F_{xx}^3 F_y^2 - 30 F_x F_{xx} F_{xxx} F_y^2 - \\
&\quad 20 F_x^3 F_{xx} F_{xxx} F_y^2 + 4 F_{xxxx} F_y^2 + 6 F_x^2 F_{xxxx} F_y^2 + 2 F_x^4 F_{xxxx} F_y^2 + 6 F_{xxyy} F_y^2 + 18 F_x^2 F_{xxyy} F_y^2 + 18 F_x^4 F_{xxyy} F_y^2 + \\
&\quad 6 F_x^6 F_{xxyy} F_y^2 - 36 F_x^3 F_{xxy} F_{xy} F_y^2 - 36 F_x^5 F_{xxy} F_{xy} F_y^2 - 4 F_{xx} F_{xy}^2 F_y^2 - 116 F_x^2 F_{xx} F_{xy}^2 F_y^2 + 72 F_x^4 F_{xx} F_{xy}^2 F_y^2 - \\
&\quad 18 F_x^3 F_{xx} F_{xyy} F_y^2 - 18 F_x^5 F_{xx} F_{xyy} F_y^2 - 12 F_x F_{xxxy} F_y^3 - 16 F_x^3 F_{xxxy} F_y^3 - 4 F_x^5 F_{xxxy} F_y^3 - 18 F_{xx} F_{xxy} F_y^3 + \\
&\quad 36 F_x^2 F_{xx} F_{xxy} F_y^3 + 24 F_x^4 F_{xx} F_{xxy} F_y^3 + 96 F_x F_{xx}^2 F_{xy} F_y^3 - 60 F_x^3 F_{xx}^2 F_{xy} F_y^3 - 12 F_{xxx} F_{xy} F_y^3 + 24 F_x^2 F_{xxx} F_{xy} F_y^3 + \\
&\quad 16 F_x^4 F_{xxx} F_{xy} F_y^3 + 16 F_x F_{xy}^3 F_y^3 + 96 F_x^3 F_{xy}^3 F_y^3 - 24 F_{xy} F_{xyy} F_y^3 - 36 F_x^2 F_{xy} F_{xyy} F_y^3 - 12 F_x^4 F_{xy} F_{xyy} F_y^3 - \\
&\quad 8 F_x F_{xyyy} F_y^3 - 16 F_x^3 F_{xyyy} F_y^3 - 8 F_x^5 F_{xyyy} F_y^3 - 9 F_{xx}^3 F_y^4 + 15 F_x^2 F_{xx}^3 F_y^4 - 30 F_x F_{xx} F_{xxx} F_y^4 - 10 F_x^3 F_{xx} F_{xxx} F_y^4 + \\
&\quad 6 F_{xxxx} F_y^4 + 6 F_x^2 F_{xxxx} F_y^4 + F_x^4 F_{xxxx} F_y^4 + 6 F_{xxyy} F_y^4 + 18 F_x^2 F_{xxyy} F_y^4 + 12 F_x^4 F_{xxyy} F_y^4 + 36 F_x F_{xxy} F_{xy} F_y^4 - \\
&\quad 12 F_x^3 F_{xxy} F_{xy} F_y^4 + 16 F_{xx} F_{xy}^2 F_y^4 - 132 F_x^2 F_{xx} F_{xy}^2 F_y^4 + 18 F_x F_{xx} F_{xyy} F_y^4 - 6 F_x^3 F_{xx} F_{xyy} F_y^4 - 12 F_x F_{xxxy} F_y^5 - \\
&\quad 8 F_x^3 F_{xxxy} F_y^5 - 18 F_{xx} F_{xxy} F_y^5 + 18 F_x^2 F_{xx} F_{xxy} F_y^5 + 48 F_x F_{xx}^2 F_{xy} F_y^5 - 12 F_{xxx} F_{xy} F_y^5 + 12 F_x^2 F_{xxx} F_{xy} F_y^5 - \\
&\quad 24 F_x F_{xy}^3 F_y^5 - 12 F_{xy} F_{xyy} F_y^5 - 36 F_x^2 F_{xy} F_{xyy} F_y^5 - 4 F_x F_{xyyy} F_y^5 - 4 F_x^3 F_{xyyy} F_y^5 - 3 F_{xx}^3 F_y^6 - 10 F_x F_{xx} F_{xxx} F_y^6 + \\
&\quad 4 F_{xxxx} F_y^6 + 2 F_x^2 F_{xxxx} F_y^6 + 2 F_{xxyy} F_y^6 + 6 F_x^2 F_{xxyy} F_y^6 + 24 F_x F_{xxy} F_{xy} F_y^6 + 12 F_{xx} F_{xy}^2 F_y^6 + 12 F_x F_{xx} F_{xyy} F_y^6 - \\
&\quad 4 F_x F_{xxxy} F_y^7 - 6 F_{xx} F_{xxy} F_y^7 - 4 F_{xxx} F_{xy} F_y^7 + F_{xxxx} F_y^8 + ((1+F_x^2)^2 (-1+3F_x^2) F_{xx} + 48 F_x (1+F_x^2)^2 F_{xy} F_y + \\
&\qquad (5 - 28 F_x^2 - 33 F_x^4) F_{xx} F_y^2 - 60 F_x (1+F_x^2) F_{xy} F_y^3 + 6 (1+3F_x^2) F_{xx} F_y^4) F_{yy}^2 + 3 (1+F_x^2)^2 (-1 - F_x^2 + 5 F_y^2) F_{yy}^3 - \\
&\quad 4 F_x F_{xy} F_{yyy} - 12 F_x^3 F_{xy} F_{yyy} - 12 F_x^5 F_{xy} F_{yyy} - 4 F_x^7 F_{xy} F_{yyy} - 2 F_{xx} F_y F_{yyy} + 6 F_x^4 F_{xx} F_y F_{yyy} + \\
&\quad 4 F_x^6 F_{xx} F_y F_{yyy} + 12 F_x F_{xy} F_y^2 F_{yyy} + 24 F_x^3 F_{xy} F_y^2 F_{yyy} + 12 F_x^5 F_{xy} F_y^2 F_{yyy} - 4 F_{xx} F_y^3 F_{yyy} - \\
&\quad 6 F_x^2 F_{xx} F_y^3 F_{yyy} - 2 F_x^4 F_{xx} F_y^3 F_{yyy} + 16 F_x F_{xy} F_y^4 F_{yyy} + 16 F_x^3 F_{xy} F_y^4 F_{yyy} - 2 F_{xx} F_y^5 F_{yyy} - 6 F_x^2 F_{xx} F_y^5 F_{yyy} + \\
&\quad F_{yy} (-(1+F_x^2) (F_{xx}^2 - 6 F_x^2 F_{xx}^2 - 4 (-2 + F_x^2 + 3 F_x^4) F_{xy}^2 + 6 F_x^5 F_{xyy} + 2 F_x (F_{xxx} + 3 F_{xyy}) + 2 F_x^3 (F_{xxx} + 6 F_{xyy})) + \\
&\qquad (F_{xx}^2 - 28 F_x^2 F_{xx}^2 + 18 F_x^4 F_{xx}^2 - 4 (-4 + 29 F_x^2 + 33 F_x^4) F_{xy}^2 + 18 F_x F_{xyy} + 6 F_x^5 (-F_{xxx} + 3 F_{xyy}) + 6 F_x^3 (-F_{xxx} + 6 F_{xyy})) \\
&\qquad F_y^2 + (5 F_{xx}^2 - 33 F_x^2 F_{xx}^2 + 24 (1 + 3F_x^2) F_{xy}^2 + 6 F_x (F_{xxx} + 4 F_{xyy}) + F_x^3 (-2 F_{xxx} + 24 F_{xyy})) F_y^4 - \\
&\qquad 6 (F_{xxy} + 3 F_x^2 F_{xxy} + 6 F_x F_{xx} F_{xy}) F_y^5 + (3 F_{xx}^2 + 4 F_x F_{xxx}) F_y^6 - \\
&\qquad 2 F_y^3 (6 F_{xxy} + 9 F_x^2 F_{xxy} + 3 F_x^4 F_{xxy} - 10 F_x F_{xx} F_{xy} - 72 F_x^3 F_{xx} F_{xy} + 5 (1+F_x^2)^2 F_{yyy}) - 2 (1+F_x^2) F_y \\
&\qquad (3 F_{xxy} - 3 F_x^2 F_{xxy} - 6 F_x^4 F_{xxy} - 28 F_x F_{xx} F_{xy} + 18 F_x^3 F_{xx} F_{xy} + 5 (1+F_x^2)^2 F_{yyy})) + (1+F_x^2)^2 (1+F_x^2+F_y^2)^2 F_{yyyy})
\end{aligned}
$$

Fig. 3.2 The Laplacian of mean curvature in Cartesian coordinates

The first term in this equation is $\nabla_\beta \nabla^\beta B_\alpha^\alpha$, the *surface Laplacian* $\nabla_\beta \nabla^\beta$ of mean curvature B_α^α. The surface Laplacian is discussed in Chap. 11. Let us now try to imagine the mayhem that the Euler–Lagrange equation would be if the standard analysis of calculus of variations were applied directly to equation (3.11). The expression for $\nabla_\beta \nabla^\beta B_\alpha^\alpha$, a relatively simple geometric quantity, is shown in Fig. 3.2. Despite its apparent impenetrability, it is endowed with geometric meaning: it is the *Laplacian of mean curvature*. The Laplacian of mean curvature is an important object that arises often in the analysis of surfaces. We shall denote it by $\nabla_\alpha \nabla^\alpha B_\beta^\beta$. It is hard to image that someone would carry their analysis this far! If one did, it is even harder to imagine that someone would be able to recognize this expression as the Laplacian of mean curvature! These unmanageable expressions arise because the formulation of the minimization problem (3.11) in a Cartesian coordinate system—or any other particular coordinate system, for that matter—is ineffective.

3.5 The Role of Tensor Calculus

The central idea of tensor calculus is to acknowledge the need for coordinate systems and, at the same time, to avoid selecting a particular coordinate system for as long as possible in the course of solving a problem. Tensor calculus provides algorithms for constructing analytical expressions that are valid in *all* coordinate systems simultaneously. You will soon learn that tensor expressions, such as B^{α}_{α}, translate to consistent and detailed recipes for calculating geometric quantities in particular coordinate systems. If an expression evaluates to the same value in all coordinate systems, then *it must be endowed with geometric meaning*. Tensor calculus has proven to be a remarkably effective framework in which Lagrange's *mutual forces* can be lent between algebra and geometry.

3.6 A Catalog of Coordinate Systems

In this section, we describe the most commonly used coordinates systems in Euclidean spaces: Cartesian, affine, polar, cylindrical and spherical. A good way to illustrate a coordinate system graphically is by drawing the *coordinate lines* in two dimensions (e.g. Fig. 3.3), and *coordinate surfaces* in three dimensions (e.g. Fig. 3.4). A coordinate line or surface is a set of points for which the value of a particular coordinate is fixed. For example, in Cartesian coordinates x, y, z in three

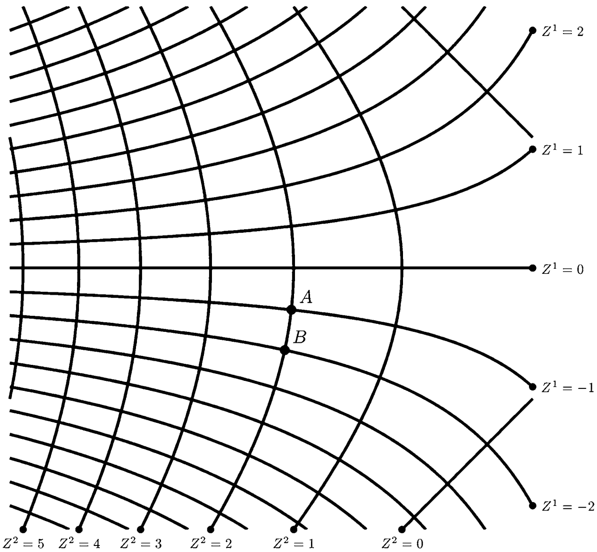


Fig. 3.3 Coordinate lines for a typical curvilinear coordinate system

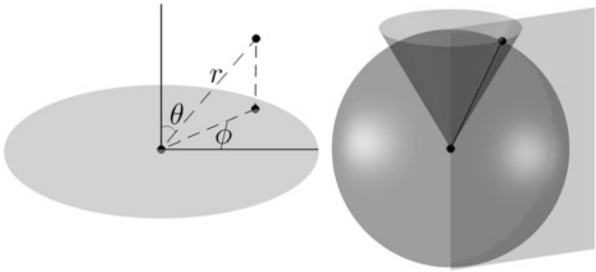


Fig. 3.4 Spherical coordinates. The plot on the right shows an example of a coordinate surface for each variable

dimensions, the coordinate surface corresponding to $x = 1$ is a plane parallel to the yz-plane. In spherical coordinates, the coordinate surface $r = 2$ is a sphere of radius 2 centered at the origin.

The figures below show coordinate lines corresponding to integer values of the fixed coordinate. In other words, the display coordinate lines are spaced one *coordinate unit* apart. When a Euclidean space is referred to a coordinate system, the term *unit* becomes ambiguous: it may refer to a coordinate unit or a unit of Euclidean length. For example, in Fig. 3.3 depicting a generic coordinate system Z^1, Z^2, the points A and B have coordinates $(-1, 2)$ and $(-2, 2)$ and are therefore one coordinate unit apart. However, the Euclidean distance between A and B may be quite different.

In all likelihood, you are already familiar with the most common coordinate systems described below. Nevertheless, we invite you to read the following sections, since we describe each coordinates system in absolute terms, rather than (as done in most texts) by a reference to a background Cartesian grid. This absolute approach is more true to the spirit of tensor calculus.

3.6.1 Cartesian Coordinates

We start with the legendary *Cartesian coordinates*. Cartesian coordinates are the most used and the most overused coordinate system. It is a natural choice in many situations and, in an a number of ways, the easiest coordinate system to use. It is also an unnatural choice in many situations where it is used anyway, especially those situations where one need not choose a coordinate system at all. It is one of the goals of this book to rid the reader of the acquired dependency on Cartesian coordinates.

Cartesian coordinates x, y refer the Euclidean plane to a regular square grid with coordinate lines one Euclidean unit apart, as illustrated in Fig. 3.5. A point with coordinates $x = x_0$ and $y = y_0$ is denoted by (x_0, y_0). The point O with coordinates $(0, 0)$ is called the *origin*.

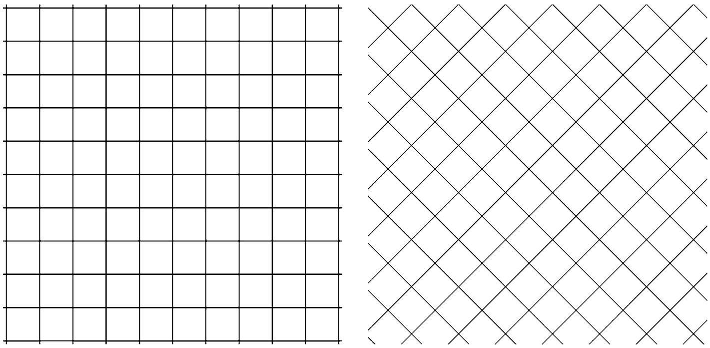


Fig. 3.5 Two different Cartesian coordinate systems

Cartesian coordinates (as well as *affine* coordinates described next) can be characterized by a natural *coordinate basis* $\mathbf{i}, \mathbf{j}$. The vector $\mathbf{i}$ connects the origin to the point with coordinates $(1, 0)$ while the vector $\mathbf{j}$ connects the origin to the point with coordinates $(0, 1)$. A vector $\mathbf{V}$ connecting the origin to the point with coordinates (x_0, y_0) can be expressed by the linear combination

$$\mathbf{V} = x_0\,\mathbf{i} + y_0\,\mathbf{j}. \tag{3.13}$$

Thus, a Cartesian coordinate yields an orthonormal basis $\mathbf{i}, \mathbf{j}$ for vectors in the plane. Conversely, any orthonormal basis $\mathbf{i}, \mathbf{j}$, combined with the location of the origin, corresponds to a Cartesian coordinate system.

The requirement that the coordinate unit equals the Euclidean unit of length is essential to the definition of a Cartesian coordinate system. A similarly constructed coordinate systems with integer coordinate lines two Euclidean units of length apart would no longer be considered Cartesian. It would instead be characterized as an *orthogonal affine* coordinate system.

There are infinitely many Cartesian coordinate systems in the plane. A Cartesian coordinate system is characterized by three continuous degrees of freedom: the location of the origin and direction of the x-axis. Furthermore, the role of axes can be reversed. That is, $\mathbf{i}$ and $\mathbf{j}$ can be chosen such that the shortest path from $\mathbf{i}$ to $\mathbf{j}$ is either counterclockwise or clockwise. In the former case, the coordinates are called *right-handed*, in the latter—*left-handed*.

Exercise 19. Generalize the contents of the section to three dimensions.

Exercise 20. Describe a set of six degrees of freedom in choosing Cartesian coordinates in the three-dimensional space.

The concept of orientation extends to arbitrary dimensions. In three dimensions, a Cartesian coordinate system is called right-handed if the coordinate triplet $\mathbf{i}, \mathbf{j}, \mathbf{k}$ satisfies the right-hand rule: when one curls the fingers of the right hand from the vector $\mathbf{i}$ to the vector $\mathbf{j}$, the thumb points in the direction of vector $\mathbf{k}$. If the thumb points in the direction opposite of the vector $\mathbf{k}$, the coordinate system is called left-handed.

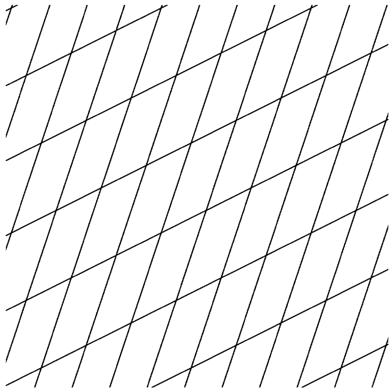


Fig. 3.6 An affine coordinate system

3.6.2 Affine Coordinates

Affine coordinates are characterized by a skewed but otherwise regular grid of coordinate lines, as in Fig. 3.6. The easiest way to introduce affine coordinates is by selecting an origin O and an arbitrary triplet of vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ which may point in any direction and have any length. To determine the affine coordinates of a point A, consider a vector $\mathbf{V} = \overrightarrow{OA}$ and decompose $\mathbf{V}$ with respect to the basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$

$$\mathbf{V} = x_0\,\mathbf{i} + y_0\,\mathbf{j} + z_0\mathbf{k}. \tag{3.14}$$

Then the point A is assigned coordinates x_0, y_0, z_0.

The triplet $\mathbf{i}, \mathbf{j}, \mathbf{k}$ is the coordinate basis. The vector $\mathbf{i}$ points from the origin to the point with coordinates $(1, 0, 0)$, just as in the case of Cartesian coordinates. Cartesian coordinates are a special case of affine coordinates with an orthonormal coordinate basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$. Any two affine coordinate systems are related by an invertible linear transformations with a shift. In particular, any affine coordinate system is related to a Cartesian coordinate system by a linear transformation with a shift. The last statement may also be taken as the definition of Cartesian coordinates.

Affine coordinates can also be referred to as *rectilinear coordinates*, that is, coordinate systems where the coordinate lines are straight. Coordinate systems where the coordinate lines are not straight are called *curvilinear*. We now turn to the description of the most fundamental curvilinear coordinates.

3.6.3 Polar Coordinates

The construction of a polar coordinate system in the plane is illustrated in Fig. 3.7. An arbitrary point O is designated as the *pole* or *origin* and a ray emanating from O is selected. This ray is known as the *polar axis*. Each point A in the plane is

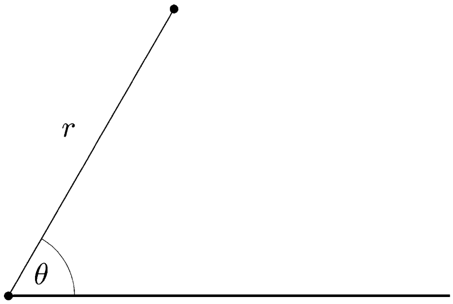


Fig. 3.7 A polar coordinates system

assigned two coordinates, r and θ, where r is the Euclidean distance from the pole and θ is the signed angle that the segment OA forms with the polar axis. According to the typical sign convention, the angle θ is positive if the angle is formed in the *counterclockwise* direction. Each point in the plane is characterized by a unique value of r, but infinitely many values of θ. That is, coordinates (r, θ) and $(r, \theta + 2\pi n)$ for any integer n refer to the same point in the plane. Finally, the angle θ is not defined at the origin O.

Notice that we described polar coordinates in absolute terms, rather than by referencing some previously established Cartesian coordinate system. This is important from two points of view. First, it is very much in the spirit of tensor calculus, which proclaims that all coordinate systems are created equal and it is only the specifics of the problem at hand that dictate the choice of a particular coordinate system. Second, it is good to see an explicit reference to Euclidean distance and angle and to realize that polar coordinates are intrinsically linked to the Euclidean nature of the plane. In particular, r is measured in Euclidean units. A rescaled polar coordinate system would no longer be called *polar coordinates*, but rather *scaled polar coordinates*.

3.6.4 Cylindrical Coordinates

Cylindrical coordinates extend polar coordinates to three dimensions. Cylindrical coordinates are constructed as follows. First a plane is selected and a polar coordinate system is constructed in that plane. This special plane is referred to as the *coordinate plane*. Then each point A in the Euclidean space is described by the triplet of numbers (r, θ, z), where r and θ are the polar coordinates of the orthogonal projection of A onto the coordinate plane and z is the signed Euclidean distance between A and the coordinate axis. The sign convention usually follows the right-hand rule. The origin O of the polar coordinates in the coordinate plane is also called the origin with regard to the cylindrical coordinates.

The straight line that passes through O and is orthogonal to the coordinate plane consists of points for which $r = 0$. It is known as the *cylindrical* or *longitudinal*

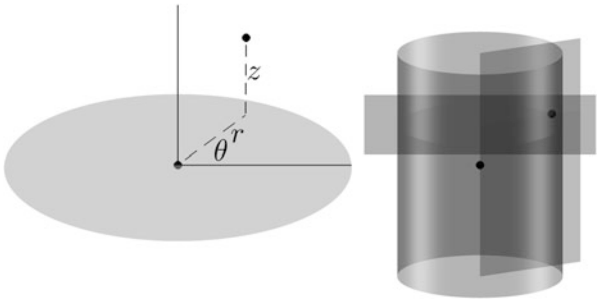


Fig. 3.8 Cylindrical coordinates. The plot on the right gives an example of a coordinate surface for each variable

axis. An alternative interpretation of r is the Euclidean distance from this axis. The term *cylindrical* comes from the fact that points described by constant r form a cylinder (Fig. 3.8).

3.6.5 Spherical Coordinates

Spherical coordinates r, θ, ϕ are perfectly intuitive because the angles θ and ϕ correspond, respectively, to longitude and latitude on the surface of the Earth, and r is the distance to the center of the Earth. A formal construction of spherical coordinates starts by selecting a coordinate plane with an origin O and a polar axis that establishes the azimuthal angle ϕ. The coordinate r is defined as the Euclidean distance to the origin O. The set of points corresponding to a given value of r is a sphere of that radius with the center at the origin. Such spheres are called *coordinate spheres*. Fixed r and θ correspond to points that form a meridian on a coordinate sphere of radius r. The final coordinate θ, known as the *longitudinal angle*, varies from 0 to π and gives the angle between OA and the longitudinal axis, a straight line orthogonal to the coordinate plane that passes through the origin O. The plane of the equator corresponds to $\theta = \pi/2$. Neither angle is determined at the origin O. The azimuthal angle ϕ is undefined along the longitudinal axis.

3.6.6 Relationships Among Common Coordinate Systems

Exercise 21. Given a polar coordinate system, introduce Cartesian coordinates that originate at the same pole and for which the x-axis coincides with the polar axis. Show that the Cartesian coordinates (x, y) are related to the polar coordinates (r, θ) by the following identities:

$$x(r,\theta) = r\cos\theta \tag{3.15}$$

$$y(r,\theta) = r\sin\theta. \tag{3.16}$$

Exercise 22. Show that the inverse relationship is

$$r(x,y) = \sqrt{x^2+y^2} \tag{3.17}$$

$$\theta(x,y) = \arctan\frac{y}{x}. \tag{3.18}$$

Exercise 23. Align a Cartesian coordinate system x, y, z with a given cylindrical coordinate system in a natural way. Show that x, y, z and r, θ, z (we hope it is OK to reuse the letter z) are related as follows

$$x(r,\theta,z) = r\cos\theta \tag{3.19}$$

$$y(r,\theta,z) = r\sin\theta \tag{3.20}$$

$$z(r,\theta,z) = z. \tag{3.21}$$

Exercise 24. Show that the inverse relationships are

$$r(x,y,z) = \sqrt{x^2+y^2} \tag{3.22}$$

$$\theta(x,y,z) = \arctan\frac{y}{x} \tag{3.23}$$

$$z(x,y,z) = z. \tag{3.24}$$

Exercise 25. Align a Cartesian coordinate system with a spherical coordinate system in a natural way. Show that x, y, z and r, θ, ϕ are related by

$$x(r,\theta,\phi) = r\sin\theta\cos\phi \tag{3.25}$$

$$y(r,\theta,\phi) = r\sin\theta\sin\phi \tag{3.26}$$

$$z(r,\theta,\phi) = r\cos\theta. \tag{3.27}$$

Exercise 26. Show that the inverse relationships is

$$r(x,y,z) = \sqrt{x^2+y^2+z^2} \tag{3.28}$$

$$\theta(x,y,z) = \arcsin\frac{z}{\sqrt{x^2+y^2+z^2}} \tag{3.29}$$

$$\phi(x,y,z) = \arctan\frac{y}{x}. \tag{3.30}$$

3.7 Summary

In this chapter we highlighted some of the most important advantages of coordinate systems. Of course, the coordinate approach hardly needs a champion. However, there are many perils of using coordinate systems that include loss of geometric insight and unmanageable complexity of calculations. We will find that tensor calculus goes a long way towards overcoming these challenges and thus helps make the use of coordinate systems even more powerful.

There are a handful of coordinate systems that are used in solving problems for special geometries. Several of these—Cartesian, affine, polar, cylindrical, and spherical—were described in this chapter.

Chapter 4
Change of Coordinates

4.1 Preview

Tensor calculus achieves invariance by establishing rules for forming expressions that evaluate to the same value in all coordinate systems. In order to construct expressions that do not depend on the choice of coordinates, one must understand how individual elements transform under a change of coordinate. This chapter is therefore devoted to the study of coordinate changes and the corresponding changes in calculated quantities. Looking ahead, we discover that certain objects transform according to a special rule that makes it particularly easy to form invariant combinations. Such objects are called *tensors*.

In this chapter, we also introduce the essential elements of the tensor notation. We begin to use indices in earnest and show their great utility in forming very compact and readable expressions.

4.2 An Example of a Coordinate Change

Consider the relationship between Cartesian coordinates x, y and polar coordinates r, θ in the plane:

$$r(x,y) = \sqrt{x^2 + y^2} \tag{4.1}$$

$$\theta(x,y) = \arctan\frac{y}{x}. \tag{4.2}$$

The inverse relationship is

$$x(r,\theta) = r\cos\theta \tag{4.3}$$

$$y(r,\theta) = r\sin\theta. \tag{4.4}$$

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_4,

We note that in these equations, some letters denote more than one object. In equations (4.1) and (4.2), the letters x and y denote independent variables while r and θ denote functions of x and y. In equations (4.3) and (4.4), the letters x and y denote functions or r and θ, while r and θ are independent variables. One may find it helpful to use new letters, f and g, to denote these functions:

$$r = f(x, y) = \sqrt{x^2 + y^2} \tag{4.5}$$

$$\theta = g(x, y) = \arctan \frac{y}{x} \tag{4.6}$$

and two more letters, h and i, for the inverse relationships:

$$x = h(r, \theta) = r \cos \theta \tag{4.7}$$

$$y = i(r, \theta) = r \sin \theta \tag{4.8}$$

We do not use extra letters in our discussion. Instead, we use another notational device for distinguishing between objects such as r-the-independent-variable and r-the-function-of-x-and-y: when we mean r-the-function, we include the functional arguments and write $r(x, y)$. That makes it clear that we are talking about r-the-function rather than r-the-independent-variable.

4.3 A Jacobian Example

Let us construct the matrix $J(x, y)$, often denoted in literature by $\partial(r, \theta)/\partial(x, y)$, which consists of the partial derivatives of the functions $r(x, y)$ and $\theta(x, y)$ with respect to each of the Cartesian variables:

$$J(x, y) = \begin{bmatrix} \frac{\partial r(x,y)}{\partial x} & \frac{\partial r(x,y)}{\partial y} \\ \frac{\partial \theta(x,y)}{\partial x} & \frac{\partial \theta(x,y)}{\partial y} \end{bmatrix}. \tag{4.9}$$

This matrix is called the *Jacobian of the coordinate transformation*. Evaluating the partial derivatives, we find

$$J(x, y) = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{bmatrix}. \tag{4.10}$$

Next, calculate the Jacobian J' for the reverse transformation of coordinates:

$$J'(r, \theta) = \begin{bmatrix} \frac{\partial x(r,\theta)}{\partial r} & \frac{\partial x(r,\theta)}{\partial \theta} \\ \frac{\partial y(r,\theta)}{\partial r} & \frac{\partial y(r,\theta)}{\partial \theta} \end{bmatrix}. \tag{4.11}$$

Evaluating the partial derivatives, we find

$$J' = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}. \tag{4.12}$$

Exercise 27. Calculate the determinant $|J|$.

Exercise 28. Evaluate J at the point with Cartesian coordinates $x = 1$, $y = 1$?

Exercise 29. Calculate the determinant $|J'|$. How does it relate to $|J|$?

Exercise 30. Evaluate J' at the point with polar coordinates $r = \sqrt{2}$, $\theta = \pi/4$ (which is the same as the point with Cartesian coordinates $x = 1$, $y = 1$).

Exercise 31. Evaluate the matrix product

$$JJ' \tag{4.13}$$

at the point with Cartesian coordinates $x = 1$, $y = 1$ or, equivalently, polar coordinates $r = \sqrt{2}$, $\theta = \pi/4$. Did your calculation yield the identity matrix?

4.4 The Inverse Relationship Between the Jacobians

In this section, we find that the Jacobians J and J' evaluated at the same point in space are matrix inverses of each other:

$$JJ' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.14}$$

In order to carry out the multiplication, we must express both Jacobians in the same coordinates. We have treated $J(x, y)$ as a function of x and y and $J'(r, \theta)$ as a function of r and θ. But you may also think of J and J' as functions of points in the Euclidean space. As such, they may be expressed in either coordinates. Express the Jacobian $J(x, y)$ in polar coordinates:

$$J(r,\theta) = \begin{bmatrix} \frac{x(r,\theta)}{\sqrt{x(r,\theta)^2+y(r,\theta)^2}} & \frac{y(r,\theta)}{\sqrt{x(r,\theta)^2+y(r,\theta)^2}} \\ \frac{-y(r,\theta)}{x(r,\theta)^2+y(r,\theta)^2} & \frac{x(r,\theta)}{x(r,\theta)^2+y(r,\theta)^2} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\frac{\sin\theta}{r} & \frac{\cos\theta}{r} \end{bmatrix}. \tag{4.15}$$

and evaluate the matrix product JJ':

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\frac{\sin\theta}{r} & \frac{\cos\theta}{r} \end{bmatrix} \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.16}$$

As predicted, the result is the identity matrix. The remainder of this chapter is devoted to demonstrating that this special relationship is holds for arbitrary coordinate changes. This task gives us an opportunity to introduce the key elements of the tensor notation.

Exercise 32. Show that J and J' are the inverses of each other by transforming $J'(r,\theta)$ to Cartesian coordinates.

Exercise 33. In three dimensions, analyze the transformation from Cartesian to spherical coordinates and confirm that the associated Jacobians are matrix inverses of each other.

4.5 The Chain Rule in Tensor Notation

Suppose that F is a function of three variables a, b, and c:

$$F = F(a,b,c). \tag{4.17}$$

Suppose further, that we have three functions of two variables μ and ν:

$$A(\mu,\nu) \tag{4.18}$$
$$B(\mu,\nu) \tag{4.19}$$
$$C(\mu,\nu). \tag{4.20}$$

Form a new function $f(\mu,\nu)$ by composition

$$f(\mu,\nu) = F(A(\mu,\nu), B(\mu,\nu), C(\mu,\nu)). \tag{4.21}$$

Use the chain rule to evaluate the derivative of $f(\mu,\nu)$ with respect to each of the arguments:

$$\frac{\partial f(\mu,\nu)}{\partial \mu} = \frac{\partial F}{\partial a}\frac{\partial A}{\partial \mu} + \frac{\partial F}{\partial b}\frac{\partial B}{\partial \mu} + \frac{\partial F}{\partial c}\frac{\partial C}{\partial \mu} \tag{4.22}$$
$$\frac{\partial f(\mu,\nu)}{\partial \nu} = \frac{\partial F}{\partial a}\frac{\partial A}{\partial \nu} + \frac{\partial F}{\partial b}\frac{\partial B}{\partial \nu} + \frac{\partial F}{\partial c}\frac{\partial C}{\partial \nu}. \tag{4.23}$$

Exercise 34. Evaluate

$$\frac{\partial^2 f(\mu,\nu)}{\partial \mu^2}, \frac{\partial^2 f(\mu,\nu)}{\partial \mu \partial \nu}, \text{ and } \frac{\partial f^2(\mu,\nu)}{\partial \nu^2}. \tag{4.24}$$

Produce expression similar to equations (4.22) and (4.23).

Exercise 35. Derive the third-order derivative $\partial^3 f(\mu,\nu)/\partial^2\mu\partial\nu$.

We now take several decisive steps towards converting equations (4.22) and (4.23) to tensor form. The first step is to rename the variables a, b, and c into a^1, a^2, and a^3. Collectively, we refer to these variables as a^i. Note that we use an *upper index*, or *superscript*, and that it denotes enumeration and not exponentiation.

The fact that F is a function of a^1, a^2, and a^3 can be expressed in one of three ways:

$$F = F\left(a^1, a^2, a^3\right) \tag{4.25}$$

$$F = F\left(a^i\right) \tag{4.26}$$

$$F = F(a). \tag{4.27}$$

We prefer equation (4.27). That is, we represent the entire collection of arguments by a single letter and suppress the superscript i. It is important *not* to think of the symbol a as a vector with components a^1, a^2, and a^3. (A vector is a directed segment, after all, not a triplet of numbers). Rather, $F(a)$ is a shorthand notation for $F\left(a^1, a^2, a^3\right)$.

The next step in converting equations (4.22) and (4.23) to tensor form is to use the symbols A^1, A^2, and A^3 for the three functions A, B, and C:

$$A^1(\mu,\nu) \tag{4.28}$$

$$A^2(\mu,\nu) \tag{4.29}$$

$$A^3(\mu,\nu). \tag{4.30}$$

Collectively, these functions can be denoted by $A^i(\mu,\nu)$. Once again, we use a superscript for enumeration.

The function $f(\mu,\nu)$, constructed by composing $F(a)$—that is $F\left(a^1, a^2, a^3\right)$, with $A^i(\mu,\nu)$—that is $A^1(\mu,\nu)$, $A^2(\mu,\nu)$ and $A^3(\mu,\nu)$, can now be expressed by the following compact equation:

$$f(\mu,\nu) = F(A(\mu,\nu)). \tag{4.31}$$

When expanded, this expression reads

$$f(\mu,\nu) = F\left(A^1(\mu,\nu), A^2(\mu,\nu), A^3(\mu,\nu)\right). \tag{4.32}$$

With the use of superscripts, the equations (4.22) and (4.23) representing the chain rule, can be rewritten more compactly way with the help of the summation symbol $\sum$:

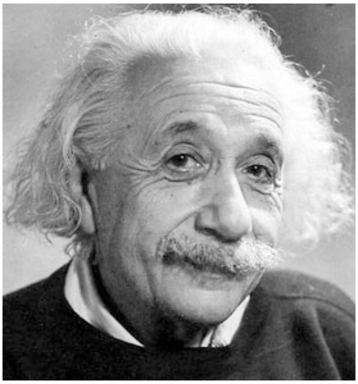


Fig. 4.1 Albert Einstein (1879–1955) was a celebrated twentieth-century scientist and one of the earliest proponents of tensor calculus. Einstein suggested to suppress the summation sign in tensor expressions

$$\frac{\partial f(\mu,\nu)}{\partial \mu} = \sum_{i=1}^{3} \frac{\partial F}{\partial a^i}\frac{\partial A^i}{\partial \mu} \tag{4.33}$$

$$\frac{\partial f(\mu,\nu)}{\partial \nu} = \sum_{i=1}^{3} \frac{\partial F}{\partial a^i}\frac{\partial A^i}{\partial \nu} \tag{4.34}$$

We are well on our way towards tensor notation. In the expression $\partial F/\partial a^i$, the index i is a *lower index*, or *subscript*, since it is a superscript in the denominator. If the expression $\partial F/\partial a^i$ were denoted by a special symbol, such as $\nabla_i F$ or $F_{|i}$, the index would necessarily appear as a subscript.

The next important element is the elimination of the summation sign. According to the *Einstein notation* (or the *Einstein summation convention*), summation is *implied* when an index appears twice, once as a subscript and once as a superscript. Albert Einstein whose portrait appears in Fig. 4.1 suggested the convention in 1916 [10] and later joked in a letter to a friend [25]: *I have made a great discovery in mathematics; I have suppressed the summation sign each time that the summation must be made over an index that occurs twice…*

With the Einstein convention, equations (4.33) and (4.34) become

$$\frac{\partial f(\mu,\nu)}{\partial \mu} = \frac{\partial F}{\partial a^i}\frac{\partial A^i}{\partial \mu} \tag{4.35}$$

$$\frac{\partial f(\mu,\nu)}{\partial \nu} = \frac{\partial F}{\partial a^i}\frac{\partial A^i}{\partial \nu}. \tag{4.36}$$

The repeated index i, which implies summation, is called a *dummy* index, or a *repeated* index, or a *contracted* index. Summation over a dummy index is called a *contraction*.

Exercise 36. Derive the expressions for

$$\frac{\partial^2 f(\mu,\nu)}{\partial\mu^2}, \frac{\partial^2 f(\mu,\nu)}{\partial\mu\partial\nu}, \text{ and } \frac{\partial f(\mu,\nu)}{\partial\nu^2} \tag{4.37}$$

similar to equations (4.35) and (4.36).

The final step in converting equations (4.22) and (4.23) to tensor form is to combine them into a single expression. Denote the independent variables μ and ν by μ^1 and μ^2, or collectively μ^α. We utilize a letter from the Greek alphabet because this index assumes values $1, 2$, compared to the values $1, 2, 3$ assumed by Latin indices.

The fact that f is a function of μ^1 and μ^2 can also be expressed in one of three ways,

$$f\left(\mu^1,\mu^2\right) \tag{4.38}$$

$$f\left(\mu^\alpha\right) \tag{4.39}$$

$$f(\mu), \tag{4.40}$$

and we once again choose the third way, where the index of the functional argument is dropped. Similarly, the fact that each A^i is a function of μ^1 and μ^2 is denoted by $A^i(\mu)$.

The composition by which $f(\mu)$ is formed can now be expressed quite concisely

$$f(\mu) = F(A(\mu)). \tag{4.41}$$

Note that we suppressed both indices of functional arguments. The compact equation (4.41) in actuality represents

$$f\left(\mu^1,\mu^2\right) = F\left(A^1\left(\mu^1,\mu^2\right), A^2\left(\mu^1,\mu^2\right), A^3\left(\mu^1,\mu^2\right)\right). \tag{4.42}$$

We are now able to capture the two partial derivatives of f by a single elegant expression:

$$\frac{\partial f}{\partial\mu^\alpha} = \frac{\partial F}{\partial a^i}\frac{\partial A^i}{\partial\mu^\alpha}. \tag{4.43}$$

Let us review the main features of equation (4.43). First, this single expression represents two relationships, one for $\alpha = 1$ and one for $\alpha = 2$. This makes α a *live* index. Second, on the right-hand side, there is an implied summation over the dummy index i because it is repeated as a subscript and a superscript.

Exercise 37. Show that the tensor expression for the collection of second-order partial derivatives is

$$\frac{\partial^2 f}{\partial \mu^\alpha \partial \mu^\beta} = \frac{\partial^2 F}{\partial A^i \partial A^j} \frac{\partial A^i}{\partial \mu^\alpha} \frac{\partial A^j}{\partial \mu^\beta} + \frac{\partial F}{\partial A^i} \frac{\partial^2 A^i}{\partial \mu^\alpha \partial \mu^\beta}. \tag{4.44}$$

How many identities does the resulting tensor expression represent?

Exercise 38. Derive the tensor expression for the collection of third-order partial derivatives

$$\frac{\partial^3 f}{\partial \mu^\alpha \partial \mu^\beta \partial \mu^\gamma}. \tag{4.45}$$

In this section, we introduced the essential features of the tensor notation. Over the next few chapters we will discover just how robust the tensor notation is. It is capable of expressing a broad range of operations in mathematics, including the matrix product and the operation of the transpose.

4.6 Inverse Functions

The inverse matrix relationship between the Jacobians J and J' is a consequence of the fact that the transformations from Cartesian to polar coordinates and back are the functional inverses of each other. Let us first explore the inverse relationship between functions of one variable.

What does it mean for two functions $f(x)$ and $g(x)$ to be the inverses of each other? It means that applying f to $g(x)$ yields x:

$$g(f(x)) \equiv x. \tag{4.46}$$

From this identity, we can obtain a relationship between the derivatives f' and g'. Differentiating with respect to x yields

$$g'(f(x))\, f'(x) = 1, \tag{4.47}$$

or

$$f'(x) = \frac{1}{g'(f(x))}. \tag{4.48}$$

Thus, with some care, we can state that g' and f' are algebraic inverses of each other. We must remember that, in order for this statement to be true, g' must be evaluated at $f(x)$ rather than x. For example, the derivative of $f(x) = e^{2x}$ at $x = 3$ is $2e^6$. Meanwhile, the derivative of $g(x)\,\frac{1}{2} = \ln x$ at $f(3) = e^6$ is $\frac{1}{2e^6}$.

Exercise 39. Derive that

$$\frac{d\arccos x}{dx} = \pm\frac{1}{\sqrt{1-x^2}} \tag{4.49}$$

by differentiating the identity $\cos\arccos x = x$.

Exercise 40. Suppose that $f(x)$ is the inverse of $g(x)$. Show that the second derivatives f'' and g'' satisfy the identity

$$g''(f(x))\,f'(x)^2 + g'(f(x))\,f''(x) = 0. \tag{4.50}$$

Exercise 41. Verify equation (4.50) for $f(x) = e^x$ and $g(x) = \ln x$.

Exercise 42. Verify equation (4.50) for $f(x) = \arccos x$ and $g(x) = \cos x$.

Exercise 43. Differentiate the identity (4.50) one more time with respect to x to obtain a new expressions relating higher-order derivatives of $f(x)$ and $g(x)$.

4.7 Inverse Functions of Several Variables

Equation (4.48) is the prototype of the matrix relationship between the Jacobians J and J'. Let us next generalize equation (4.48) for two sets of functions of two variables. Following that, we use the tensor notation to generalize to arbitrary dimension.

How does the concept of inverse functions generalize to two dimensions? For example, consider the transformations between Cartesian and polar coordinates. The two transformations (4.1), (4.2) and (4.3), (4.4) are the inverses of each other. That means that substituting $x(r,\theta)$ and $y(r,\theta)$ into $r(x,y)$ will recover r. (As an exercise, confirm that this is so.). Similarly, substituting $x(r,\theta)$ and $y(r,\theta)$ into $r(x,y)$ will recover θ. Express these relationships analytically:

$$r(x(r,\theta), y(r,\theta)) \equiv r \tag{4.51}$$

$$\theta(x(r,\theta), y(r,\theta)) \equiv \theta, \tag{4.52}$$

More generally, suppose that the sets of functions $f(x,y)$, $g(x,y)$ and $F(X,Y)$, $G(X,Y)$ are the inverses of each other. By analogy with (4.51) and (4.52), the inverse relationship can be captured by the identities

$$F(f(x,y), g(x,y)) = x \tag{4.53}$$

$$G(f(x,y), g(x,y)) = y. \tag{4.54}$$

From these two identities we can obtain four relationships for the partial derivatives of F, G, f and g. These relationships are by differentiating each of the identities with respect to each independent variables. Differentiating the first identity (4.53) with respect to each variable yields

$$\frac{\partial F}{\partial X}\frac{\partial f}{\partial x}+\frac{\partial F}{\partial Y}\frac{\partial g}{\partial x}=1 \tag{4.55}$$

$$\frac{\partial F}{\partial X}\frac{\partial f}{\partial y}+\frac{\partial F}{\partial Y}\frac{\partial g}{\partial y}=0. \tag{4.56}$$

Differentiating the second identity (4.54) with respect to each variable yields

$$\frac{\partial G}{\partial X}\frac{\partial f}{\partial x}+\frac{\partial G}{\partial Y}\frac{\partial g}{\partial x}=0 \tag{4.57}$$

$$\frac{\partial G}{\partial X}\frac{\partial f}{\partial y}+\frac{\partial G}{\partial Y}\frac{\partial g}{\partial y}=1. \tag{4.58}$$

Combining the four identities into a single matrix relationship yields

$$\begin{bmatrix} \frac{\partial F}{\partial X} & \frac{\partial F}{\partial Y} \\ \frac{\partial G}{\partial X} & \frac{\partial G}{\partial Y} \end{bmatrix}\begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{4.59}$$

which is precisely the relationship we were hoping to discover. It shows that the Jacobians J and J' for general coordinate changes are the matrix inverses of each another, provided they are evaluated at the same physical point. The one exciting task that remains is to generalize this relationship to an arbitrary dimension. We will do so in the next section with the help of the tensor notation. You will find that the tensor notation actually makes the N-dimensional case simpler than the two-dimensional case!

Exercise 44. What is the equivalent of equations (4.51) where the roles of Cartesian and polar coordinates are reversed?

$$x\,(\ldots)=x \tag{4.60}$$

$$y\,(\ldots)=y? \tag{4.61}$$

Exercise 45. Show that the second-order partial derivatives of F and G, and f and g are related by the equation

$$\frac{\partial^2 F}{\partial X^2}\left(\frac{\partial f}{\partial x}\right)^2+2\frac{\partial^2 F}{\partial X\partial Y}\frac{\partial f}{\partial x}\frac{\partial g}{\partial x}+\frac{\partial^2 F}{\partial Y^2}\left(\frac{\partial g}{\partial x}\right)^2+\frac{\partial F}{\partial X}\frac{\partial^2 f}{\partial x^2}+\frac{\partial F}{\partial Y}\frac{\partial^2 g}{\partial x^2}=0. \tag{4.62}$$

Derive the remaining three second-order relationships.

4.8 The Jacobian Property in Tensor Notation

In this section we generalize the inverse matrix relationship between Jacobians to arbitrary dimension. Along the way, we introduce an important new element of the tensor notation: the Kronecker delta symbol δ^i_j.

Consider two alternative coordinate systems Z^i and $Z^{i'}$ in an N-dimensional space. Notice that we placed the prime next to the index rather than the letter Z. Let us call the coordinates Z^i *unprimed* and the coordinates $Z^{i'}$ *primed*. We also use the symbols $Z^{i'}$ and Z^i to denote the functions that express the relationships between the coordinates:

$$Z^{i'} = Z^{i'}(Z) \tag{4.63}$$

$$Z^i = Z^i\left(Z'\right). \tag{4.64}$$

The first expression represents N relationships analogous to equations (4.1) and (4.2) and the second represents N relationships analogous to equations (4.3) and (4.4).

As sets of functions, Z^i and $Z^{i'}$ are the inverses of each other. This fact can be expressed either by the identity

$$Z^i\left(Z'(Z)\right) \equiv Z^i \tag{4.65}$$

or by the identity

$$Z^{i'}\left(Z\left(Z'\right)\right) \equiv Z^{i'}. \tag{4.66}$$

Note all the tensor conventions in play here. The indices of function arguments are suppressed. Several relationships are captured by a single expression with a live index. Note also, that in the first identity, the symbols Z^i (on the left-hand side) and Z' (which is $Z^{i'}$ with the index suppressed) represent functions, while the symbols Z (which is Z^i with the index suppressed) and Z^i (on the right-hand side) represent the independent variables. Equation (4.65) is analogous to equations (4.53) and (4.54) while equation (4.66) is analogous to equations

$$f\left(F(X,Y), G(X,Y)\right) = X \tag{4.67}$$

$$g\left(F(X,Y), G(X,Y)\right) = Y. \tag{4.68}$$

Let us analyze the first identity $Z^i\left(Z'(Z)\right) = Z^i$. It represents N relationships and each of the N relationships can be differentiated with respect to each of the N independent variables. This will yield N^2 relationships for the first-order partial derivatives of the functions Z^i and $Z^{i'}$.

With the help of the tensor notation, all of the operations can be carried out in a single step. We differentiate the identity $Z^i\left(Z'(Z)\right) = Z^i$ with respect to Z^j. It is essential that the differentiation is to take place with respect to Z^j rather than Z^i, because our intention is to differentiate *each* of the identities in $Z^i(Z'(Z)) = Z^i$ with respect to *each* of the variables. The resulting expression will have two *live* indices i and j.

The result of the differentiation reads

$$\frac{\partial Z^i}{\partial Z^{i'}} \frac{\partial Z^{i'}}{\partial Z^j} = \frac{\partial Z^i}{\partial Z^j}. \tag{4.69}$$

This single expression with two live indices i and j is analogous to equations (4.55)–(4.58). The objects $\partial Z^i / \partial Z^{i'}$ and $\partial Z^{i'} / \partial Z^j$ are precisely the Jacobians of the coordinate transformations. Let us now denote these Jacobians by the indicial symbols $J^i_{i'}$ and $J^{i'}_i$:

$$J^i_{i'} = \frac{\partial Z^i(Z')}{\partial Z^{i'}} \tag{4.70}$$

$$J^{i'}_i = \frac{\partial Z^{i'}(Z)}{\partial Z^i}. \tag{4.71}$$

Incorporating these symbols into equation (4.69), it reads

$$J^i_{i'} J^{i'}_j = \frac{\partial Z^i}{\partial Z^j} \tag{4.72}$$

and we must keep in mind that $J^i_{i'}$ and $J^{i'}_i$ must be evaluated at the same physical point.

Now, what of the partial derivative $\partial Z^i / \partial Z^j$? The symbol Z^i on the right-hand side of equation (4.65) represents the independent variable or the collection of the independent variables. The symbol Z^j also represents the independent variable. Therefore, $\partial Z^i / \partial Z^j$ is 1 if Z^i and Z^j are the same variable, and 0 otherwise. There is a special symbol that denotes such a quantity: the *Kronecker delta* δ^i_j also known as the *Kronecker symbol*. It is defined as follows

$$\delta^i_j = \begin{cases} 1, \text{ if } i = j \\ 0, \text{ if } i \neq j. \end{cases} \tag{4.73}$$

From the linear algebra point of view, the Kronecker delta can be thought of as the identity matrix. It is an object of utmost importance and will play a crucial role in the remainder of the textbook.

With the help of the Kronecker delta, we can rewrite equation (4.72) in its final form

$$J^i_{i'} J^{i'}_j = \delta^i_j. \tag{4.74}$$

This equation expresses the inverse relationships between the Jacobians that we have set out to demonstrate.

Exercise 46. By differentiating

$$Z^{i'}\left(Z\left(Z'\right)\right) = Z^{i'} \tag{4.75}$$

show that

$$J_i^{i'} J_{j'}^i = \delta_{j'}^{i'}. \tag{4.76}$$

Exercise 47. Derive equation (4.76) from (4.74) by multiplying both sides by $J_{j'}^j$.

Exercise 48. Introduce the symbols $J_{i'j'}^i$ and $J_{ij}^{i'}$ for the second-order partial derivatives

$$J_{i'j'}^i = \frac{\partial^2 Z^i\left(Z'\right)}{\partial Z^{i'} \partial Z^{j'}} \tag{4.77}$$

$$J_{ij}^{i'} = \frac{\partial^2 Z^{i'}\left(Z\right)}{\partial Z^i \partial Z^j}. \tag{4.78}$$

Show that

$$J_{i'j'}^i J_j^{i'} J_k^{j'} + J_{i'}^i J_{jk}^{i'} = 0. \tag{4.79}$$

How many identities does this tensor relationship represent?

Exercise 49. What is $J_{i'j'}^i$ for a transformation from one affine coordinate system to another?

Exercise 50. Derive the identity

$$J_{i'j'}^i J_j^{i'} + J_{i'}^i J_{jk}^{i'} J_{j'}^k = 0 \tag{4.80}$$

by contracting both sides of equation (4.79) with $J_{k'}^k$ and subsequently renaming $j' \to k'$. Verify this relationship for the transformation from Cartesian to polar coordinates.

Exercise 51. Show that

$$J_{i'j'}^i + J_{kj}^{k'} J_{k'}^i J_{j'}^j J_{i'}^k = 0. \tag{4.81}$$

Equations (4.79), (4.80), (4.81) will prove to be of critical importance in Chap. 8.

Exercise 52. Derive third-order relationships of the same kind.

Exercise 53. Consider three changes of variables: from Z^i to $Z^{i'}$ to $Z^{i''}$ and back to Z^i. Show that

$$J_{i'}^i J_{i''}^{i'} J_j^{i''} = \delta_j^i. \tag{4.82}$$

4.9 Several Notes on the Tensor Notation

We have already encountered many elements of the tensor notation—and seen many of its benefits. In this section, we point out a few properties of the notation that have already been used in this chapter, but could be clarified further.

4.9.1 *The Naming of Indices*

Indices are frequently needed to be renamed depending on the context in which an object is being used. Sometimes the name of the index does not matter at all. For example, the collection of the partial derivatives of a scalar field F can be referred to as $\partial F/\partial Z^i$ or $\partial F/\partial Z^j$—any letter (preferably a lowercase letter from the English alphabet) can be used to denote the index. Similarly, the Jacobian $J^i_{i'}$ can be called $J^j_{j'}$ or $J^i_{j'}$ or $J^j_{i'}$—the names of indices do not matter until the object is used in an expression. When referring to an object T^i, we may call it T^j or T^k or T^n depending on what fits the current context.

A repeated index always denotes contraction. In particular, a repeated index must appear exactly twice, once as upper and once as lower. For example, the collection of second-order partial derivatives

$$\frac{\partial^2 F}{\partial Z^i \partial Z^j} \tag{4.83}$$

cannot be denoted by

$$\frac{\partial^2 F}{\partial Z^i \partial Z^i}. \tag{4.84}$$

This expression *could* be interpreted as a second-order derivative with respect to the same variable, for instance, $\partial^2 F/\partial r^2$ and $\partial^2 F/\partial \theta^2$. However, tensor calculus does not interpret it this way. By convention, objects with repeated indices are disallowed, except in a contraction.

When objects are combined to form longer expressions, renaming indices becomes more restricted. Given a contraction, such as

$$\frac{\partial F}{\partial Z^i}\frac{\partial Z^i}{\partial \mu^\alpha}, \tag{4.85}$$

the index i, if renamed, should be renamed in both instances:

$$\frac{\partial F}{\partial Z^i}\frac{\partial Z^i}{\partial \mu^\alpha} = \frac{\partial F}{\partial Z^j}\frac{\partial Z^j}{\partial \mu^\alpha}. \tag{4.86}$$

The repeated index indicating contraction can be renamed arbitrarily, as long as the new name does not appear in another part of the term, causing a clash. For example, in the expression

$$\frac{\partial F}{\partial Z^i}\frac{\partial Z^i}{\partial \mu^\alpha}\frac{\partial F}{\partial Z^j}\frac{\partial Z^j}{\partial \mu^\beta} \tag{4.87}$$

the index i can be changed to any letter, except j!

Finally, in tensor identities, *live* indices must be renamed consistently. For example, suppose that

$$T_\alpha = \frac{\partial F}{\partial Z^i}\frac{\partial Z^i}{\partial \mu^\alpha}. \tag{4.88}$$

Then α can be changed to β on both sides of the identity:

$$T_\beta = \frac{\partial F}{\partial Z^j}\frac{\partial Z^j}{\partial \mu^\beta}. \tag{4.89}$$

Notice that, for good measure, we also renamed the dummy index from i to j.

4.9.2 Commutativity of Contractions

We have already encountered several expressions that include two or more contractions. One such expression was $J^i_{i'}J^{i'}_{jk}J^k_{j'}$. Should we have been concerned with the order of contractions? The answer is *no* because contractions commute, that is, they can be carried out in any order.

For an illustration, consider an object T^{ij}_{kl} with four indices and the double contraction T^{ij}_{ij}. The expression T^{ij}_{ij} can be interpreted in three ways. First, we can imagine that contraction on i takes place first, yielding the object $U^j_l = T^{ij}_{il}$. Fully expanding, U^j_l is the sum

$$U^j_l = T^{1j}_{1l} + T^{2j}_{2l} + T^{3j}_{3l}. \tag{4.90}$$

Subsequently, contraction on the remaining indices yields $U^j_j = T^{ij}_{ij}$:

$$T^{ij}_{ij} = T^{11}_{11} + T^{12}_{12} + T^{13}_{13} + T^{21}_{21} + T^{22}_{22} + T^{23}_{23} + T^{31}_{31} + T^{32}_{32} + T^{33}_{33}. \tag{4.91}$$

Alternatively, we can contract on j first (yielding a temporary object $V^i_k = T^{ij}_{kj}$) and subsequently on i. You can confirm that this sequence will result in the same sum. Finally, the contractions can be viewed as simultaneous. That is, the terms in the sum are formed by letting the indices i and j run through the N^2 possible combinations. It is clear that in this case the end result is also the same.

4.9.3 More on the Kronecker Symbol

The Kronecker delta δ^i_j will prove to be an extremely important object. Its importance goes beyond its linear algebra interpretation as the identity matrix. For now, we would like to say a few words about its index-renaming effect in indicial expressions.

Consider the expression $T_i \delta^i_j$. What is the value of this object for $j = 1$? It is

$$T_i \delta^i_1 = T_1 \delta^1_1 + T_2 \delta^2_1 + T_3 \delta^3_1. \tag{4.92}$$

The only surviving term is $T_1 \delta^1_1$ which equals T_1. Similarly for $j = 2$, $T_i \delta^i_j$ is T_2. And so forth. Summarizing,

$$T_i \delta^i_j = T_j. \tag{4.93}$$

Thus, the effect of contracting T_i with δ^i_j is to rename the index i to j. Similarly,

$$T^i \delta^j_i = T^j. \tag{4.94}$$

The last two equations show why the Kronecker delta can also be viewed as the identity matrix: when contracted with an object, the result is the original object.

Exercise 54. Simplify the expression $A_{ij} \delta^i_k$.

Exercise 55. Simplify the expression $\delta^i_j \delta^j_k$.

Exercise 56. Evaluate the expression $\delta^i_j \delta^j_i$.

Exercise 57. Evaluate the expression $\delta^i_i \delta^j_j$.

Exercise 58. Simplify the expression $\delta^i_j S_i S^j$.

Exercise 59. Show that the object $\delta^r_i \delta^s_j - \delta^s_i \delta^r_j$ is skew-symmetric in i and j as well as in r and s.

4.10 Orientation-Preserving Coordinate Changes

Let J be the determinant of the Jacobian $J^i_{i'}$. We express this relationship in the following way

$$J = \left| J^{\cdot}_{\cdot'} \right|. \tag{4.95}$$

The coordinate change is called *orientation-preserving* if $J > 0$. The coordinate change is called *orientation-reversing* if $J < 0$. These definitions are local: the orientation-preserving property may vary from one point to another. Since the Jacobians $J^i_{i'}$ and $J^{i'}_i$ are matrix inverses of each other, the determinant of $J^{i'}_i$ is J^{-1}. Therefore, if a change of coordinates is orientation-preserving, then so is its inverse. The orientation-preserving property of coordinate changes is particularly important for the discussion of the volume element and the Levi-Civita symbols in Chap. 9.

If two coordinate systems are related by an orientation-preserving change of coordinates, then one of those coordinate systems is said to be *positively oriented* with respect to the other. In the three-dimensional space, the right-hand rule gives us an opportunity to characterize the orientation of a coordinate system in an absolute sense. A Cartesian coordinate system is *positive* if its coordinate basis vectors **i**, **j**, **k** are ordered according to the right-hand rule. Any other coordinate system is called *positive* in the absolute sense if it is positively oriented with respect a positive Cartesian coordinate system.

4.11 Summary

In this chapter, we discussed coordinate changes. Tensor calculus achieves invariance by constructing expressions where the eventual values remain unchanged under a change of coordinates. This is why the discussion of coordinate changes is critical. In the next chapter, we begin to introduce some of the most important objects in Euclidean spaces. In Chap. 6, we study how these objects transform under a change of coordinates. We find that some objects, characterized as tensors, transform according to a very special rule. Other objects (e.g., partial derivatives of tensors) do not transform according to the same rule. This lack of the tensor property poses a problem in constructing invariant expressions. This problem is addressed in Chap. 8 which introduces a new kind of tensor-property-preserving differential operator: the covariant derivative.

In this chapter we also presented all the essential elements of the tensor notation. We saw how the indicial notation leads to very compact expressions that are easily interpreted. We also observed that indices naturally arise as lower indices and upper indices, and that the two types of indices seem to be in perfect balance. This is not a coincidence: the placement of the index helps distinguish between covariant and contravariant tensors.

Chapter 5
The Tensor Description of Euclidean Spaces

5.1 Preview

Armed with the tensor notation introduced in Chap. 4, we present a differential description of Euclidean spaces in tensor terms. The objects presented here are of great importance in the coordinate-based description of space and play a central role in the remainder of this book.

All objects, except the position vector **R**, introduced in this chapter depend on the choice of coordinates. That is, if the same definition is applied in different coordinates, the resulting values are different. Such objects are called *variants*. Objects, such as the position vector **R**, the values of which are the same in all coordinates, are called *invariants*. For some variants, their values in different coordinate systems are related in a particular way. Such objects are called *tensors*. The discussion of the tensor property takes place in Chap. 6.

5.2 The Position Vector R

The *position vector* (or the *radius vector*) is a vector **R** that represents the position of points in the Euclidean space with respect to an arbitrarily selected point O, known as the *origin*. The concept of the position vector is possible only in Euclidean spaces. After all, the position vector is straight, and straightness is a characteristic of the Euclidean space. If we were studying the surface of a sphere and were allowed to consider only the points on the surface and not in the surrounding space, we would not be able to introduce the position vector.

The position vector **R** is introduced without the help of coordinates. Objects that can be constructed without a reference to a coordinate system are called *geometric*. The vector gradient of a scalar field is another example of a geometric object. Some of the objects we construct with the help of coordinate systems turn out to be independent of coordinates. Such objects are called *invariant*. Experience

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*, DOI 10.1007/978-1-4614-7867-6_5, © Springer Science+Business Media New York 2013

shows that all *invariant* objects are *geometric*. As a result, the terms *geometric* and *invariant* are often used interchangeably. That is not to say that finding the geometric interpretation for an invariant object is always straightforward.

5.3 The Position Vector as a Function of Coordinates

Refer the Euclidean space to a coordinate system Z^i. In the spirit of tensor calculus, the coordinate system Z^i is *arbitrary*. In other words, we do not assume that it possesses any special characteristics. In particular, it is certainly not assumed Cartesian or affine. In the three-dimensional Euclidean space, the index i has values $1, 2, 3$ and Z^i stands for Z^1, Z^2, Z^3. When reading this chapter, it is probably best to keep the three-dimensional case in mind. Of course, all of the concepts are valid in any number of dimensions.

In a Euclidean space referred to a coordinate system, the position vector $\mathbf{R}$ is a function of the coordinates Z^i. That is, to each valid combination of Z^1, Z^2, and Z^3, there corresponds a specific value of the position vector $\mathbf{R}$. The coordinates uniquely determine the point in the Euclidean space which, in turn, corresponds to a particular value of the position vector $\mathbf{R}$. Denote this function by $\mathbf{R}(Z)$:

$$\mathbf{R} = \mathbf{R}(Z). \tag{5.1}$$

This notation is consistent with the convention described in Sect. 4.5 according to which we suppress the indices of the function arguments.

The function $\mathbf{R}(Z)$ is the starting point in an important sequence of identities. It is therefore important to have a clear understanding of each of the elements of equation (5.1). This equation features the symbol $\mathbf{R}$ in two different roles. On the left-hand side, $\mathbf{R}$ represents the geometric position vector introduced in Sect. 5.2. On the right-hand side, $\mathbf{R}$ stands for the vector-valued function that yields the position vector for every valid combination of coordinates. As it is often the case, the same letter is used to represent two different objects. We distinguish among the objects denoted by the same letter by the full *signature* of the expression. In the case of equation (5.1), the signature includes the collection of arguments. Thus, $\mathbf{R}$ is the invariant position vector while $\mathbf{R}(Z)$ is a vector-valued function of three variables.

The function $\mathbf{R}(Z)$ can be differentiated with respect to each of its arguments in the sense of Sect. 2.9. We reiterate that it is essential to treat $\mathbf{R}$ as a primary object that is subject to its own set of rules and operations. It is counterproductive to think of $\mathbf{R}$ as a triplet of components with respect to a some Cartesian grid. We have not and will not (unless it is convenient for a particular problem) introduce a background Cartesian coordinate system. Thus, there is no basis with respect to which $\mathbf{R}$ could be decomposed. Fortunately, we can perform quite a number of operations on vector quantities such as $\mathbf{R}$, including differentiation with respect to a parameter.

Exercise 60. Consider a curve parametrized by its arc length s, and consider the function $\mathbf{R}(s)$. Explain, on the basis of your geometric intuition, why $\mathbf{T} = \mathbf{R}'(s)$ is the unit tangent vector. Note that the condition $|\mathbf{R}'(s)| = 1$ may be taken as the definition of arc length s.

5.4 The Covariant Basis $\mathbf{Z}_i$

The *covariant basis* is obtained from the position vector $\mathbf{R}(Z)$ by differentiation with respect to each of the coordinates:

$$\mathbf{Z}_i = \frac{\partial \mathbf{R}(Z)}{\partial Z^i}. \tag{5.2}$$

The term *covariant* refers to the special way in which the vectors $\mathbf{Z}_i$ transform under a change of coordinates. This is studied in Chap. 6.

The covariant basis is a generalization of the *affine coordinate basis* to curvilinear coordinate systems. It is called the *local coordinate basis* since it varies from one point to another. It is sometimes described as the *local basis imposed by the coordinate system*. The letter Z is now used to denote an additional object. However, the symbols $\mathbf{Z}_i$ and Z^i cannot be mistaken for each other since the bold $\mathbf{Z}$ in $\mathbf{Z}_i$ is a vector while the plain Z in Z^i is a scalar.

The tensor notation may still be quite new to you. Therefore, let us expand the relationships captured by equation (5.2). Since the index i is live, equation (5.2) defines three vectors $\mathbf{Z}_1$, $\mathbf{Z}_2$, and $\mathbf{Z}_3$:

$$\mathbf{Z}_1 = \frac{\partial \mathbf{R}\left(Z^1, Z^2, Z^3\right)}{\partial Z^1}; \quad \mathbf{Z}_2 = \frac{\partial \mathbf{R}\left(Z^1, Z^2, Z^3\right)}{\partial Z^2}; \quad \mathbf{Z}_3 = \frac{\partial \mathbf{R}\left(Z^1, Z^2, Z^3\right)}{\partial Z^3}. \tag{5.3}$$

The argument-collapsing convention of Sect. 4.5 shortens these equation to

$$\mathbf{Z}_1 = \frac{\partial \mathbf{R}(Z)}{\partial Z^1}; \quad \mathbf{Z}_2 = \frac{\partial \mathbf{R}(Z)}{\partial Z^2}; \quad \mathbf{Z}_3 = \frac{\partial \mathbf{R}(Z)}{\partial Z^3}. \tag{5.4}$$

Finally, encoding this set of identities with the index i transforms them to the ultimate tensor form (5.2).

Figure 5.1 shows the covariant basis $\mathbf{Z}_i$ in the plane referred to Cartesian, affine, and polar coordinate systems. In affine coordinates, the covariant basis $\mathbf{Z}_i$ coincides with the coordinate basis $\mathbf{i}, \mathbf{j}$. In curvilinear coordinates, the covariant basis varies from point to point.

The covariant basis $\mathbf{Z}_i$ is a *primary* object. Once again, you should not imagine a background Cartesian coordinate system and decompose the vectors $\mathbf{Z}_i$ with respect to that system's coordinate basis. It is counterproductive to think of $\mathbf{Z}_i$ as a set of N^2 coordinates, where N is the dimension of the space. The vectors $\mathbf{Z}_i$ are just that: vectors.

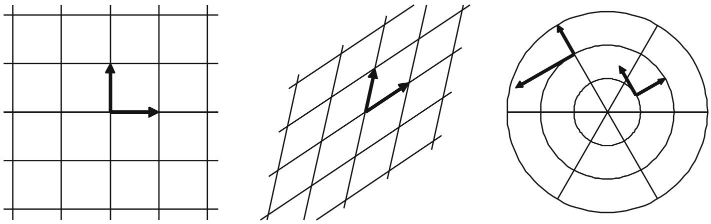


Fig. 5.1 The covariant basis $\mathbf{Z}_i$ in Cartesian, affine, and polar coordinates in two dimensions

At all points in the Euclidean space, the covariant basis $\mathbf{Z}_i$ provides a convenient basis for decomposing other vectors. The components V^i of a vector $\mathbf{V}$ are the scalar values that produce $\mathbf{V}$ when used in a linear combination with the vectors $\mathbf{Z}_i$:

$$\mathbf{V} = V^1\mathbf{Z}_1 + V^2\mathbf{Z}_2 + V^3\mathbf{Z}_3. \tag{5.5}$$

In the tensor notation

$$\mathbf{V} = V^i\mathbf{Z}_i. \tag{5.6}$$

The values V^i are called the *contravariant components* of the vector $\mathbf{V}$. The term contravariant refers to the way in which the object V^i transforms under a change of coordinates. This is discussed in Chap. 6. Since $\mathbf{Z}_i$ varies from one point to another, two identical vectors $\mathbf{U}$ and $\mathbf{V}$ decomposed at different points of the Euclidean space may have different contravariant components.

Exercise 61. What are the components of the vectors $\mathbf{Z}_i$ with respect to the covariant basis $\mathbf{Z}_j$? The answer can be captured with a single symbol introduced in Chap. 4.

5.5 The Covariant Metric Tensor Z_{ij}

By definition, the *covariant metric tensor* Z_{ij} consists of the pairwise dot products of the covariant basis vectors:

$$Z_{ij} = \mathbf{Z}_i \cdot \mathbf{Z}_j. \tag{5.7}$$

The metric tensor is one of the central objects in tensor calculus. It is often referred to as the *fundamental tensor*. In linear algebra, the covariant metric tensor is known as the *inner product matrix* or the *Gram matrix*. Because the dot product is commutative ($\mathbf{Z}_i \cdot \mathbf{Z}_j = \mathbf{Z}_j \cdot \mathbf{Z}_i$), the metric tensor is symmetric:

$$Z_{ij} = Z_{ji}. \tag{5.8}$$

It carries complete information about the dot product and is therefore the main tool in measuring lengths, areas, and volumes. Suppose that two vectors **U** and **V** are located at the same point and that their components are U^i and V^i. Then the dot product $\mathbf{U} \cdot \mathbf{V}$ is given by

$$\mathbf{U} \cdot \mathbf{V} = Z_{ij} U^i V^j. \tag{5.9}$$

This can be demonstrated as follows:

$$\mathbf{U} \cdot \mathbf{W} = U^i \mathbf{Z}_i \cdot V^j \mathbf{Z}_j = \left(\mathbf{Z}_i \cdot \mathbf{Z}_j\right) U^i V^j = Z_{ij} U^i V^j.$$

Exercise 62. Explain why the covariant basis, interpreted as a matrix, is positive definite.

Exercise 63. Show that the length $|\mathbf{V}|$ of a vector **V** is given by

$$|\mathbf{V}| = \sqrt{Z_{ij} V^i V^j}. \tag{5.10}$$

5.6 The Contravariant Metric Tensor Z^{ij}

The *contravariant metric tensor* Z^{ij} is the matrix inverse of Z_{ij}:

$$Z^{ij} Z_{jk} = \delta^i_k. \tag{5.11}$$

From linear algebra, we know that $AB = I$ implies $BA = I$ (the left inverse is the right inverse). Therefore, contracting on the other index also yields in the Kronecker symbol:

$$Z^{ij} Z_{kj} = \delta^i_k. \tag{5.12}$$

This relationship also follows from the symmetry of Z_{ij}, but the inverse argument is more general.

Since the inverse of a symmetric matrix is symmetric (prove this), the contravariant metric tensor is symmetric

$$Z^{ij} = Z^{ji}. \tag{5.13}$$

Exercise 64. Prove that Z^{ij} is positive definite.

5.7 The Contravariant Basis $\mathbf{Z}^i$

The *contravariant basis* $\mathbf{Z}^i$ is defined as

$$\mathbf{Z}^i = Z^{ij}\mathbf{Z}_j. \tag{5.14}$$

Since Z^{ij} is symmetric, it makes no difference which index is contracted in equation (5.14). Thus we may also write

$$\mathbf{Z}^i = Z^{ji}\mathbf{Z}_j. \tag{5.15}$$

The bases $\mathbf{Z}_i$ and $\mathbf{Z}^i$ are mutually orthonormal:

$$\mathbf{Z}^i \cdot \mathbf{Z}_j = \delta^i_j. \tag{5.16}$$

That is, each vector $\mathbf{Z}^i$ is orthogonal for each $\mathbf{Z}_{j,}$ for which $i \neq j$. Furthermore, the dot product of $\mathbf{Z}^i$ with $\mathbf{Z}_i$ is 1. The latter relationship cannot be written as $\mathbf{Z}^i \cdot \mathbf{Z}_i = 1$ since the repeated index would invoke the summation convention. Sometimes, uppercase indices are used to indicate a relationship for a single value of the repeated index: $\mathbf{Z}^I \cdot \mathbf{Z}_I = 1$.

Exercise 65. Demonstrate equation (5.16). Hint: $\mathbf{Z}^i \cdot \mathbf{Z}_j = Z^{ik}\mathbf{Z}_k \cdot \mathbf{Z}_j$.

Exercise 66. Show that the angle between $\mathbf{Z}^1$ and $\mathbf{Z}_1$ is less than $\pi/2$.

Exercise 67. Show that the product of the lengths $\left|\mathbf{Z}^1\right| |\mathbf{Z}_1|$ is at least 1.

Exercise 68. Show that the covariant basis $\mathbf{Z}_i$ can be obtained from the contravariant basis $\mathbf{Z}^j$ by

$$\mathbf{Z}_i = Z_{ij}\mathbf{Z}^j. \tag{5.17}$$

Hint: Multiply both sides of equation (5.14) by Z_{ik}.

Exercise 69. Show that the pairwise dot products of the contravariant basis elements yields the contravariant metric tensor:

$$\mathbf{Z}^i \cdot \mathbf{Z}^j = Z^{ij}. \tag{5.18}$$

Exercise 70. Argue from the linear algebra point of view that equation (5.16) uniquely determines the vectors $\mathbf{Z}^i$.

Exercise 71. Therefore, equation (5.16) can be taken as the definition of the contravariant basis $\mathbf{Z}^i$ and equation (5.18) can be taken as the definition of the contravariant basis Z^{ij}. Show that the property (5.11) follows these definitions.

5.8 The Metric Tensor and Measuring Lengths

The metric tensor plays a fundamental role in calculations. It is responsible for calculating lengths of vectors since $|\mathbf{V}| = \sqrt{Z_{ij}V^iV^j}$. It is also the key to calculating lengths of curves. You are familiar with the following formula for the length L of a curve given in Cartesian coordinates by $(x(t), y(t), z(t))$:

$$L = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} dt. \tag{5.19}$$

Note that this formula may also be taken as the definition of the length of a curve. (See Exercise 60 for an alternative equivalent approach.) What is the generalization of equation (5.19) to arbitrary coordinates Z^i in which the curve is given by $\left(Z^1(t), Z^2(t), Z^3(t)\right)$. The general formula for L is

$$L = \int_a^b \sqrt{Z_{ij}\frac{dZ^i}{dt}\frac{dZ^j}{dt}} dt. \tag{5.20}$$

Why is the formula (5.20) correct? In Cartesian coordinates, the metric tensor Z_{ij} is represented by the identity matrix. Therefore, since only the *diagonal* terms survive, equation (5.20) agrees with equation (5.19). So we are on the right track! The complete proof of equation (5.20) is found in Chap. 6. The key to the proof is the tensor property of all the elements in equation (5.20).

Equation (5.20) can be interpreted in another fascinating way. Suppose you find yourself in a Euclidean space referred to some coordinate system Z^i. Suppose also, that the only measurement you are allowed to take is the length of any curve segment. From equation (5.20), it follows that this would allow you to calculate the metric tensor at any point. Thus, lengths of curves can be calculated from the metric tensor and *vice versa*.

Consider the point $P = \left(Z_0^1, Z_0^2, Z_0^3\right)$ and the curve given parametrically by $\left(Z_0^1 + t, Z_0^2, Z_0^3\right)$. This curve is the coordinate line corresponding to Z^1 that passes through P at $t = 0$. Denote the length of this curve from the point P by $L_{11}(t)$. Since only a single term survives under the square root in equation (5.20),

$$L_{11}(t) = \int_0^t \sqrt{Z_{11}(h)} dh. \tag{5.21}$$

By the fundamental theorem of calculus,

$$L'_{11}(0) = \sqrt{Z_{11}} \text{ at } P. \tag{5.22}$$

We can similarly compute the remaining *diagonal* entries of Z_{ij}.

To compute the off-diagonal terms, say Z_{12}, consider the curve $\left(Z_0^1 + t, Z_0^2 + t, Z_0^3\right)$ and denote its length from the point P by $L_{12}(t)$. This time, three terms survive under the square root in equation (5.20):

$$L_{12}(t) = \int_0^t \sqrt{Z_{11}(h) + 2Z_{12}(h) + Z_{22}(h)}\, dh. \tag{5.23}$$

Thus,

$$L'_{12}(0) = \sqrt{Z_{11} + 2Z_{12} + Z_{22}} \text{ at } P. \tag{5.24}$$

Since the diagonal terms Z_{11} and Z_{22} are already available, we are now able to compute Z_{12} and, similarly, all the remaining entries of the metric tensor.

5.9 Intrinsic Objects and Riemann Spaces

Objects, such as the metric tensor, which can be obtained by measuring distances and computing the derivatives of those distances, are called *intrinsic*. Therefore, all objects that can be expressed in terms of the metric tensor and *its* derivatives are also intrinsic. Most notably, the Christoffel symbol presented below and the Riemann–Christoffel tensor presented in Chap. 8 are intrinsic. Intrinsic objects are particularly important in general relativity and in the study of embedded surfaces to which Parts II and III of this book are devoted.

It turns out that much of differential geometry and tensor calculus can be constructed from the metric tensor as the starting point, bypassing the position vector $\mathbf{R}$ and the covariant basis $\mathbf{Z}_i$. Differential geometry that starts with the metric tensor is called *Riemannian geometry*. Its exclusive focus is on intrinsic objects. In Riemannian geometry, the dot product is defined in terms of the metric tensor, rather than the other way around. Thus, the metric tensor is a primary concept and equation (5.9) is the *definition* of the dot product. The properties of the space are dictated by the choice of the metric tensor. The metric tensor need not come from a Euclidean dot product, but can be rather arbitrarily assigned subject to certain conditions such as symmetry and positive definiteness. A coordinate space in which the metric tensor field is a priori given is called a *Riemann space*. The Riemannian perspective is critically important, even in Euclidean spaces.

5.10 Decomposition with Respect to a Basis by Dot Product

Linear algebra offers an elegant algorithm for calculating the components of a vector with respect to an *orthogonal basis*. Suppose that $\mathbf{e}_1, \ldots, \mathbf{e}_N$ is an orthogonal basis, that is

$$\left(\mathbf{e}_i, \mathbf{e}_j\right) = 0, \text{ if } i \neq j. \tag{5.25}$$

Let V_i be the components of a vector $\mathbf{V}$ with respect to this basis:

$$\mathbf{V} = V_1\mathbf{e}_1 + \cdots + V_N\mathbf{e}_N. \tag{5.26}$$

Dot both sides of this equation with $\mathbf{e}_i$. Since all but one term on the right-hand side vanish by the orthogonality of the basis, we find

$$(\mathbf{V}, \mathbf{e}_i) = V_i\,(\mathbf{e}_i, \mathbf{e}_i). \tag{5.27}$$

Therefore, the coefficient V_i is given by the simple expression

$$V_i = \frac{(\mathbf{V}, \mathbf{e}_i)}{(\mathbf{e}_i, \mathbf{e}_i)}. \tag{5.28}$$

This expression simplifies further if the basis is *orthonormal*, that is, each basis vector is unit length:

$$(\mathbf{e}_i, \mathbf{e}_i) = 1. \tag{5.29}$$

With respect to an orthonormal basis, V_i is given by a single dot product

$$V_i = (\mathbf{V}, \mathbf{e}_i). \tag{5.30}$$

If the basis $\mathbf{e}_1, \ldots, \mathbf{e}_N$ is not orthogonal, it is still possible to calculate V_i by evaluating dot products. However, the simplicity of the algorithm is lost. Instead of being able to determine each component V_i individually, one needs to solve an $N \times N$ system of coupled equations

$$\begin{bmatrix} (\mathbf{e}_1, \mathbf{e}_1) & \cdots & (\mathbf{e}_1, \mathbf{e}_N) \\ \cdots & \cdots & \cdots \\ (\mathbf{e}_N, \mathbf{e}_1) & \cdots & (\mathbf{e}_N, \mathbf{e}_N) \end{bmatrix} \begin{bmatrix} V_1 \\ \cdots \\ V_N \end{bmatrix} = \begin{bmatrix} (\mathbf{V}, \mathbf{e}_1) \\ \cdots \\ (\mathbf{V}, \mathbf{e}_N) \end{bmatrix}. \tag{5.31}$$

Note that the matrix on the left-hand side corresponds to the metric tensor Z_{ij}.

In tensor calculus, this relatively complicated procedure is captured by a formula that is as simple as equation (5.30). Namely, suppose that V^i are the contravariant components of the vector $\mathbf{V}$ with respect to $\mathbf{Z}_i$. Then

$$V^i = \mathbf{V} \cdot \mathbf{Z}^i. \tag{5.32}$$

Note that $\mathbf{V}$ is dotted with the *contravariant* basis $\mathbf{Z}^i$. The proof of this relationship is simple

$$\mathbf{V} \cdot \mathbf{Z}^i = V^j \mathbf{Z}_j \cdot \mathbf{Z}^i = V^j \delta^i_j = V^i. \tag{5.33}$$

The formula in equation (5.32) is equivalent to equation (5.31) but has the simplicity of equation (5.30). Equation (5.32) states that V^i can be calculated by single dot product even if the basis is not orthogonal. Of course, the single dot product involves the contravariant vector $\mathbf{Z}^i$, the calculation of which involved the matrix inversion of the covariant metric tensor. Therefore, the algorithm implied by equation (5.32) is no simpler than that implied by equation (5.31). It is the same algorithm. However, its expression in tensor calculus is far more effective and concise.

5.11 The Fundamental Elements in Various Coordinates

This section catalogs the fundamental elements in the Euclidean space for the most common coordinate systems. In affine and polar coordinates, all elements can be calculated easily by geometric reasoning. Spherical coordinates pose a little bit more of a challenge. In order to calculate covariant basis $\mathbf{Z}_i$ in spherical coordinates, a background Cartesian coordinate system is introduced. Earlier, we emphatically argued against background Cartesian coordinates. On the other hand, this section deals with a *particular* problem that can be solved more easily by choosing a convenient coordinate system. However, it is important to continue to think of the bases $\mathbf{Z}_i$ and $\mathbf{Z}^i$ as primary objects that do not require another, more fundamental, coordinate system in order to be expressed. To reinforce this perspective, we illustrate the covariant basis $\mathbf{Z}_i$ in Fig. 5.5 without the scaffolding of the background Cartesian coordinates.

5.11.1 Cartesian Coordinates

The covariant basis $\mathbf{Z}_i$ coincides with the coordinate basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$. The coordinate basis arises naturally in affine (and therefore Cartesian) coordinates as a set of vectors that point along the coordinate axes and have the length of one corresponding coordinate unit.

The geometric algorithm, by which the derivative $\mathbf{Z}_1 = \partial \mathbf{R} / \partial x$ is evaluated, is illustrated in Fig. 5.2. For any h, the vector

$$\mathbf{V} = \frac{\mathbf{R}(x+h, y, z) - \mathbf{R}(x, y, z)}{h} \tag{5.34}$$

is the unit vector that points in the direction of the x-axis. Therefore, $\mathbf{V}$ is the Cartesian basis element $\mathbf{i}$. The limit of $\mathbf{V}$ as $h \to 0$ is also $\mathbf{i}$, and we conclude that the covariant basis $\mathbf{Z}_i$ coincides with the coordinate basis:

$$\mathbf{Z}_1 = \mathbf{i}; \quad \mathbf{Z}_2 = \mathbf{j}; \quad \mathbf{Z}_3 = \mathbf{k}.$$

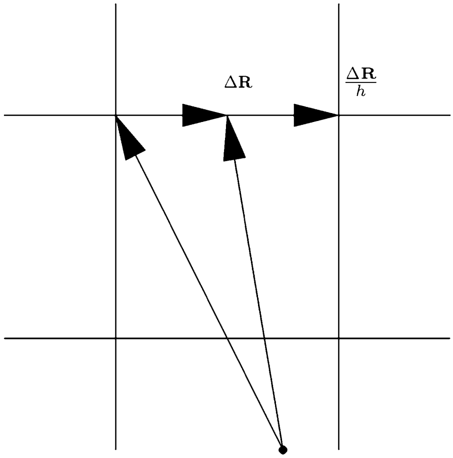


Fig. 5.2 Geometric construction of the basis vector $\mathbf{Z}_1 = \partial\mathbf{R}/\partial x$ in Cartesian coordinates

Consequently, the covariant and the contravariant metric tensors have only three nonzero entries:

$$Z_{11} = Z_{22} = Z_{33} = 1 \tag{5.35}$$

$$Z^{11} = Z^{22} = Z^{33} = 1. \tag{5.36}$$

In the matrix notation, the metric tensors are presented by the identity matrix:

$$Z_{ij}, Z^{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.37}$$

The contravariant basis $\mathbf{Z}^i$ consists of the same vectors as the covariant basis $\mathbf{Z}_i$:

$$\mathbf{Z}^1 = \mathbf{i}; \quad \mathbf{Z}^2 = \mathbf{j}; \quad \mathbf{Z}^3 = \mathbf{k}. \tag{5.38}$$

5.11.2 Affine Coordinates

A similar geometric argument shows that in affine coordinates, the covariant basis $\mathbf{Z}_i$ coincides with the coordinate basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$. The metric tensor Z_{ij} is therefore unchanged from one point to another. Its entries are pairwise dot products of the coordinate basis vectors:

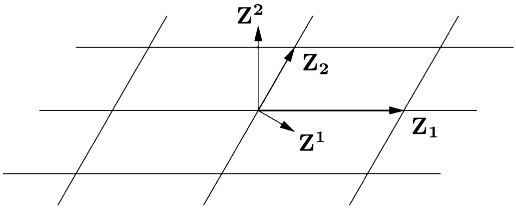


Fig. 5.3 In affine coordinates, the covariant basis $\mathbf{Z}_i$ lines up perfectly with the coordinate grid. This figure also shows the contravariant basis $\mathbf{Z}^i$. Note the orthogonality of the vectors $\mathbf{Z}_1 = \mathbf{i}$ and $\mathbf{Z}^1$, and $\mathbf{Z}_2 = \mathbf{j}$ and $\mathbf{Z}^2$

$$Z_{ij} = \begin{bmatrix} \mathbf{i}\cdot\mathbf{i} & \mathbf{i}\cdot\mathbf{j} & \mathbf{i}\cdot\mathbf{k} \\ \mathbf{j}\cdot\mathbf{i} & \mathbf{j}\cdot\mathbf{j} & \mathbf{j}\cdot\mathbf{k} \\ \mathbf{k}\cdot\mathbf{i} & \mathbf{k}\cdot\mathbf{j} & \mathbf{k}\cdot\mathbf{k} \end{bmatrix}. \tag{5.39}$$

The metric tensor is diagonal only if the affine coordinate system is orthogonal. The contravariant metric Z^{ij} is obtained by inverting the matrix (5.39) and the contravariant basis is obtained by contracting Z^{ij} with $\mathbf{Z}_i$.

We illustrate this calculation with an example in two dimensions illustrated in Fig. 5.3. Suppose that the affine coordinates are such that the angle between $\mathbf{Z}_1 = \mathbf{i}$ and $\mathbf{Z}_2 = \mathbf{j}$ is $\pi/3$ and that $|\mathbf{Z}_1| = 2$ and $|\mathbf{Z}_2| = 1$. Then

$$Z_{ij} = \begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix} \tag{5.40}$$

and

$$Z^{ij} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{4}{3} \end{bmatrix}. \tag{5.41}$$

The contravariant basis vectors are therefore given by

$$\mathbf{Z}^1 = \frac{1}{3}\mathbf{i} - \frac{1}{3}\mathbf{j} \tag{5.42}$$

$$\mathbf{Z}^2 = -\frac{1}{3}\mathbf{i} + \frac{4}{3}\mathbf{j}. \tag{5.43}$$

Exercise 72. Confirm that $\mathbf{Z}^1$ in equation (5.42) is orthogonal to $\mathbf{Z}_2$ and that $\mathbf{Z}^2$ in equation (5.43) is orthogonal to $\mathbf{Z}_1$.

Exercise 73. Confirm that $\mathbf{Z}^1 \cdot \mathbf{Z}_1 = \mathbf{Z}^2 \cdot \mathbf{Z}_2 = 1$.

Exercise 74. Confirm that $\mathbf{Z}^1 \cdot \mathbf{Z}^2 = Z^{12} = -1/3$.

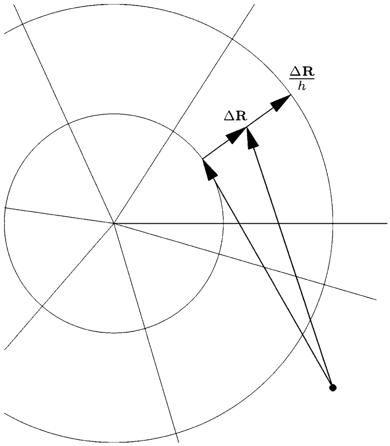


Fig. 5.4 Geometric construction of the vector $\mathbf{Z}_1 = \partial\mathbf{R}/\partial r$ in polar coordinates

5.11.3 Polar and Cylindrical Coordinates

Consider polar coordinates in the plane. The geometric construction that leads to $\mathbf{Z}_1$ is illustrated in Fig. 5.4. It is clear that the vector

$$\mathbf{V} = \frac{\mathbf{R}(r+h,\theta) - \mathbf{R}(r,\theta)}{h} \tag{5.44}$$

points radially away from the origin O. Furthermore, the length of the vector $\Delta\mathbf{R} = \mathbf{R}(r+h,\theta) - \mathbf{R}(r,\theta)$ in the denominator is h. Therefore, $\mathbf{V} = \Delta\mathbf{R}/h$ is unit length for all h. We conclude that $\mathbf{Z}_1$ is a unit vector that points in the radial direction.

The vector $\mathbf{Z}_2$ is more challenging, but we have already performed the calculation that gives $\mathbf{Z}_2$ in Sect. 2.9 of Chap. 2. In that section, we showed that, for $r = 1$, $\mathbf{Z}_2$ is the unit vector in the direction orthogonal to $\mathbf{Z}_1$. It is also easy to see that the length of $\mathbf{Z}_2$ is proportional to r by noticing that the denominator in

$$\mathbf{V} = \frac{\mathbf{R}(r,\theta+h) - \mathbf{R}(r,\theta)}{h} \tag{5.45}$$

is exactly twice as large at $r = 2$ as it is at $r = 1$. This relationship also holds in the limit as $h \to 0$ and we conclude that $\mathbf{Z}_2$ is a vector of length r orthogonal to $\mathbf{Z}_1$. The covariant basis for polar coordinates is seen in Fig. 5.1.

A coordinate system with an orthogonal covariant basis is called *orthogonal*. Thus, polar coordinates are orthogonal. Orthogonal coordinate systems are characterized by diagonal metric tensors. The metric tensors for polar coordinates are given by

$$Z_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix}; \quad Z^{ij} = \begin{bmatrix} 1 & 0 \\ 0 & r^{-2} \end{bmatrix}. \tag{5.46}$$

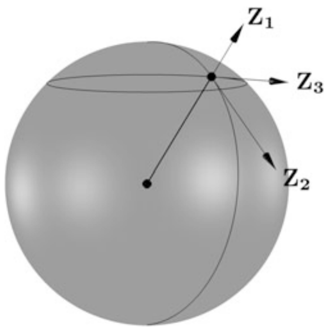


Fig. 5.5 The covariant basis Z_i at the point $R = 2$, $\theta = \pi/4$, $\phi = 0$

The contravariant basis vector $\mathbf{Z}^1$ equals $\mathbf{Z}_1$, while $\mathbf{Z}^2$ is colinear with $\mathbf{Z}_2$ and has length $1/r$.

Exercise 75. Show that $\mathbf{Z}_3$ in cylindrical coordinates is a unit vector that points in the direction of the z-axis.

Exercise 76. Show that metric tensors in polar coordinates are given by

$$Z_{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad Z^{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & 1 \end{bmatrix}; \tag{5.47}$$

Exercise 77. Show that $\mathbf{Z}^3 = \mathbf{Z}_3$.

5.11.4 Spherical Coordinates

Align a background Cartesian grid with the spherical coordinates r, θ, ϕ in the natural way: the z-axis coincide with the azimuthal axis $\theta = 0$ and let the x-axis coincide with the azimuthal axis $\phi = 0$. The position vector $\mathbf{R}$ is given by

$$\mathbf{R}(r, \theta, \phi) = r \sin\theta \cos\phi \; \mathbf{i} + r \sin\theta \sin\phi \; \mathbf{j} + r \cos\theta \; \mathbf{k}. \tag{5.48}$$

The covariant basis is obtained by partial differentiation with respect to r, θ, and ϕ:

$$\mathbf{Z}_1 = \sin\theta \cos\phi \; \mathbf{i} + \sin\theta \sin\phi \; \mathbf{j} + \cos\theta \; \mathbf{k} \tag{5.49}$$

$$\mathbf{Z}_2 = r \cos\theta \cos\phi \; \mathbf{i} + r \cos\theta \sin\phi \; \mathbf{j} - r \sin\theta \; \mathbf{k} \tag{5.50}$$

$$\mathbf{Z}_3 = -r \sin\theta \sin\phi \; \mathbf{i} + r \sin\theta \cos\phi \; \mathbf{j}. \tag{5.51}$$

The covariant basis is illustrated in Fig. 5.5.

The metric tensors are given by

$$Z_{ij} = \begin{bmatrix} 1 & & \\ & r^2 & \\ & & r^2 \sin^2 \theta \end{bmatrix}; \quad Z^{ij} = \begin{bmatrix} 1 & & \\ & r^{-2} & \\ & & r^{-2} \sin^{-2} \theta \end{bmatrix}. \tag{5.52}$$

The contravariant basis $\mathbf{Z}^i$ is

$$\mathbf{Z}^1 = \sin\theta\cos\phi\, \mathbf{i} + \sin\theta\sin\phi\, \mathbf{j} + \cos\theta\, \mathbf{k} \tag{5.53}$$

$$\mathbf{Z}^2 = r^{-1}\cos\theta\cos\phi\, \mathbf{i} + r^{-1}\cos\theta\sin\phi\, \mathbf{j} - r^{-1}\sin\theta\, \mathbf{k} \tag{5.54}$$

$$\mathbf{Z}^3 = -r^{-1}\sin^{-1}\theta\sin\phi\, \mathbf{i} + r^{-1}\sin^{-1}\theta\cos\phi\, \mathbf{j}. \tag{5.55}$$

5.12 The Christoffel Symbol Γ_{ij}^k

In affine coordinates, the covariant basis $\mathbf{Z}_i$ is the same at all points. In curvilinear coordinate systems, the basis varies from one point to another. The variation can be described by the partial derivatives $\partial \mathbf{Z}_i / \partial Z^j$. The expression $\partial \mathbf{Z}_i / \partial Z^j$ represents N^2 vectors: each of the basis elements differentiated with respect to each of the coordinates. Each of the N^2 vectors $\partial \mathbf{Z}_i / \partial Z^j$ can be decomposed with respect to the covariant basis $\mathbf{Z}_k$. The resulting N^3 coefficients form the Christoffel symbol Γ_{ij}^k:

$$\frac{\partial \mathbf{Z}_i}{\partial Z^j} = \Gamma_{ij}^k \mathbf{Z}_k. \tag{5.56}$$

This equation represents N^2 vector identities that define N^3 scalar entries of Γ_{ij}^k.

The Christoffel symbol first appeared in Elwin Bruno Christoffel's masterpiece [7]. This work was a precursor to Gregorio Ricci and Tullio Levi-Civita's paper [34] that announced tensor calculus. The Christoffel symbol plays an extraordinarily important behind-the-scenes role. Its value becomes apparent in Chap. 8 in which the covariant derivative is introduced.

The object $\partial \mathbf{Z}_i / \partial Z^j$ is the second derivative of the position vector $\mathbf{R}$

$$\frac{\partial \mathbf{Z}_i}{\partial Z^j} = \frac{\partial^2 \mathbf{R}}{\partial Z^j \partial Z^i}. \tag{5.57}$$

Therefore,

$$\frac{\partial \mathbf{Z}_i}{\partial Z^j} = \frac{\partial \mathbf{Z}_j}{\partial Z^i}, \tag{5.58}$$

Fig. 5.6 Elwin Bruno Christoffel (1829–1900), German physicist and mathematician, was a central figure in the origins of tensor calculus

which implies that the Christoffel symbol is symmetric in the lower indices:

$$\Gamma^k_{ij} = \Gamma^k_{ji}. \tag{5.59}$$

The explicit expression for the Christoffel symbol is

$$\Gamma^k_{ij} = \mathbf{Z}^k \cdot \frac{\partial \mathbf{Z}_i}{\partial Z^j}. \tag{5.60}$$

The equation follows immediately from the discussion in Sect. 5.10. Start with equation (5.56) and dot both sides with $\mathbf{Z}^l$:

$$\mathbf{Z}^l \cdot \frac{\partial \mathbf{Z}_i}{\partial Z^j} = \Gamma^k_{ij} \mathbf{Z}_k \cdot \mathbf{Z}^l. \tag{5.61}$$

By equation (5.16), $\mathbf{Z}_k \cdot \mathbf{Z}^l = \delta^l_k$, therefore

$$\mathbf{Z}^l \cdot \frac{\partial \mathbf{Z}_i}{\partial Z^j} = \Gamma^k_{ij} \delta^l_k = \Gamma^l_{ij}. \tag{5.62}$$

Renaming the index l into k yields equation (5.60).

In Euclidean spaces, equations (5.56) and (5.60) are equivalent. In Chap. 10, where we discuss differentiation on curved surfaces, we discover that equation (5.60) is a more universal approach to defining the Christoffel symbol than equation (5.56).

The Christoffel symbol also appears in the decomposition of the partial derivatives $\partial \mathbf{Z}^i / \partial Z^j$ of the *contravariant* basis $\mathbf{Z}^i$. Transform equation (5.60) by the product rule

$$\Gamma^k_{ij} = \frac{\partial \left(\mathbf{Z}^k \cdot \mathbf{Z}_i \right)}{\partial Z^j} - \frac{\partial \mathbf{Z}^k}{\partial Z^j} \cdot \mathbf{Z}_i. \tag{5.63}$$

The product $\mathbf{Z}^k \cdot \mathbf{Z}_i$ is the Kronecker symbol δ^k_j. Since δ^k_j has constant elements, its partial derivatives vanish. This yields a new expression for the Christoffel symbol

$$\Gamma^k_{ij} = -\frac{\partial \mathbf{Z}^k}{\partial Z^j} \cdot \mathbf{Z}_i, \tag{5.64}$$

from which it can be seen that $\partial \mathbf{Z}^k / \partial Z^j$ is given by

$$\frac{\partial \mathbf{Z}^k}{\partial Z^j} = -\Gamma^k_{ij} \mathbf{Z}^i. \tag{5.65}$$

Note the similarities and the differences between equations (5.56) and (5.65). In both cases, the Christoffel symbol is the key to expressing the partial derivatives of the basis elements. On the other hand, the equations differ in sign—this sign difference between covariant and contravariant objects permeates all of tensor calculus. The second important difference is the roles of the indices. In equation (5.56), the two lower indices (i and j) are free and the upper index participates in the contraction. In equation (5.65), the upper and lower indices (k and j) are free and the other lower index is contracted.

We now mention the remarkable property that the Christoffel symbol is *intrinsic*. The expression for Γ^k_{ij} in terms of the metric tensor reads

$$\Gamma^k_{ij} = \frac{1}{2} Z^{km} \left(\frac{\partial Z_{mi}}{\partial Z^j} + \frac{\partial Z_{mj}}{\partial Z^i} - \frac{\partial Z_{ij}}{\partial Z^m} \right). \tag{5.66}$$

In our approach, the Christoffel symbol was defined *extrinsically* in terms of the bases $\mathbf{Z}_i$ and $\mathbf{Z}^i$. Therefore, the fact that it is actually an intrinsic object is surprising. In Riemannian geometry, where the starting point is the metric tensor, equation (5.66) serves as the definition of the Christoffel symbol. The derivation of equation (5.66) is given in the following exercises.

Exercise 78. Show that

$$\frac{\partial Z_{ij}}{\partial Z^k} = Z_{li} \Gamma^l_{jk} + Z_{lj} \Gamma^l_{ik}. \tag{5.67}$$

The combination $Z_{li}\Gamma^l_{jk}$ can be denoted by $\Gamma_{i,jk}$. This is an example of *index juggling* introduced in the next chapter. Some texts refer to $\Gamma_{i,jk}$ as the Christoffel symbol of the first kind and Γ^i_{jk} as the Christoffel symbol of the second kind. In terms of the Christoffel symbol of the first kind, the expression for $\partial Z_{ij} / \partial Z^k$ reads

$$\frac{\partial Z_{ij}}{\partial Z^k} = \Gamma_{i,jk} + \Gamma_{j,ik}. \tag{5.68}$$

Exercise 79. By renaming indices, obtain an expression for each of the terms in parentheses of equation (5.66).

Exercise 80. Combine all expressions in the preceding exercise to derive equation (5.66).

Exercise 81. Consider a material particle moving along a curve with parametrization with respect to time given by

$$Z^i \equiv Z^i(t). \tag{5.69}$$

Suppose that the velocity of the material particle is given by $\mathbf{V}(t)$. Show that the component $V^i(t)$ of $\mathbf{V}(t)$ is given by

$$V^i(t) = \frac{dZ^i(t)}{dt}. \tag{5.70}$$

Hint: $\mathbf{V}(t) = \mathbf{R}'(t)$.

Exercise 82. Show that the component $A^i(t)$ of acceleration $\mathbf{A}(t)$ of the particle from the preceding exercise is given by

$$A^i = \frac{dV^i}{dt} + \Gamma^i_{jk} V^j V^k. \tag{5.71}$$

Exercise 83. Let $\mathbf{B}(t) = \mathbf{A}'(t)$ be the rate of change in acceleration $\mathbf{A}(t)$. Show that $B^i(t)$ is given by

$$B^i = \frac{dA^i}{dt} + \Gamma^i_{jk} A^j V^k. \tag{5.72}$$

Exercise 84. For a general vector field $\mathbf{U}(t)$ defined along the curve $Z^i \equiv Z^i(t)$, show that

$$\mathbf{U}'(t) = \left(\frac{dU^i}{dt} + V^j \Gamma^i_{jk} U^k\right) \mathbf{Z}_i. \tag{5.73}$$

Therefore, if $\mathbf{U}(t)$ is constant along the curve, we have

$$\frac{dU^i}{dt} + V^j \Gamma^i_{jk} U^k = 0. \tag{5.74}$$

5.13 The Order of Indices

With the introduction of the Christoffel symbol, we have started using the tensor notation in earnest. We must appreciate the ease with which complex operations are expressed in tensor notation. Consider the now familiar change of identities

$$\Gamma^k_{ij} = \mathbf{Z}^k \cdot \frac{\partial \mathbf{Z}_i}{\partial Z^j} = \frac{\partial \left(\mathbf{Z}^k \cdot \mathbf{Z}_i\right)}{\partial Z^j} - \frac{\partial \mathbf{Z}^k}{\partial Z^j} \cdot \mathbf{Z}_i = -\frac{\partial \mathbf{Z}^k}{\partial Z^j} \cdot \mathbf{Z}_i. \tag{5.75}$$

In three dimensions, this equation represents 27 simultaneous calculations. Yet, the calculation flows as easily as if it were a single calculation.

The Christoffel symbol has three indices, the largest number of indices we have encountered so far. The second Bianchi identity (12.15) in Chap. 12 includes an object with five free indices. It is impractical to treat these objects as multidimensional matrices with a specialized operation of multiplication. In tensor calculus, it is best to think of variants, regardless of dimension, as an indexed collection of values. In other words, a variant is an object that yields a value for every valid combination of indices.

The order of indices is important. For example, what is the value of the Christoffel symbol in spherical coordinates at the point $r = 1, \theta = \pi/2, \phi = 0$ for the index values $1, 3, 2$? This question is unclear because it is not indicated which index is meant to have which value. It is often impractical to refer to indices by name since indices can be renamed. A better way to eliminate the ambiguity is to agree on the order of the indices. For the Christoffel symbol Γ^k_{ij}, there are two competing conventions regarding the order of the indices. According to the convention adopted here, the upper index is first and the two lower indices are second and third.

For variants with indices that are all lower or all upper, the order is self-evident. Such variants include the metric tensors Z_{ij} and Z^{ij}. For some variants, where the index placements are mixed, the order does not matter. For example, the Kronecker symbol δ^i_j is zero for index values 1 and 2 regardless of which index has value 1 and which has value 2. For all other variants the order is important and must be maintained either by convention, as in the case of the Christoffel symbol, or by a special notational device such as the dot place keeper. For example, in order to indicate that the upper index is first for a variant A^i_j we can write it as $A^i_{\cdot j}$. If we want the upper index to be second, we write it as $A^{\cdot i}_j$. The dot approach is discussed further in Chap. 6 in the context of *index juggling*.

5.14 The Christoffel Symbol in Various Coordinates

5.14.1 Cartesian and Affine Coordinates

In affine coordinates, and Cartesian in particular, the Christoffel symbol vanishes at all points

$$\Gamma^i_{jk} = 0. \tag{5.76}$$

This is a differential way of saying that the covariant basis, and therefore the metric tensors, are unchanged from one point to another.

5.14.2 Cylindrical Coordinates

The nonzero Christoffel elements are given by

$$\Gamma^1_{22} = -r \tag{5.77}$$

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r}. \tag{5.78}$$

5.14.3 Spherical Coordinates

Letting $R = 1$, $\Theta = 2$, and $\Phi = 3$, the nonzero Christoffel symbol entries are given by

$$\Gamma^R_{\Theta\Theta} = -r \tag{5.79}$$

$$\Gamma^R_{\Phi\Phi} = -r\sin^2\theta \tag{5.80}$$

$$\Gamma^\Theta_{R\Theta} = \Gamma^\Theta_{\Theta R} = \frac{1}{r} \tag{5.81}$$

$$\Gamma^\Theta_{\Phi\Phi} = -\sin\theta\cos\theta \tag{5.82}$$

$$\Gamma^\Phi_{R\Phi} = \Gamma^\Phi_{\Phi R} = \frac{1}{r} \tag{5.83}$$

$$\Gamma^\Theta_{\Theta\Phi} = \Gamma^\Theta_{\Phi\Theta} = \cot\theta. \tag{5.84}$$

Exercise 85. Derive the Christoffel symbols in cylindrical and spherical coordinates by equation (5.60).

Exercise 86. Derive the Christoffel symbols in cylindrical and spherical coordinates by equation (5.66).

5.15 Summary

In this chapter, we introduced the most important differential elements in the Euclidean space: the position vector $\mathbf{R}$, the covariant and contravariant bases $\mathbf{Z}_i$ and $\mathbf{Z}^i$, the covariant and contravariant metric tensors Z_{ij} and Z^{ij}, and the volume element $\sqrt{Z}$. These objects are collectively known as the *metrics*. There are two more objects that are considered metrics: the volume element $\sqrt{Z}$ and the Levi-Civita symbols ε^{ijk} are discussed. Both are introduced in Chap. 9. The volume element plays a particularly crucial role in the subject of integrals (Chap. 14).

In the next chapter we study how variants transform under a change of coordinates. We learn that the metrics and a number of other objects are tensors (that is objects that transform according to a special rule) and discuss the important implications of the tensor property.

Chapter 6
The Tensor Property

6.1 Preview

In this chapter we get down to the business of tensors. We introduce the concept of a variant and define tensors as variants that transform from one coordinate system to another by a very special rule. We define covariance and contravariance and demonstrate how these opposite modes of transformation under a change of coordinates lead to the main goal of tensor calculus: invariance.

6.2 Variants

A *variant* is an object that can be constructed by a similar rule in various coordinate systems. In the preceding chapter, we encountered a number of variants. For example, the covariant basis $\mathbf{Z}_i$ is a variant because it is obtained in any coordinate system by the same rule—partial differentiation of the position vector $\mathbf{R}$. The metric tensor Z_{ij} is also a variant because it is formed by pairwise dot products of the elements of the covariant basis. Similarly, every object introduced in the preceding chapter is a variant, including the Christoffel symbol Γ^k_{ij}.

The Jacobian $J^{i'}_i$ is not a variant. It expresses a relationship between two coordinate systems and therefore cannot be constructed given one coordinate system. Similarly, second-order derivatives $J^{i'}_{ij}$ are not variants. If one of the coordinate systems is fixed and the other one is allowed to change, then $J^{i'}_i$ and $J^{i'}_{ij}$ can be considered variants with respect to the changing coordinate system.

Naturally, when the same rule is applied in different coordinate systems, the results differ. This is the motivation behind calling these objects *variants*. Among variants, there stands out a very special class of objects that transform according to a special rule. These objects are called *tensors*, and we now turn to the study of these crucially important objects. We note in advance that all metrics $\mathbf{Z}_i$, $\mathbf{Z}^i$, Z_{ij} and Z^{ij} are tensors.

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_6, © Springer Science+Business Media New York 2013

6.3 Definitions and Essential Ideas

6.3.1 Tensors of Order One

A variant T_i is called a *covariant tensor* if its values T_i and $T_{i'}$ in the coordinate systems Z^i and $Z^{i'}$ are related by

$$T_{i'} = T_i J^i_{i'}, \tag{6.1}$$

where $J^i_{i'}$ is the Jacobian defined in equation (4.70). A variant T^i is called a *contravariant tensor* if T^i and $T^{i'}$ are related by the Jacobian $J^{i'}_i$ defined in (4.71)

$$T^{i'} = T^i J^{i'}_i. \tag{6.2}$$

The term covariant means *transforms in the same way as the basis* $\mathbf{Z}_i$ and the term *contravariant* means transforms in the *opposite* way. Recall that the Jacobians $J^i_{i'}$ and $J^{i'}_i$ are the matrix inverses of each other. Covariant tensors are denoted by lower indices and contravariant tensors are denoted by upper indices. In addition to numerous other advantages, the placement of the index makes it easy to remember whether $J^{i'}_i$ or $J^i_{i'}$ is to be used.

The fundamental importance of the tensor property is rooted in the fact that $J^i_{i'}$ and $J^{i'}_i$ are the matrix inverses of each other. This relationship is expressed in equations (4.74) and (4.76). Let us repeat these identities here since we will make frequent use of them in this chapter:

$$J^i_{i'} J^{i'}_j = \delta^i_j \tag{4.74}$$

$$J^{i'}_i J^i_{j'} = \delta^{i'}_{j'}. \tag{4.76}$$

6.3.2 Tensors Are the Key to Invariance

To appreciate the importance of the tensor property, let us prove a special case of the *contraction theorem*. It is the fundamental theorem of tensor calculus and it gives us a recipe for producing geometrically meaningful invariants. Suppose that S_i is a covariant tensor and T^i is a contravariant tensor. Then the contraction

$$U = S_i T^i \tag{6.3}$$

is *invariant*. That is, U evaluates to the same value in all coordinate systems!

To demonstrate this property, consider the quantity U' evaluated in the primed coordinates according to the same contraction as U in the unprimed coordinates:

$$U' = S_{i'} T^{i'}. \tag{6.4}$$

Relate the values $S_{i'}$ and $T^{i'}$ to their values S_i and T^i in the unprimed coordinates:

$$U' = S_i J^i_{i'} T^k J^{i'}_k . \tag{6.5}$$

We introduced a new index k in expressing $T^{i'}$ since i was used for $S_{i'}$. Note, that because of the relationship (4.74) between the Jacobians $J^i_{i'}$ and $J^{i'}_k$, U' is given by

$$U' = S_i T^k \delta^i_k , \tag{6.6}$$

or, since $T^k \delta^i_k = T^i$, we find

$$U' = S_i T^i . \tag{6.7}$$

In other words,

$$U' = U \tag{6.8}$$

and we have shown that the variant U evaluates to the same value in all coordinate systems. The keys to the *invariance* are the tensor property of S_i and T^i and the inverse relationship between the Jacobians.

6.3.3 *The Tensor Property of* $\mathbf{Z}_i$

We now show that the covariant basis $\mathbf{Z}_i$ is indeed a covariant tensor and so the term *covariant* is justified. The demonstration of the tensor property usually proceeds as follows. We start with the definition of the variant in the primed coordinate system and then relate its values in the primed coordinates to its values in the unprimed coordinates. If the relationship is of type (6.1) or (6.2) then the variant is a tensor of the corresponding kind.

Let us first look at the one-line proof of the tensor property $\mathbf{Z}_i$ and then discuss each step:

$$\mathbf{Z}_{i'} = \frac{\partial \mathbf{R}\left(Z'\right)}{\partial Z^{i'}} = \frac{\partial \mathbf{R}\left(Z\left(Z'\right)\right)}{\partial Z^{i'}} = \frac{\partial \mathbf{R}}{\partial Z^i} \frac{\partial Z^i}{\partial Z^{i'}} = \mathbf{Z}_i J^i_{i'} . \tag{6.9}$$

Now, one step at a time. The covariant basis $\mathbf{Z}_{i'}$ in the primed coordinate system is given by the partial derivative of the position vector $\mathbf{R}$ with respect to the primed coordinate $Z^{i'}$:

$$\mathbf{Z}_{i'} = \frac{\partial \mathbf{R}\left(Z'\right)}{\partial Z^{i'}} . \tag{6.10}$$

The function $\mathbf{R}(Z')$ gives the position vector $\mathbf{R}$ as a function of the coordinates $Z^{i'}$. The following is the critical step from which the ultimate relationship is derived. The function $\mathbf{R}(Z')$ can also be constructed by substituting $Z^i(Z')$—the expression for the unprimed coordinates in terms of the primed coordinates—into $\mathbf{R}(Z)$, that is the function that gives the position vector $\mathbf{R}$ as a function of the unprimed coordinates Z^i:

$$\mathbf{R}\left(Z'\right) = \mathbf{R}\left(Z\left(Z'\right)\right). \tag{6.11}$$

The next step is differentiation by the chain rule, the tensor form of which you mastered in Chap. 4:

$$\frac{\partial \mathbf{R}\left(Z\left(Z'\right)\right)}{\partial Z^{i'}} = \frac{\partial \mathbf{R}}{\partial Z^i}\frac{\partial Z^i}{\partial Z^{i'}}. \tag{6.12}$$

Finally, we recognize that $\partial \mathbf{R}/\partial Z^i$ is $\mathbf{Z}_i$ and $\partial Z^i/\partial Z^{i'}$ is $J^i_{i'}$. Summarizing

$$\mathbf{Z}_{i'} = \mathbf{Z}_i J^i_{i'}, \tag{6.13}$$

which shows that $\mathbf{Z}_i$ is indeed a covariant tensor.

6.3.4 *The Reverse Tensor Relationship*

Suppose that T_i is a covariant tensor, that is, $T_{i'} = T_i J^i_{i'}$. What is T_i in terms of $T_{i'}$? For a moment, consider this question in the language of linear algebra: if $T' = JT$, then what is T in terms of T'? It is obtained by left-multiplying both sides of the identity by J^{-1}, yielding $T = J^{-1}T'$. In tensor algebra terms, the same effect is achieved by multiplying both sides of the identity $T_{i'} = T_i J^i_{i'}$ by $J^{i'}_j$:

$$T_{i'} J^{i'}_j = T_i J^i_{i'} J^{i'}_j. \tag{6.14}$$

Since $J^i_{i'} J^{i'}_j = \delta^i_j$ by equation (4.74), we have

$$T_{i'} J^{i'}_j = T_i \delta^i_j = T_j. \tag{6.15}$$

Finally, renaming the index j to i, we find

$$T_i = T_{i'} J^{i'}_i. \tag{6.16}$$

Similarly, if T^i is the contravariant tensor, then the reverse of $T^{i'} = T^i J^{i'}_i$ reads

$$T^i = T^{i'} J^i_{i'}. \tag{6.17}$$

Equations (6.16), (6.17), in conjunction with (6.1), (6.2), offer greater flexibility in demonstrating the tensor property. On some occasions, it is easier to start with the unprimed components T_i or T^i and relate them to the primed components $T_{i'}$ and $T^{i'}$. On other occasions, its more natural to go the other way. Equations (6.1), (6.2), (6.16) and (6.17) help move effortlessly in either direction.

There is a noteworthy elegance in equations (6.1), (6.2), (6.16), (6.17) which is found in all tensor relationships: the placement of the indices leads to the proper choice of objects. For example, you may not find it easy to remember whether $T^{i'}$ and T^i are related by $J^i_{i'}$ or its inverse $J^{i'}_i$. The matching of indices naturally suggests the correct equations (6.2) and (6.17). This property of the tensor notation is one of its greatest assets.

6.3.5 Tensor Property of Vector Components

We next show that components V^i of a vector $\mathbf{V}$ with respect to the covariant basis $\mathbf{Z}_i$ form a contravariant tensor. As always, begin by writing the definition of $V^{i'}$ in the primed coordinates:

$$\mathbf{V} = V^{i'} \mathbf{Z}_{i'}. \tag{6.18}$$

By the tensor property of $\mathbf{Z}_i$, we have

$$\mathbf{V} = V^{i'} J^i_{i'} \mathbf{Z}_i. \tag{6.19}$$

In the unprimed coordinates, the vector $\mathbf{V}$ is given by

$$\mathbf{V} = V^i \mathbf{Z}_i. \tag{6.20}$$

Since decomposition with respect to a basis is unique, the last two equations yield the following relationship between V^i and $V^{i'}$:

$$V^i = V^{i'} J^i_{i'}. \tag{6.21}$$

This proves the contravariant property of vector components V^i.

We have previously shown that if S_i and T^i are a covariant tensor and a contravariant tensor, then the contraction $S_i T^i$ is an invariant. In the foregoing proof, we have essentially demonstrated a converse: If $S_i T^i$ is invariant and one of the terms in the product is a tensor, then the other term is also a tensor. This is a special case of a more general *quotient theorem*.

6.3.6 The Tensor Property of Z^i

The tensor property of the contravariant basis $\mathbf{Z}^i$ can be proved by similar reasoning. Recall the following relationship between the covariant and contravariant bases:

$$\mathbf{Z}^i \cdot \mathbf{Z}_j = \delta^i_j. \tag{5.16}$$

Multiply both sides by $J^{i'}_i J^j_{j'}$

$$\mathbf{Z}^i J^{i'}_i \cdot \mathbf{Z}_j J^j_{j'} = \delta^i_j J^{i'}_i J^j_{j'} \tag{6.22}$$

As we stressed in Chap. 4, the order of the multiplicative terms is immaterial. Furthermore, the placement of the dot $\cdot$ in the dot product is arbitrary, and the result is not affected by the order of operations.

By the tensor property of $\mathbf{Z}_j$, the contraction $\mathbf{Z}_j J^j_{j'}$ is $\mathbf{Z}_{j'}$. Furthermore $\delta^i_j J^{i'}_i J^j_{j'}$ is $\delta^{i'}_{j'}$ (why?), yielding:

$$\mathbf{Z}^i J^{i'}_i \cdot \mathbf{Z}_{j'} = \delta^{i'}_{j'}. \tag{6.23}$$

This identity tells us that the set of three vectors $\mathbf{Z}^i J^{i'}_i$ is mutually orthogonal to $\mathbf{Z}_{j'}$. Therefore, on linear algebra grounds, the vector $\mathbf{Z}^i J^{i'}_i$ must be $\mathbf{Z}^{i'}$:

$$\mathbf{Z}^{i'} = \mathbf{Z}^i J^{i'}_i. \tag{6.24}$$

This proves the contravariant property of $\mathbf{Z}^i$.

6.4 Tensors of Higher Order

The definition of tensor generalizes to an arbitrary number of indices. For example, a variant T_{ij} is a doubly covariant tensor if its primed values $T_{i'j'}$ relate to the unprimed T_{ij} by the identity

$$T_{i'j'} = T_{ij} J^i_{i'} J^j_{j'}. \tag{6.25}$$

Similarly, T^{ij} is a doubly contravariant tensor if $T^{i'j'}$ and T^{ij} are related by

$$T^{i'j'} = T^{ij} J^{i'}_i J^{j'}_j. \tag{6.26}$$

Finally, T^i_j is tensor of one contravariant order and one covariant order if

$$T^{i'}_{j'} = T^i_j J^{i'}_i J^j_{j'}. \tag{6.27}$$

When stating that a particular variant is a tensor, one need not mention the *flavor* (covariant or contravariant) of the index since it is indicated by the placement of the index. Therefore, T_{ij}, T^{ij}, and T^i_j are all described simply as second-order tensors.

The tensor definition is easily extended to tensors of any order by operating on each index with the proper Jacobian. For instance R_{ijkl} is a fourth-order tensor if

$$R_{i'j'k'l'} = R_{ijkl} J^i_{i'} J^j_{j'} J^k_{k'} J^l_{l'}. \tag{6.28}$$

6.4.1 The Tensor Property of Z_{ij} and Z^{ij}

We now show that Z_{ij} and Z^{ij} are tensors. For the covariant metric tensor Z_{ij}, its components in the primed coordinates are, by definition, the dot product of the primed covariant basis vectors:

$$Z_{i'j'} = \mathbf{Z}_{i'} \cdot \mathbf{Z}_{j'}. \tag{6.29}$$

Utilizing the covariant property of the covariant basis, we have

$$Z_{i'j'} = \mathbf{Z}_i J^i_{i'} \cdot \mathbf{Z}_j J^j_{j'}. \tag{6.30}$$

Recognizing that $\mathbf{Z}_i \cdot \mathbf{Z}_j = Z_{ij}$ we arrive at the identity

$$Z_{i'j'} = Z_{ij} J^i_{i'} J^j_{j'}, \tag{6.31}$$

which proves the tensor property of Z_{ij}.

The tensor property of Z^{ij} can be established in two ways. Since we have already established the tensor property of $\mathbf{Z}^i$, the tensor property of Z^{ij} follows easily by analogy with Z_{ij}:

$$Z^{i'j'} = \mathbf{Z}^{i'} \cdot \mathbf{Z}^{j'} = \mathbf{Z}^i J^{i'}_i \cdot \mathbf{Z}^j J^{j'}_j = Z^{ij} J^{i'}_i J^{j'}_j. \tag{6.32}$$

On the other hand, the tensor property of Z^{ij} can be established by a quotient argument similar to the argument that showed the tensor property of $\mathbf{Z}^i$.

Exercise 87. Prove the tensor property of Z^{ij} from the identity $Z^{ij} Z_{jk} = \delta^i_k$

6.4.2 The Tensor Property of δ^i_j

The Kronecker symbol δ^i_j may seem like an invariant: it has the same values in all coordinates. However, the concept of invariant only applies to variants of order zero. The Kronecker symbol as a variant of order two. But is it a tensor? The answer is yes. Below, we prove that a product (or a dot product) of two tensors is a tensor.

This implies that δ^i_j is a tensor because it is the dot product of two tensors: $\delta^i_j = \mathbf{Z}^i \cdot \mathbf{Z}_j$. On the other hand, we can show the tensor property of the Kronecker symbol directly. Multiply δ^i_j by the combination $J^{i'}_i J^j_{j'}$. The result is $\delta^{i'}_{j'}$

$$\delta^i_j J^{i'}_i J^j_{j'} = J^{i'}_j J^j_{j'} = \delta^{i'}_{j'}, \tag{6.33}$$

which shows that the values of the Kronecker symbol in different coordinate systems relate by the tensor rule. Therefore, the Kronecker symbol is a tensor.

6.5 Exercises

Exercise 88. Show that, for a scalar field F, the collection of partial derivatives $\partial F/\partial Z^i$ is a covariant tensor.

Exercise 89. Given a scalar field F, show that the collection of second-order partial derivatives $\partial^2 F/\partial Z^i \partial Z^j$ is not a tensor. More generally, generally show that, for a covariant tensor field T_i, the variant $\partial T_i/\partial Z^j$ is not a tensor.

Exercise 90. Show that the skew-symmetric part S_{ij}

$$S_{ij} = \frac{\partial T_i}{\partial Z^j} - \frac{\partial T_j}{\partial Z^i} \tag{6.34}$$

of the variant $\partial T_i/\partial Z^j$ is a tensor.

Exercise 91. Similarly, for a contravariant tensor T^i, derive the transformation rule for $\partial T^i/\partial Z^j$, and show that it is not a tensor.

Exercise 92. Derive the transformation rule for the Christoffel symbol Γ^k_{ij} on the basis of equation (5.60) and show that is not a tensor.

Exercise 93. For a tensor T_{ij}, derive the transformation rule for $\partial T_{ij}/\partial Z^k$ and show that it is not a tensor. Apply the obtained relationship to the metric tensor Z_{ij} and use the result to obtain the transformation rule for the Christoffel symbol Γ^k_{ij} on the basis of equation (5.66).

Exercise 94. Show that if a tensor vanishes in one coordinate system, then it vanishes in all coordinate systems.

The tensor property may sometimes be considered with respect to a subclass of transformations. That is, a variant may not be a tensor in the general sense, but still a tensor with respect to a narrower class of coordinate transformations, such as linear transformations characterized by Jacobians with positive determinants.

Exercise 95. Which of the objects considered in the preceding exercises are tensors with respect to linear transformations

$$Z^{i'} = A^{i'}_{i} Z^{i} + b^{i'}, \tag{6.35}$$

without being tensors in the general sense?

6.6 The Fundamental Properties of Tensors

6.6.1 Sum of Tensors

The sum of two tensors is a tensor. It is implied that the two tensors being added have identical index signatures and the result is a tensor with the same signature. We prove this property by considering a representative indicial signature. Let A^{i}_{jk} be the sum of B^{i}_{jk} and C^{i}_{jk}:

$$A^{i}_{jk} = B^{i}_{jk} + C^{i}_{jk}. \tag{6.36}$$

Then

$$A^{i'}_{j'k'} = B^{i'}_{j'k'} + C^{i'}_{j'k'} = B^{i}_{jk} J^{i'}_{i} J^{j}_{j'} J^{k}_{k'} + C^{i}_{jk} J^{i'}_{i} J^{j}_{j'} J^{k}_{k'} = A^{i}_{jk} J^{i'}_{i} J^{j}_{j'} J^{k}_{k'}. \tag{6.37}$$

Summarizing,

$$A^{i'}_{j'k'} = A^{i}_{jk} J^{i'}_{i} J^{j}_{j'} J^{k}_{k'}, \tag{6.38}$$

which proves that A^{i}_{jk} is a tensor.

Exercise 96. Show that a linear combination of tensors is a tensor.

6.6.2 Product of Tensors

Any two variants can be multiplied together to form a new variant. The result is called the *tensor product* and it has as many indices as the multiplicative terms combined. For example,

$$A^{i}_{jkl} = B^{i}_{j} C_{kl}. \tag{6.39}$$

The tensor product of tensors is a tensor. The proof for representative indicial signatures is completely straightforward:

$$A^{i'}_{j'k'l'} = B^{i'}_{j'} C_{k'l'} = B^{i}_{j} C_{kl} J^{i'}_{i} J^{j}_{j'} J^{k}_{k'} J^{l}_{l'} = A^{i}_{jkl} J^{i'}_{i} J^{j}_{j'} J^{k}_{k'} J^{l}_{l'}. \tag{6.40}$$

Similarly, the dot product of two vector-valued tensors is a tensor, and the product of a scalar-valued tensor with a vector-valued tensor is a tensor.

6.6.3 The Contraction Theorem

We have already discussed a special case of the contraction theorem and learned that it is the key to forming invariants. Here we give a general form of the theorem and discuss its important implications. The theorem states: contraction of a tensor is a tensor. Of course, a contraction is only valid when it is over one covariant and one contravariant index.

The proof once again proceeds by example. The key to the proof is, of course, the inverse relationships between the Jacobians. Suppose that S^{i}_{jk} is a tensor of order three. We will show that $T_k = S^{i}_{ik}$ is a tensor. We first express the fact that S^{i}_{jk} is a tensor

$$S^{i'}_{j'k'} = S^{i}_{jk} J^{i'}_{i} J^{j}_{j'} J^{k}_{k'}. \tag{6.41}$$

Contracting both sides on i' and j' yields

$$S^{i'}_{i'k'} = S^{i}_{jk} J^{i'}_{i} J^{j}_{i'} J^{k}_{k'}. \tag{6.42}$$

Since $J^{i'}_{i} J^{j}_{i'} = \delta^{j}_{i}$, we find

$$S^{i'}_{i'k'} = S^{i}_{jk} \delta^{j}_{i} J^{k}_{k'}, \tag{6.43}$$

or

$$S^{i'}_{i'k'} = S^{i}_{ik} J^{k}_{k'}. \tag{6.44}$$

In other words,

$$T_{k'} = T_k J^{k}_{k'} \tag{6.45}$$

and the desired relationship is obtained. This demonstration for a particular *index signature* should convince you that the theorem holds for arbitrary tensors.

6.6.4 *The Important Implications of the Contraction Theorem*

We have shown that a variant built up from tensors is a tensor. *Building up an expression includes forming linear combinations, products and contractions.* Importantly, in this scheme, invariants can be treated as *tensors of order zero*. For example, if U is an invariant, then by definition

$$U' = U, \tag{6.46}$$

which is consistent with the definition of a tensor with zero indices.

Therefore, by the contraction theorem, when all indices are contracted away, the result is an invariant. For example, if S^i_j is a tensor, then S^i_i is an invariant. If S^{ij}_{kl} is a tensor, then S^{ij}_{ij} is an invariant and S^{ij}_{ji} is an invariant as well, but a different one. If S_i is a tensor and T^{ij}_k is a tensor, then $S_i T^{ij}_j$ is an invariant by combinations of the product and contraction theorems. If S^i_j is a tensor, then $\left(\delta^k_i \delta^l_j - \delta^l_i \delta^k_j\right) S^i_k S^j_l$ is an invariant by combination of the sum, product, and contraction theorems.

And so, we developed a strategy for forming invariants: stick tensors and tensor operations—linear combinations, products, and contractions. This is the central idea of tensor calculus and an important moment in this book. Earlier we argued that while coordinate systems are an irreplaceable tool for solving problem, they often lead to loss of geometric insight and unmanageable calculations. We have just found a remedy to many of the coordinate ills! The tensor framework provides an approach in which calculations can take full advantage of coordinate systems while maintaining the geometric insight.

A tensor is a beautiful object. It is not an invariant and, in a particular coordinate system, it exhibits some features that reflect its geometric origins and some features that are strictly artifacts of the coordinate system. Nevertheless, the artifacts of the coordinate system are present in a systematic tensor fashion and are removed by contraction, leaving us with a geometric object. This is what makes tensors meaningful.

Many a poetic analogy can be made for tensors. One is that of a photographic negative. A negative is not a photograph but it has all the information that is needed to make one. The equipment and the materials required to make a photograph represent the contraction in the analogy. Importantly, while a negative is not a photograph, it is actually better than a photograph for a number of photograph-making tasks. For example, a negative can be used to make a photograph of many reasonable sizes without loss of quality. A negative is not meaningful by itself (neither is a tensor), but in combination with proper equipment (a tensor needs contraction) it can produce a photograph (tensors lead to invariants).

6.7 Exercises

Exercise 97. Suppose that S_i and T^{jk} are tensors. Prove that $S_i T^{ij}$ is a tensor.

Exercise 98. Since δ^i_j is a tensor, δ^i_i must be an invariant. What is the value of δ^i_i and what is its geometric meaning?

Exercise 99. Suppose that $\mathbf{V}_{ij}$ is a tensor with vector elements. Show that V^k_{ij}, the components $\mathbf{V}_{ij}$ with respect to the covariant basis $\mathbf{Z}_i$, is a tensor.

Exercise 100. Give a variant arbitrary values in some coordinate system and let it transform by the proper tensor rules. Show that this construction produces a tensor. More specifically, single out a particular coordinate system $Z^{\bar{\imath}}$ and choose an arbitrary collection of values for $T^{\bar{\imath}\bar{j}}_{\bar{k}}$. Then, in coordinates Z^i, define T^{ij}_k by

$$T^{ij}_k = T^{\bar{\imath}\bar{j}}_{\bar{k}} J^i_{\bar{\imath}} J^j_{\bar{j}} J^{\bar{k}}_k. \tag{6.47}$$

Show that T^{ij}_k is a tensor.

6.8 The Gradient Revisited and Fixed

We now have all the tools to address the issue of the gradient that was raised earlier. Recall that the gradient ∇F of a scalar field F has a perfectly clear invariant definition: it is a vector that points in the direction of the greatest increase of F and its magnitude equals the directional derivative of F in that direction.

The problem arose when we attempted to find an analytical expression for the gradient that yielded the correct result in all coordinate systems. The candidate expression was the combination

$$\nabla F = \sum_i \frac{\partial F}{\partial Z^i} \mathbf{Z}_i. \tag{6.48}$$

Note that in this expression we must use the summation sign because both indices are lower and the summation convention does not apply. We learned that this expression gives the right answer only in Cartesian coordinates. The tensor perspective offers both the explanation of why expression (6.48) fails and a fix.

Equation (6.48) has several elements, of which all but one are proper from the tensor points of view. The collection of partial derivatives $\partial F / \partial Z^i$ is a tensor. The covariant tensor $\mathbf{Z}_i$ is a tensor. Their product $\left(\partial F / \partial Z^j\right) \mathbf{Z}_i$ is a tensor of order two. However, the contraction is invalid because both tensors are covariant. The result, therefore, is not guaranteed to be an invariant. In fact, it is virtually guaranteed not to.

The fix dictates itself. In order to produce an invariant we must mix covariant and contravariant tensors. The correct analytical expression for the gradient is

$$\nabla F = \frac{\partial F}{\partial Z^i} \mathbf{Z}^i. \tag{6.49}$$

By the product and contraction theorems, this expression is an invariant. Furthermore, in Cartesian coordinates it reduces to

$$\nabla F = \frac{\partial F}{\partial x}\mathbf{e}_1 + \frac{\partial F}{\partial y}\mathbf{e}_2 + \frac{\partial F}{\partial z}\mathbf{e}_3, \tag{6.50}$$

which gives the correct vector. Since the combination in (6.49) gives the same answer in all coordinate systems and the correct answer in Cartesian coordinates, it must give the correct answer in all coordinates!

This way of reasoning, which resorts to special coordinate systems, is frequently used in tensor calculus. It is often preferable and more geometrically insightful to furnish a proof that does not require particular coordinate systems.

Exercise 101. Explain why the Laplacian ΔF of a scalar function F cannot be defined as

$$\Delta F = Z^{ij}\frac{\partial^2 F}{\partial Z^i \partial Z^j}? \tag{6.51}$$

Hint: Is $\partial^2 F/\partial Z^i \partial Z^j$ a tensor?

6.9 The Directional Derivative Identity

Now that we have a solid analytical definition of the gradient, we are able to give the proper analytical derivation of equation (2.11)

$$\frac{dF}{dl} = \nabla F \cdot \mathbf{L}. \tag{6.52}$$

Suppose l is a straight line passing through the point P in the direction of the unit vector $\mathbf{L}$. Suppose that l is parametrized by its arc length s:

$$Z^i = Z^i(s). \tag{6.53}$$

Then, along l, the function $F(s)$ is given by the composition

$$F(s) = F(Z(s)). \tag{6.54}$$

Differentiating this identity by the chain rule yields

$$\frac{dF}{ds} = \frac{\partial F}{\partial Z^i}\frac{dZ^i}{ds}. \tag{6.55}$$

The three elements of this equation can be identified with the three elements in equation (2.11). dF/ds is by definition the directional derivative dF/dl, $\partial F/\partial Z^i$ is the gradient $\nabla_i F$, and, finally, dZ^i/ds represents the unit tangent $\mathbf{L}$. Therefore, equation (2.11) is justified.

6.10 Index Juggling

We have already encountered index juggling in Chap. 5 where we introduced the contravariant basis $\mathbf{Z}^i$ as $Z^{ij}\mathbf{Z}_j$. This is an example of *raising the index*. In general, the *raising of the index* is the contraction of a variant with the contravariant metric tensor:

$$T^j = T_i Z^{ij}. \tag{6.56}$$

The variant T^j is closely related to the variant T_i. In fact, although the elements of T^j and T_i are different, they can be thought of as different manifestations of the same object.

Contraction with the *covariant* metric tensor effects the *lowering of the index*. Not surprisingly, the two operations are the inverses of each other, so no information is lost or created by the operation. It is simply a notational convenience, but as far as notation conveniences go, it is on par with the Einstein summation convention in terms of effectiveness. The operations of lowering and raising indices are referred to collectively as *index juggling*.

Any covariant index can be raised and any contravariant index can be lowered. Recall that in every tensor, the order of the indices must be clearly specified. When all the indices are either covariant, as in T_{ijk}, or contravariant, as in T^{ijk}, the natural order of the indices is easy to recognize: for T_{ijk} and T^{ijk}, i is the first index, j is the second, and k is the third.

When the flavors of the indices are mixed, as in T^i_j, the order of the indices is more arbitrary. For example, we may agree that in T^i_j, the upper index is first and the lower index is second. However, once the order is established, it must be maintained in subsequent computations. If indices can be juggled, a notational device is need to maintain the order of the indices. To illustrate the potential problem, consider the situations in which the upper index of T^i_j is lowered and the lower is raised:

$$T^k_l = T^i_j Z_{il} Z^{jk}. \tag{6.57}$$

Then what is the entry T^1_2? Well, it is unclear whether T^1_2 is referring to the original variant or the one with the indices juggled.

The solution is to allocate a slot for each index and use the dot $\cdot$ to spread out the indices. The original variant can be denoted by $T^i_{\cdot j}$ clearly indicating that the upper index is first and the lower is second. The upper part of the second slot occupied by j could also be marked with a dot $T^{i\cdot}_{\cdot j}$, but the second dot is not necessary. If the first index is lowered, dots are not necessary since both indices are lower:

$$T_{kj} = T^i_{\cdot j} Z_{ik}. \tag{6.58}$$

Subsequently, if the second index is raised, the dot reappears in the first slot:

$$T_{k}^{\cdot l} = T_{kj} Z^{jl}. \tag{6.59}$$

With this notation it is clear that $T_{\cdot 2}^{1}$ refers to the original form of the variant, T_{12} refers to the intermediate form, while $T_{1}^{\cdot 2}$ refers to the juggled form.

There are several types of variants for which tracking the order of indices is handled by different means. First, are symmetric variants. If

$$T_{ij} = T_{ji} \tag{6.60}$$

then

$$T_{\cdot j}^{i} = T_{j}^{\cdot i}, \tag{6.61}$$

and it can therefore be written as T_{j}^{i} without confusion. Note, as discussed in the following section, the Kronecker symbol δ_{j}^{i} with a lowered index is Z_{ij}. Thus, $\delta_{\cdot j}^{i} = \delta_{j}^{\cdot i}$ and it can be denoted by δ_{j}^{i} with no ambiguity.

Finally, there are a number of special variants for which the order of indices is defined by convention example, for the Christoffel symbol Γ_{jk}^{i}, introduced in Chap. 8, we adopt the convention that the upper index is first. In practice, there does not arise a situation where the lower indices need to be raised. On the other hand, the first Christoffel index is commonly juggled, and the Christoffel index with the lowered first index is written as $\Gamma_{i,jk}$. Note that even the comma is unnecessary. However, some texts use the convention where the upper index is actually last. The comma therefore is meant to signal a casual reader of this book as to our convention. If the situation ever arises where second and third indices are raised, the $\cdot$ notation would be advisable.

Exercise 102. Consider a coordinate system where the metric tensor at point A is

$$Z_{ij} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \tag{6.62}$$

Suppose that $T_{\cdot jkl}^{i}$ is a object whose only nonzero value is

$$T_{\cdot 212}^{1} = \pi. \tag{6.63}$$

Determine T_{1212} and T_{2212}

Exercise 103. Show that the operations of raising and lowering indices are the inverses of each other. That is, if

$$S_{i} = S^{j} Z_{jk}, \tag{6.64}$$

then

$$S^i = S_j Z^{jk}. \tag{6.65}$$

Exercise 104. Show that the flavors of indices can be swapped in a contraction:

$$S_i T^i = S^i T_i. \tag{6.66}$$

Exercise 105. Show that in a tensor identity, a live covariant (contravariant) index can be raised (lowered) on both sides. That is, if, for instance,

$$S^i = T^{jki} U_{jk}, \tag{6.67}$$

then

$$S_i = T^{jk}_{\cdot\cdot i} U_{jk}. \tag{6.68}$$

Exercise 106. Show that for symmetric tensors T_{ij} $(T_{ij} = T_{ji})$, we have

$$T^i_{\cdot j} = T^{\cdot j}_i. \tag{6.69}$$

Therefore

$$\delta^i_{\cdot j} = \delta^{\cdot i}_j, \tag{6.70}$$

as we are about to show.

6.11 The Equivalence of δ^i_j and Z_{ij}

Objects related by index juggling are considered equivalent. From this point of view, the Kronecker symbol and the metric tensor are equivalent, since δ^i_j with the contravariant index is the covariant metric tensor:

$$\delta^i_j Z_{ik} = Z_{jk}. \tag{6.71}$$

Therefore, the object that could be denoted by δ_{jk} is actually Z_{jk}. Similarly, what is Z_{ij} with the index raised? It is the Kronecker symbol, since

$$Z_{ij} Z^{jk} = \delta^k_i. \tag{6.72}$$

Finally, raising the covariant index in δ^i_j yields the contravariant metric tensor:

$$\delta^i_j Z^{jk} = Z^{ik}. \tag{6.73}$$

Thus, Z_{ij}, δ^i_j, and Z^{ij} represent one and the same family of tensors related by index juggling. In fact, many texts denote Z_{ij} by δ_{ij} and Z^{ij} by δ^{ij}. Furthermore, some texts denote the Kronecker symbol by Z^i_j. Since the Kronecker delta symbol is a form of the metric tensor, each of these choices are perfectly valid. On the other hand, will continue to denote the Kronecker delta by δ^i_j and covariant and contravariant metric tensors by Z_{ij} and Z^{ij}.

6.12 The Effect of Index Juggling on the Tensor Notation

Index juggling leads to further compacting of the tensor notation. As a result of index juggling, the metric tensor—perhaps the central object in tensor calculus—almost never appears in tensor relationships explicitly. Instead, the contraction of the metric tensor with other variants is reflected in the location of the index. Consider, for instance, the dot product of two vectors $\mathbf{U} = U^i \mathbf{Z}_i$ and $\mathbf{V} = V^j Z_j$:

$$\mathbf{U} \cdot \mathbf{V} = U^i \mathbf{Z}_i \cdot V^j \mathbf{Z}_j = Z_{ij} U^i V^j. \tag{6.74}$$

With index juggling, $\mathbf{U} \cdot \mathbf{V}$ can be written as

$$\mathbf{U} \cdot \mathbf{V} = U_i V^i, \tag{6.75}$$

or

$$\mathbf{U} \cdot \mathbf{V} = U^i V_i. \tag{6.76}$$

Equations (6.75) and (6.76) are deceptively simple and make the dot produce appear computationally cheaper than it really is. For example, in three dimensions, equation (6.75) reads

$$\mathbf{U} \cdot \mathbf{V} = U_1 V^1 + U_2 V^2 + U_3 V^3, \tag{6.77}$$

and appears to be three multiplications and two additions—that is, as cheap as if the basis was Cartesian. Of course, that is not the case, because a contraction with the metric tensor is required in order to obtain the covariant coordinates U_i out of the contravariant coordinates U^i. Therefore, as we would expect, index juggling does not lead to more efficient computational algorithms. It does, however, lead to more attractive and compact notation that hides the non-Cartesian nature of the local basis.

6.13 Summary

This chapter introduced the paramount tensor property. The tensor property leads to the key recipe for constructing invariant expressions: stick to tensor expressions and contract away all indices. We proved that some the essential objects in Euclidean spaces are tensors. These include $\mathbf{Z}_i$, $\mathbf{Z}^i$, Z_{ij} and Z^{ij}. On the other hand, partial derivatives of tensors result in variants that are not tensors. Therefore, while we were able to give an invariant formulation for the *gradient* in equation (6.49), we were unable to give do the same for the Laplacian since the collection of second derivatives $\partial^2 F/\partial Z^i \partial Z^j$ is not a tensor. This problem is fixed by introducing the *covariant derivative*—a new operator that produces tensors out of tensors. It is the subject of Chap. 8.

Chapter 7
Elements of Linear Algebra in Tensor Notation

7.1 Preview

Linear algebra and tensor algebra are inextricably linked. The mechanics of these subjects are similar: linear combinations dominate calculations and change of coordinates (referred to as *change of basis* in linear algebra) is of primary interest. Some ideas are best expressed with matrix notation while others are best expressed with tensors. Matrix notation is more effective at expressing the overall algebraic structure of a calculation. Tensor notation is a better choice when it is necessary to refer to the individual entries of an indexed set of values. For example, the inverse rule

$$(AB)^{-1} = B^{-1}A^{-1}, \tag{7.1}$$

is best expressed with matrices. Meanwhile, the minimization of a quadratic form

$$\frac{1}{2}x^T Ax - x^T b \tag{7.2}$$

should be analyzed in the tensor notation.

This chapter is devoted to exploring the correspondence between these notational systems as well as highlighting the several unique insights that the tensor notation brings to the fundamental topics in linear algebra.

7.2 The Correspondence Between Contraction and Matrix Multiplication

The values of a variant in a given coordinate system form an indexed set of numbers. How one chooses to organize these numbers into a table is a matter of convention. For first- and second-order systems, we adopt a convention that is

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_7,

consistent with that of linear algebra. A variant of order one, T^i or T_i, is represented by a column. The flavor of the index (covariant or contravariant) has no bearing on the representation. A variant of order two, T_{ij}, T^i_j, or T^{ij}, is represented by a matrix in which the first index indicates the row and the second index indicates the column. Thus, it is important to clearly identify the order of the indices in a variant.

For a variant of order higher than two, there is no single commonly accepted convention. For variants of order three, one may imagine a three-dimensional cube of numbers where the first index indicates the slice and the other two indices indicate the row and the column within the slice. Another approach, employed by several software packages, works for any order. The idea is to organize the values as tables within tables. For example, the system $T_{ijk} = i^2 jk$ of order 3 can be represented as

$$\begin{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} & \begin{bmatrix} 2 \\ 4 \end{bmatrix} \\ \begin{bmatrix} 4 \\ 8 \end{bmatrix} & \begin{bmatrix} 8 \\ 16 \end{bmatrix} \end{bmatrix}. \tag{7.3}$$

In this 2×2 matrix of columns, the first two indices determine the row and the column, and the final index determines the position within the innermost column. Similarly, the fourth-order system $T_{ijkl} = i^2 jkl$ can be represented by the table

$$\begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} & \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix} \\ \begin{bmatrix} 4 & 8 \\ 8 & 16 \end{bmatrix} & \begin{bmatrix} 8 & 16 \\ 16 & 32 \end{bmatrix} \end{bmatrix}. \tag{7.4}$$

The first two indices determine the row and the column of the overall table, and the remaining two indices determine the row and the column within the innermost matrix.

The matrix notation works best for two-dimensional systems. A matrix is usually denoted by a single letter. The product AB of two matrices A and B implies a specific contraction of the entries of A and B, as well as a specific way of arranging the result. Consider the product

$$C = AB. \tag{7.5}$$

In indicial notation, this product reads

$$C_{ij} = \sum_k A_{ik} B_{kj}. \tag{7.6}$$

In words, the *second* index of A is contracted with the *first* index of B. The result of this contraction for fixed i and j is placed in the ith row and jth column of the resulting matrix.

Second-order systems can be contracted in other ways, too:

$$C_{ij} = \sum_k A_{ik} B_{jk} \tag{7.7}$$

$$C_{ij} = \sum_k A_{kj} B_{ik} \tag{7.8}$$

$$C_{ji} = \sum_k A_{ik} B_{kj}. \tag{7.9}$$

In order to capture these operations, the matrix notation offers two devices: the order of multiplicative terms and the operation transpose. The above identities are expressed in the matrix notation as follows:

$$C = AB^T \tag{7.10}$$

$$C = BA \tag{7.11}$$

$$C^T = AB. \tag{7.12}$$

Finally, the matrix notation requires the operation of the trace in order to capture contractions within a single square matrix. The trace is defined as the sum of the diagonal entries:

$$\operatorname{Tr} A = \sum_k A_{kk}. \tag{7.13}$$

In contrast with the matrix notation, the tensor notation can be used for systems of arbitrary order and includes only two operations: the tensor product and contraction. The order of the terms is immaterial and the explicit operation of transposition is absent.

The matrix notation is more effective in situations in which the overall algebraic structure is important and when it is best to view matrices as indivisible units. Consider, for example, the proof of the fact that for two invertible matrices A and B, the matrix products AB and BA have identical eigenvalues. Suppose that λ and x are an eigenvalue and a corresponding eigenvector of AB:

$$ABx = \lambda x. \tag{7.14}$$

Multiply both sides on the left by the matrix B

$$BABx = \lambda Bx \tag{7.15}$$

and denote Bx by y:

$$BAy = \lambda y. \tag{7.16}$$

Therefore, λ is an eigenvalue of BA and the nonzero vector Bx is a corresponding eigenvector.

The matrix notation is perfect for this kind of reasoning and truly brings out the essence of the argument. While it is possible to present the same argument in tensor notation, the simplicity of the argument would be obscured by the manipulation of the indices. In short, the matrix notation is preferred in situations where the focus is on the *algebra*.

Example 107. Express the symmetric property

$$A^T = A$$

in the tensor notation.

Example 108. Express the product $C = A^T B$ in the tensor notation.

Example 109. Express the product $C^T = A^T B$ in the tensor notation.

Example 110. Express the product $C^T = A^T B^T$ in the tensor notation and explain why it follows that $(AB)^T = B^T A^T$. This is an example of a theorem in matrix algebra that is best proved in the tensor notation.

Example 111. Use the matrix notation to prove that, for invertible matrices A, B, and C,

$$(ABC)^{-1} = C^{-1} B^{-1} A^{1}. \tag{7.17}$$

7.3 The Fundamental Elements of Linear Algebra in Tensor Notation

Suppose that a basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is chosen in the n-dimensional linear space. Denote the elements of the basis collectively by $\mathbf{e}_i$. A vector $\boldsymbol{v}$ has a unique decomposition v^i with respect to the basis $\mathbf{e}_i$:

$$\mathbf{v} = v^i \mathbf{e}_i. \tag{7.18}$$

The values v^i are called the *components* of the vector v with respect to the basis $\mathbf{e}_i$. We use a superscript in v^i and employ the Einstein summation convention.

How do the components v^i transform under a change of basis? In essence, we already answered this question in Chap. 6 where we concluded that the components of a vector with respect to a covariant basis form a contravariant tensor. Our current goal is to translate those ideas into the matrix notation.

Consider a *new* basis $\mathbf{e}_{i'}$ and suppose that the two bases are related by the identities

$$\mathbf{e}_{i'} = J^{i}_{i'}\mathbf{e}_i \tag{7.19}$$

$$\mathbf{e}_i = J^{i'}_{i}\mathbf{e}_{i'}. \tag{7.20}$$

In the Jacobians $J^{i}_{i'}$ and $J^{i'}_{i}$, assume that the superscript is the first and the subscript is the second. Furthermore, denote the Jacobian $J^{i}_{i'}$ by the matrix X. Then, since $J^{i}_{i'}J^{i'}_{j} = \delta^{i}_{j}$, the Jacobian $J^{i'}_{i}$ corresponds to X^{-1}. Note that in equation (7.19), the contraction takes place on the first index of X. Therefore, in the matrix notation, the basis $\mathbf{e}_{i'}$ is given in terms of $\mathbf{e}_i$ not by the matrix X but by the transpose X^T:

$$\mathbf{E}' = X^T\mathbf{E}, \tag{7.21}$$

where $\mathbf{E}$ and $\mathbf{E}'$ are the basis vectors formally organized into a column:

$$\mathbf{E} = \begin{bmatrix} \mathbf{e}_1 \\ \cdots \\ \mathbf{e}_n \end{bmatrix}; \ \mathbf{E}' = \begin{bmatrix} \mathbf{e}_{1'} \\ \cdots \\ \mathbf{e}_{n'} \end{bmatrix} \tag{7.22}$$

From Chap. 6, we know that the components $v^{i'}$ and v^i are related by the identities

$$v^{i'} = J^{i'}_{i}v^i \tag{7.23}$$

$$v^i = J^{i}_{i'}v^{i'}. \tag{7.24}$$

Translating these relationships to the matrix notation, we find that the components v' (v' is $v^{i'}$ organized into a column) are given in terms of v (v is v^i organized into a column) by multiplication with X^{-1}. Conversely, the components v are given in terms of v' by multiplication with X:

$$v' = X^{-1}v \tag{7.25}$$

$$v = Xv'. \tag{7.26}$$

This is a vivid illustration of the term *contra*variant: the components transform in a way opposite of the bases.

In summary, tensor calculus expressions provide excellent mnemonic rules for establishing correct matrix rules for transformations under a change of basis. The foregoing discussion was not a derivation of equations (7.25) and (7.26)—the derivation took place in Chap. 6—but rather a translation of the easy to remember equations (7.23) and (7.24) into matrix form.

Equations (7.25) and (7.26) can also be derived directly in the matrix notation. Rewrite equation (7.18) in the matrix notation:

$$\mathbf{v} = v^T\mathbf{E}. \tag{7.27}$$

In the equation above, insert, between v^T and $\mathbf{E}$, the identity matrix in the form $\left(X^T\right)^{-1} X^T$:

$$\mathbf{v} = v^T \left(X^T\right)^{-1} X^T \mathbf{E} \tag{7.28}$$

and group the elements on the right-hand side as follows

$$\mathbf{v} = \left(X^{-1} v\right)^T \left(X^T \mathbf{E}\right). \tag{7.29}$$

Since $X^T \mathbf{E}$ is $\mathbf{E}'$, the combination $X^{-1} v$ must be v', which is equation (7.25).

A linear space equipped with an inner product $(\mathbf{u}, \mathbf{v})$ is called a *Euclidean space*. This is different from the sense in which we use the term *Euclidean space* in the rest of this book. The inner product can be evaluated in terms of components

$$(\mathbf{u}, \mathbf{v}) = \left(u^i \mathbf{e}_i, v^j \mathbf{e}_j\right) = \left(\mathbf{e}_i, \mathbf{e}_j\right) u^i v^j. \tag{7.30}$$

The matrix M with entries

$$M_{ij} = \left(\mathbf{e}_i, \mathbf{e}_j\right) \tag{7.31}$$

is called the *Gram matrix*. We may refer to M_{ij} as the *Gram tensor*, but of course it is nothing more than the metric tensor Z_{ij}. It can even be used to juggle indices, as we do later in this chapter. In terms of the Gram tensor, the inner product reads

$$(\mathbf{u}, \mathbf{v}) = M_{ij} u^i v^i. \tag{7.32}$$

In the matrix notation

$$(\mathbf{u}, \mathbf{v}) = u^T M v.$$

For an orthonormal basis $\mathbf{e}_i$, the Gram matrix is the identity.

Once again, tensor calculus offers a simple way to remember the formula for the transformation of the Gram matrix under a change of basis. By analogy with the metric tensor, M_{ij} is a doubly covariant tensor and therefore transforms according to the rule

$$M_{i'j'} = M_{ij} J^i_{i'} J^j_{j'}. \tag{7.33}$$

To express this identity in the matrix notation, first rewrite it as follows:

$$M_{i'j'} = J^i_{i'} M_{ij} J^j_{j'}. \tag{7.34}$$

Note that the first index of $J^i_{i'}$ is contracted with the first index of M_{ij} and the second index of M_{ij} is contracted with the first index of $J^j_{j'}$. Therefore, the matrix equivalent of equation (7.33) is

$$M' = X^T M X. \tag{7.35}$$

Finally, we discuss the matrix representation of linear transformations. Suppose that v is the image of $\mathbf{u}$ under the linear transformation A:

$$\mathbf{v} = A(\mathbf{u}). \tag{7.36}$$

In the matrix notation, this relationship is represented by the matrix product

$$v = Au. \tag{7.37}$$

In the tensor notation, since both u^i and v^i have superscripts, the natural way to express equation (7.37) is to let the first in index in the representation of A be a superscript and the other a subscript:

$$v^i = A^i_j u^j, \tag{7.38}$$

This identity immediately suggests that A^i_j is a tensor with one contravariant and one covariant dimension. This is indeed the case. Therefore A^i_j transforms according to the formula

$$A^{i'}_{j'} = A^i_j J^{i'}_i J^j_{j'}. \tag{7.39}$$

Thus, in the matrix notation, the matrices A' and A are related by

$$A' = X^{-1} A X. \tag{7.40}$$

As a direct consequence of the corresponding theorem in linear algebra, we conclude from equation (7.40) that *the eigenvalues of a tensor A^i_j are invariant* under a change of coordinates.

7.4 Self-Adjoint Transformations and Symmetry

A linear transformation A is called self-adjoint if, for any two vectors $\mathbf{u}$ and $\mathbf{v}$,

$$(A\mathbf{u}, \mathbf{v}) = (\mathbf{u}, A\mathbf{v}). \tag{7.41}$$

Self-adjoint operators are of central importance in applied mathematics, physics, and engineering.

Self-adjoint transformations are said to be represented by symmetric matrices. In actuality, this is only true when the transformation is expressed with respect to an orthonormal basis. The orthonormality of the basis is often implied or omitted altogether. If the basis is not orthonormal, the statement is generally untrue. Fortunately, the tensor notation alerts us to the potential danger. After all, a tensor A^i_j that represents the linear transformation has one superscript and one subscript. That is not a kind of system to which the term symmetric can even be applied in the usual sense. To clarify, a system T_{ij} is called symmetric if $T_{ij} = T_{ji}$. What would symmetry mean for a system such as A^i_j?

The tensor notation will help us discover the correct general characterization of matrices that represent self-adjoint transformations. Express the inner products $(A\mathbf{u}, \mathbf{v})$ and $(\mathbf{u}, A\mathbf{v})$ in the tensor notation

$$(A\mathbf{u}, \mathbf{v}) = M_{ij} A^i_k u^k v^j \tag{7.42}$$

$$(\mathbf{u}, A\mathbf{v}) = M_{ij} u^i A^j_k v^k. \tag{7.43}$$

Reorder the terms and rename the indices in the second identity as follows:

$$(\mathbf{u}, A\mathbf{v}) = M_{ki} A^i_j u^k v^j. \tag{7.44}$$

Since these identities hold for any u^k and v^j, we have

$$M_{ij} A^i_k = M_{ki} A^i_j. \tag{7.45}$$

Rename the indices so that i and j are the two live indices, and take advantage of the symmetry of M_{ij}:

$$M_{ki} A^k_j = M_{kj} A^k_i. \tag{7.46}$$

Finally, rewrite equation (7.46) in the matrix notation:

$$MA = A^T M. \tag{7.47}$$

Equivalently,

$$A = M^{-1} A^T M. \tag{7.48}$$

This is a general statement characterizing a matrix A representing a self-adjoint transformation. It generally implies that A is not symmetric, unless the basis is orthonormal ($M = I$).

The tensor notation offers an elegant way to capture equation (7.48). Use the Gram tensor to lower the superscript on both sides of the equation (7.46).

$$A_{ij} = A_{ji}. \tag{7.49}$$

What a nice identity! It states that, *with the index lowered*, the tensor A_{ij} is symmetric. Of course, that is not equivalent to the matrix A being symmetric because the matrix A has the entries A^i_j, not A_{ij}. When we raise the index i on both sides of equation (7.49), we obtain another elegant identity equivalent to (7.49)

$$A^i_{\cdot j} = A^{\cdot i}_j. \tag{7.50}$$

The tensor on the left has the same entries as the matrix A. The tensor on the right does not. Equation (7.49) illustrates the importance of maintaining or carefully ordered list of indices.

7.5 Quadratic Form Optimization

Quadratic form optimization is a problem in which access to the individual entries of a matrix is necessary in order to apply calculus. A quadratic form is a multivariable generalization of a parabola with a linear shift:

$$f(x) = \frac{1}{2}ax^2 - bx. \tag{7.51}$$

When $a > 0$, $f(x)$ has a minimum at

$$x = \frac{b}{a} \tag{7.52}$$

since

$$f'(x) = ax - b. \tag{7.53}$$

In the case of N variables $x_1, \ldots, x_N$ organized into a column called x, the quadratic form is given by the matrix expression

$$f(x) = \frac{1}{2}x^T Ax - x^T b, \tag{7.54}$$

where A is a symmetric positive definite matrix and the entries of

$$b = \begin{bmatrix} b_1 \\ \cdots \\ b_N \end{bmatrix} \tag{7.55}$$

are constants. The positive definiteness of A is analogous to $a > 0$ in equation (7.51). It is a sufficient and necessary condition in order for $f(x)$ to have a well-defined minimum. The minimum of $f(x)$ occurs at x given by

$$Ax = b, \tag{7.56}$$

which is seen to be entirely analogous to (7.52). Our goal is to derive equation (7.56) by the methods of calculus: that is, by equating the partial derivatives of $f(x)$ to zero.

Exercise 112. Linear algebra offers a derivation that does not require calculus. Introduce a new vector $r = A^{-1}x - b$ and express $f(x)$ in terms of r. Show that the minimum of the new quadratic form occurs at $r = 0$.

Exercise 113. Without the tensor notation, the calculus approach proves surprisingly cumbersome. Derive the partial derivatives of

$$f(x,y) = \frac{1}{2}A_{11}x^2 + A_{12}xy + \frac{1}{2}A_{22}y^2 - b_1x - b_2y \tag{7.57}$$

and show that this calculation leads to equation (7.56).

Exercise 114. Perform the same task in three dimensions for $f(x,y,z)$.

We now use the tensor notation to carry out the calculation outlined in the these exercises. Denote the independent variables by x^i, the matrix A by A_{ij} ($A_{ij} = A_{ji}$), and the vector b by b_i. The quadratic form (7.54) is given by the expression

$$f(x) = \frac{1}{2}A_{ij}x^ix^j - b_ix^i, \tag{7.58}$$

where we have customarily suppressed the index of the function argument. Let us calculate the partial derivative of $f(x)$ with respect to x^k. By the product rule,

$$\frac{\partial f}{\partial x^k} = \frac{1}{2}A_{ij}\frac{\partial x^i}{\partial x^k}x^j + \frac{1}{2}A_{ij}x^i\frac{\partial x^j}{\partial x^k} - b_i\frac{\partial x^i}{\partial x^k}. \tag{7.59}$$

The partial derivative $\partial x^i/\partial x^k$ is captured by the Kronecker symbol:

$$\frac{\partial x^i}{\partial x^k} = \delta^i_k. \tag{7.60}$$

Thus,

$$\frac{\partial f}{\partial x^k} = \frac{1}{2}A_{ij}\delta^i_kx^j + \frac{1}{2}A_{ij}x^i\delta^j_k - b_i\delta^i_k, \tag{7.61}$$

which simplifies to

$$\frac{\partial f}{\partial x^k} = \frac{1}{2}A_{kj}x^j + \frac{1}{2}A_{ik}x^i - b_k. \tag{7.62}$$

Since A_{ij} is symmetric, the first two terms on the right-hand side are equal:

$$\frac{\partial f}{\partial x^k} = A_{ki}x^i - b_k. \tag{7.63}$$

Equating the partial derivatives $\partial f/\partial x^k$ to zero yields

$$A_{ki}x^i = b_k, \tag{7.64}$$

which, of course, is equivalent to equation (7.56)!

Exercise 115. Show that the Hessian $\partial^2 f/\partial x^i \partial x^j$ equals

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = A_{ij}. \tag{7.65}$$

7.6 The Eigenvalue Problem

In this section, we show that, much like $Ax = b$, the eigenvalue problem

$$Ax = \lambda M x \tag{7.66}$$

can be formulated as a variational problem. The matrix M is assumed to be symmetric and positive define. The variational formulation is to find the extrema of

$$f(x) = A_{ij}x^i x^j \tag{7.67}$$

subject to the constraint that

$$M_{ij}x^i x^j = 1. \tag{7.68}$$

Geometrically, equation (7.68) states that the vector x^i unit length. Use a Lagrange multiplier λ to incorporate the constraint in the augmented function $E(x, \lambda)$

$$E(x,\lambda) = A_{ij}x^i x^j - \lambda\left(M_{ij}x^i x^j - 1\right). \tag{7.69}$$

Following earlier analysis,

$$\frac{1}{2}\frac{\partial E(x)}{\partial x^i} = A_{ij}x^j - \lambda M_{ij}x^j. \tag{7.70}$$

Equating the partial derivatives to zero yields

$$A_{ij}x^j = \lambda M_{ij}x^j \tag{7.71}$$

which is equivalent to the eigenvalue problem (7.66).

Exercise 116. Show that equation (7.71) can be written as

$$A_{ij}x^j = \lambda x_i \tag{7.72}$$

and as

$$A^i_j x^j = \lambda x^j. \tag{7.73}$$

Exercise 117. Show that the eigenvalues of the generalized equation (7.71) are invariant with respect to change of basis.

Exercise 118. Show that the eigenvalues of the generalized equation (7.71) are given by the Rayleigh quotient

$$\lambda = A_{ij}x^i x^j. \tag{7.74}$$

7.7 Summary

In this chapter, we discussed the interplay between tensor calculus and linear algebra. We discovered that a number of concepts in linear algebra are clarified with the help of tensor notation. In particular, the matrix rules for the transformations of vector components, the Gram matrix, and the matrix representing a linear operators can be determined from the tensor properties of these objects. On the other hand, tensor calculus draws on several fundamental results from linear algebra. One of those results in the invariance of eigenvalues under similarity transformations. In Chap. 13, the invariance of eigenvalues plays an important role in the study of curvature.

Chapter 8
Covariant Differentiation

8.1 Preview

Chapter 6 demonstrated that the tensor property is the key to invariance. However, a partial derivative $\partial/\partial Z^i$ of a tensor is itself not a tensor. This is a major obstacle in the way of developing differential geometry using the coordinate approach. For example, the expression $\partial T^i/\partial Z^i$ cannot be used as a definition of divergence since it evaluates to different values in different coordinates. Similarly, $Z^{ij}\partial^2 T/\partial Z^i \partial Z^j$ is not invariant and is therefore not a legitimate definition of the Laplacian.

This difficulty was overcome by Gregorio Ricci-Curbastro and his student Tullio Levi-Civita (Fig. 8.1) in their celebrated work *Methods of Absolute Differential Calculus and Their Applications* [34] in the year 1900. Levi-Civita later wrote a classic textbook on the subject entitled *The Absolute Differential Calculus* [28].

The solution comes in the form of a new differential operator, ∇_i, the *covariant derivative*. Similarly to the partial derivative, the covariant derivative satisfies the familiar sum and product rules. Furthermore, it has the desirable property that *it produces tensors out of tensors*. The resulting tensor is one covariant order greater than the original tensor. The *contravariant derivative* ∇^i is defined by raising the index on the covariant derivative:

$$\nabla^i = Z^{ij}\nabla_j. \tag{8.1}$$

The new operator has a number of other desirable properties. One of those key properties is that covariant differentiation reduces to the partial differentiation in Cartesian (more generally, affine) coordinates. This enables us to define the *divergence* of a tensor field T^i as

$$\nabla_i T^i \tag{8.2}$$

and the *Laplacian* of the invariant field T as

$$Z^{ij}\nabla_i\nabla_j T. \tag{8.3}$$

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_8,

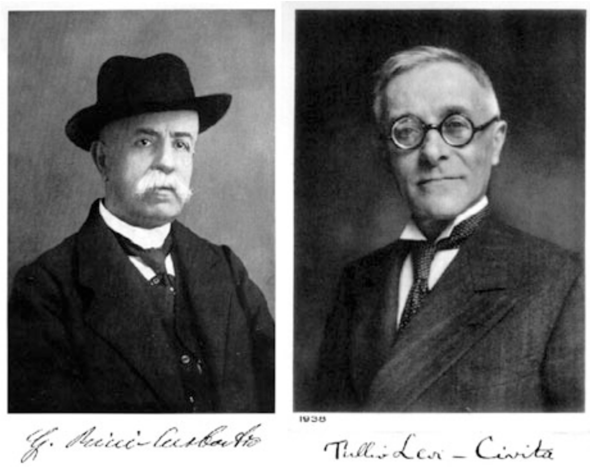


Fig. 8.1 Gregorio Ricci-Curbastro (1853–1925), son of a famous engineer, is considered the inventor of tensor calculus. Tullio Levi-Civita (1873–1941) was Ricci-Curbastro's student. A superb geometer, he was one of the key figures in mathematical foundations of general relativity. Many interesting insights into Levi-Civita's life can be found in [33]

On the one hand, these expressions are invariant. On the other hand, they coincide with the usual definitions in Cartesian coordinates. Therefore, these equations are valid in all coordinates.

The Christoffel symbol defined in Chap. 5 plays a central role in covariant differentiation. The following motivating example illustrates the role of the Christoffel symbol. Following the motivating example, we present the formal definition of the covariant derivative and devote the rest of the chapter to establishing its key properties.

8.2 A Motivating Example

Consider an invariant vector field $\mathbf{V}$ defined in the Euclidean space referred to coordinates Z^i. Then the new object obtained as the partial derivative

$$\frac{\partial \mathbf{V}}{\partial Z^j} \tag{8.4}$$

is a tensor (with vector components), as was demonstrated in Chap. 6. Decompose $\mathbf{V}$ with respect to the covariant basis $\mathbf{Z}_i$:

$$\mathbf{V} = V^i \mathbf{Z}_i. \tag{8.5}$$

Substituting this expansion in equation (8.4) and applying the product rule, we obtain

$$\frac{\partial \mathbf{V}}{\partial Z^j} = \frac{\partial V^i}{\partial Z^j}\mathbf{Z}_i + V^i \frac{\partial \mathbf{Z}_i}{\partial Z^j}. \tag{8.6}$$

The partial derivative $\partial \mathbf{Z}_i / \partial Z^j$ of the covariant basis can, of course, be expressed in terms of the Christoffel symbol Γ^k_{ij}. This is, in fact, the very definition of Γ^k_{ij} which was first given in equation (5.56) and repeated here

$$\frac{\partial \mathbf{Z}_i}{\partial Z^j} = \Gamma^k_{ij}\mathbf{Z}_k. \tag{5.56}$$

Thus

$$\frac{\partial \mathbf{V}}{\partial Z^j} = \frac{\partial V^i}{\partial Z^j}\mathbf{Z}_i + V^i \Gamma^k_{ij}\mathbf{Z}_k. \tag{8.7}$$

Our next goal is to factor out $\mathbf{Z}_i$, so that what remains in parentheses can be interpreted as the components of $\partial \mathbf{V}/\partial Z^j$ with respect to the covariant basis. Rename the dummy index k to i in the second term. In order to avoid using the index i twice in a single expression, also rename the already present dummy index i to m. Then factor out $\mathbf{Z}_i$:

$$\frac{\partial \mathbf{V}}{\partial Z^j} = \left(\frac{\partial V^i}{\partial Z^j} + \Gamma^i_{mj} V^m\right)\mathbf{Z}_i. \tag{8.8}$$

We are nearly done. In a final step, take advantage of the symmetry of the Christoffel symbol with respect to the two lower indices to rewrite this expression as

$$\frac{\partial \mathbf{V}}{\partial Z^j} = \left(\frac{\partial V^i}{\partial Z^j} + \Gamma^i_{jm} V^m\right)\mathbf{Z}_i. \tag{8.9}$$

Now comes the key point: the expression in parentheses

$$\frac{\partial V^i}{\partial Z^j} + \Gamma^i_{jm} V^m \tag{8.10}$$

is a tensor. This is because it represents the components of a tensor with respect to the covariant basis $\mathbf{Z}_i$. (In Sect. 6.3.5, we showed that the component of a vector with respect to $\mathbf{Z}_i$ is a contravariant tensor.). It is interesting that neither term in this expression is a tensor, yet they combine to produce one.

The expression $\partial V^i/\partial Z^j + \Gamma^i_{jm} V^m$ is a better measure of the rate of change in the tensor V^i than $\partial V^i/\partial Z^j$ alone. The quantity $\partial V^i/\partial Z^j$ captures the rate of change in the component V^i. But it also takes the (changing) covariant basis $\mathbf{Z}_i$ to rebuild $\mathbf{V}$ from V^i. For example, the vector field $\mathbf{V}$ may be constant but, in curvilinear coordinates, the components V^i vary. Conversely, a constant tensor field V^i in curvilinear coordinates corresponds to a varying vector field $\mathbf{V}$. However, that partial derivative $\partial V^i/\partial Z^j$ only measures the rate of change of the tensor field $\partial V^i/\partial Z^j$ and ignores the change in the basis $\mathbf{Z}_i$.

The combined expression $\partial V^i/\partial Z^j + \Gamma^i_{jm} V^m$ captures the rates of change of both elements that contribute to $\mathbf{V}$: the component V^i and the covariant basis $\mathbf{Z}_i$. And, as equation (8.9) shows, the tensor combination $\partial V^i/\partial Z^j + \Gamma^i_{jm} V^m$ represents precisely the *components* of the rate of change in the vector field $\mathbf{V}$. We therefore propose the following definition of the *covariant derivative* ∇_j of a contravariant tensor V^i:

$$\nabla_j V^i = \frac{\partial V^i}{\partial Z^j} + \Gamma^i_{jm} V^m. \tag{8.11}$$

This definition needs to be modified for covariant tensors. If $\mathbf{V}$ is decomposed with respect to the contravariant basis $\mathbf{Z}^i$

$$\mathbf{V} = V_i \mathbf{Z}^i \tag{8.12}$$

then we find

$$\frac{\partial \mathbf{V}}{\partial Z^j} = \left(\frac{\partial V_i}{\partial Z^j} - \Gamma^m_{ij} V_m \right) \mathbf{Z}^i. \tag{8.13}$$

We similarly conclude that the combination $\partial V_i/\partial Z^j - \Gamma^m_{ij} V_m$ is a tensor and that it perfectly represent the rate of change of the vector field $\mathbf{V}$. Thus, we define the covariant derivative ∇_j of a covariant tensor V_i according to

$$\nabla_j V_i = \frac{\partial V_i}{\partial Z^j} - \Gamma^m_{ij} V_m. \tag{8.14}$$

The definitions (8.11) and (8.14) are quite similar. The differences are the minus sign in the case of the covariant tensor and index on which the Christoffel symbol is contracted with original tensor.

Exercise 119. Derive equation (8.13).

Exercise 120. Explain why $\nabla_j V_i$ is a tensor.

Exercise 121. In Exercise 82, you were asked to show that the acceleration of a particle moving along the curve $Z^i \equiv Z^i(t)$ is given by

$$A^i = \frac{dV^i}{dt} + \Gamma^i_{jk} V^j V^k. \tag{8.15}$$

Conclude that A^i is a tensor.

8.3 The Laplacian

We now present the Laplacian operator—a central differential operator in applied mathematics. This discussion also illustrates the *variant* property of the covariant derivative. In Chap. 6, a variant was defined as *an object that can be constructed by a similar rule in various coordinate systems*. The covariant derivative satisfies this important requirement: it can be interpreted by a consistent algorithm across all coordinate systems.

The Laplacian ΔF of an invariant field F is usually defined by its expression in Cartesian coordinates:

$$\Delta F = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2}. \tag{8.16}$$

This definition immediately raises a number of questions, the first one being whether this expression evaluates to the same value in different Cartesian coordinates. The answer is *yes*, but to show this outside of the tensor framework requires a lengthy calculation. Nevertheless, (8.16) is a legitimate definition and is our starting point.

In tensor calculus, the Laplacian is defined by the expression

$$\Delta F = Z^{ij} \nabla_i \nabla_j F. \tag{8.17}$$

So far, we have defined the covariant derivative only for tensors of order one. For invariants, such as F, the covariant derivative is defined simply as the partial derivative

$$\nabla_i F = \frac{\partial F}{\partial Z^i}. \tag{8.18}$$

This property is discussed below in Sect. 8.6.2.

Later in this chapter, when we have proved the metrinilic, product, and contraction properties of the covariant derivative, the definition of the Laplacian is captured by an even more compact expression

$$\Delta F = \nabla_i \nabla^i F. \tag{8.19}$$

The central point is the invariance of the expression in (8.17). It is an invariant because all elements of the expression are tensors. It is therefore guaranteed to produce the same value in all coordinates. This simple point is the central idea in tensor calculus and its importance cannot be overstated.

The other important point is the algorithmic nature of the covariant derivative and the definition (8.17). The expression in (8.17) can be interpreted the same way in all coordinate systems and in each case produces an expression in terms of conventional partial derivatives that is valid in that particular coordinate system.

In Cartesian coordinates, the Christoffel symbol vanishes and the covariant derivatives reduce to partial derivatives:

$$\Delta F = Z^{ij} \frac{\partial^2 F}{\partial Z^i \partial Z^j}. \tag{8.20}$$

The contravariant metric tensor Z^{ij} has two nonzero components $Z^{11} = Z^{22} = 1$. Therefore, among the four terms in (8.20), only two survive:

$$\Delta F = \frac{\partial^2 F}{\partial Z^1 \partial Z^1} + \frac{\partial^2 F}{\partial Z^2 \partial Z^2}. \tag{8.21}$$

In other words,

$$\Delta F = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2}. \tag{8.22}$$

Thus, in Cartesian coordinates, the definition (8.17) agrees with the conventional definition (8.16). We can also conclude that the expression (8.22) evaluates to the same value in all Cartesian coordinate systems.

We now calculate the expression for the Laplacian in polar coordinates. We mention at the outset that a more effective way to calculate the Laplacian is expressed by the Voss–Weyl formula discussed in Sect. 9.8.

Recall the nonzero entries of the Christoffel symbols in polar coordinates calculated in Chap. 5:

$$\Gamma^1_{22} = -r \tag{5.77}$$

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r}. \tag{5.78}$$

The function F is now assumed to be expressed as a function of the coordinates $Z^1 = r$ and $Z^2 = \theta$. Since the covariant derivative coincides with the partial derivative for invariants, the values of $\nabla_j F$ are

$$\nabla_1 F = \frac{\partial F}{\partial r} \tag{8.23}$$

$$\nabla_2 F = \frac{\partial F}{\partial \theta}. \tag{8.24}$$

By definition, the second-order covariant derivative $\nabla_i \nabla_j F$ is given by

$$\nabla_i \nabla_j F = \frac{\partial \nabla_j F}{\partial Z^i} - \Gamma^k_{ij} \nabla_k F = \frac{\partial^2 F}{\partial Z^i \partial Z^j} - \Gamma^k_{ij} \frac{\partial F}{\partial Z^k}. \tag{8.25}$$

This indicial identity captures the following four relationships:

$$\nabla_1\nabla_1 F = \frac{\partial^2 F}{\partial r^2} - \Gamma^k_{11}\frac{\partial F}{\partial Z^k} \tag{8.26}$$

$$\nabla_1\nabla_2 F = \frac{\partial^2 F}{\partial r\partial\theta} - \Gamma^k_{12}\frac{\partial F}{\partial Z^k} \tag{8.27}$$

$$\nabla_2\nabla_1 F = \frac{\partial^2 F}{\partial\theta\partial r} - \Gamma^k_{21}\frac{\partial F}{\partial Z^k} \tag{8.28}$$

$$\nabla_2\nabla_2 F = \frac{\partial^2 F}{\partial\theta^2} - \Gamma^k_{22}\frac{\partial F}{\partial Z^k}. \tag{8.29}$$

Plugging in the values of the Christoffel symbol, we find

$$\nabla_1\nabla_1 F = \frac{\partial^2 F}{\partial r^2} \tag{8.30}$$

$$\nabla_1\nabla_2 F = \frac{\partial^2 F}{\partial r\partial\theta} - \frac{1}{r}\frac{\partial F}{\partial\theta} \tag{8.31}$$

$$\nabla_2\nabla_1 F = \frac{\partial^2 F}{\partial\theta\partial r} - \frac{1}{r}\frac{\partial F}{\partial\theta} \tag{8.32}$$

$$\nabla_2\nabla_2 F = \frac{\partial^2 F}{\partial\theta^2} + r\frac{\partial F}{\partial r}. \tag{8.33}$$

Notice that $\nabla_1\nabla_2 F$ and $\nabla_2\nabla_1 F$ are equal. This is an example of the commutative property of the covariant derivative. Commutativity is discussed in Sect. 8.6.4.

Finally, recall the nonzero entries $Z^{11} = 1$ and $Z^{22} = \frac{1}{r^2}$ of the covariant metric Z^{ij}:

$$Z^{ij}\nabla_i\nabla_j F = Z^{11}\nabla_1\nabla_1 F + Z^{22}\nabla_2\nabla_2 F = \frac{\partial^2 F}{\partial r^2} + \frac{1}{r^2}\left(\frac{\partial^2 F}{\partial\theta^2} + r\frac{\partial F}{\partial r}\right), \tag{8.34}$$

or

$$Z^{ij}\nabla_i\nabla_j F = \frac{\partial^2 F}{\partial r^2} + \frac{1}{r}\frac{\partial F}{\partial r} + \frac{1}{r^2}\frac{\partial^2 F}{\partial\theta^2}. \tag{8.35}$$

An essential point of this discussion is that the expressions (8.16) and (8.35) for the Laplacian in Cartesian and polar coordinates are different in terms of partial derivatives but can both be obtained from the tensor definition (8.17).

Exercise 122. Show that in cylindrical coordinates in three dimensions, the Laplacian is given by

$$Z^{ij}\nabla_i\nabla_j F = \frac{\partial^2 F}{\partial r^2} + \frac{1}{r}\frac{\partial F}{\partial r} + \frac{1}{r^2}\frac{\partial^2 F}{\partial\theta^2} + \frac{\partial^2 F}{\partial z^2}. \tag{8.36}$$

Show that this expression is equivalent to

$$Z^{ij}\nabla_i\nabla_j F = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial F}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 F}{\partial \theta^2} + \frac{\partial^2 F}{\partial z^2}. \tag{8.37}$$

Exercise 123. Show that in spherical coordinates in three dimensions, the Laplacian is given by

$$Z^{ij}\nabla_i\nabla_j F = \frac{\partial^2 F}{\partial r^2} + \frac{2}{r}\frac{\partial F}{\partial r} + \frac{1}{r^2}\frac{\partial^2 F}{\partial \theta^2} + \frac{\cot\theta}{r^2}\frac{\partial F}{\partial \theta} + \frac{1}{r^2\sin\theta}\frac{\partial^2 F}{\partial \phi^2}. \tag{8.38}$$

Show that this expression is equivalent to

$$Z^{ij}\nabla_i\nabla_j F = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial F}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial F}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 F}{\partial \phi^2}. \tag{8.39}$$

Exercise 124. Derive the expression for the Laplacian in two dimensions in affine coordinates X and Y related to the Cartesian coordinates x and y by stretching:

$$X = Ax \tag{8.40}$$

$$Y = By. \tag{8.41}$$

8.4 The Formula for $\nabla_i \mathbf{Z}_j$

The discussion of the Laplacian served as a good illustration of the mechanics of the covariant derivative. We would now like to give another such illustration, in which the covariant derivative is applied to the covariant basis $\mathbf{Z}_i$ in polar coordinates. This calculation illustrates the *metrinilic* property of the covariant derivative discussed in detail below in Sect. 8.6.7. As we have been consistently emphasizing, tensor operations work equally well for tensors with scalar and vector elements. The covariant derivative is yet another example of this.

Let us begin by calculating the vector $\nabla_1\mathbf{Z}_1$. By the definition of the covariant derivative, we have

$$\nabla_1\mathbf{Z}_1 = \frac{\partial \mathbf{Z}_1}{\partial r} - \Gamma^k_{11}\mathbf{Z}_k. \tag{8.42}$$

Here, we would like to point out an important characteristic of the covariant derivative: each element of $\nabla_i\mathbf{Z}_j$ involves all elements of $\mathbf{Z}_i$. This is unlike the partial derivative, where $\partial\mathbf{Z}_1/\partial r$ can be computed by itself, without referring to $\mathbf{Z}_2$.

As we established in Chap. 5, $\mathbf{Z}_1$ is a unit vector that points directly away from the origin. Therefore, for a given θ, $\mathbf{Z}_1$ does not change with r and we have

$$\frac{\partial \mathbf{Z}_1}{\partial r} = \mathbf{0}. \tag{8.43}$$

Furthermore, since both entries Γ^k_{11} of the Christoffel symbol vanish, we have

$$\nabla_1 \mathbf{Z}_1 = \mathbf{0}. \tag{8.44}$$

Next, let us calculate $\nabla_1 \mathbf{Z}_2$. First, we write out the definition:

$$\nabla_1 \mathbf{Z}_2 = \frac{\partial \mathbf{Z}_2}{\partial r} - \Gamma^k_{12} \mathbf{Z}_k. \tag{8.45}$$

The vectors $\mathbf{Z}_2$ are orthogonal $\mathbf{Z}_1$ and their length equals r. Therefore, for a given value of the coordinate θ, the vectors $\mathbf{Z}_2$ all point in the same direction and their length grows linearly with r. Thus $\partial \mathbf{Z}_2 \partial r$ is the unit vector pointing in the same direction. Since $\Gamma^1_{12} = 0$ and $\Gamma^2_{12} = r^{-1}$ [see equation (5.78)] we find that the combination $\Gamma^k_{12} \mathbf{Z}_k$ is the unit vector that points in the direction opposite of $\mathbf{Z}_2$ and therefore

$$\nabla_1 \mathbf{Z}_2 = \mathbf{0}. \tag{8.46}$$

We now turn to $\nabla_2 \mathbf{Z}_1$ which is given by

$$\nabla_2 \mathbf{Z}_1 = \frac{\partial \mathbf{Z}_1}{\partial \theta} - \Gamma^k_{21} \mathbf{Z}_k. \tag{8.47}$$

The rest of the calculation for $\nabla_2 \mathbf{Z}_1$ is contained in the following exercises.

Exercise 125. Using geometric arguments, explain why $\partial \mathbf{Z}_1 / \partial \theta$ is the unit vector that is orthogonal to $\mathbf{Z}_1$.

Exercise 126. Conclude that

$$\nabla_2 \mathbf{Z}_1 = \mathbf{0}. \tag{8.48}$$

Exercise 127. Give an alternative derivation of this identity by showing the general property that $\nabla_i \mathbf{Z}_j$ is a symmetric object

$$\nabla_i \mathbf{Z}_j = \nabla_j \mathbf{Z}_i. \tag{8.49}$$

Hint: $\partial \mathbf{Z}_j / \partial Z^i = \partial^2 \mathbf{R} / \partial Z^i \partial Z^j$ and $\Gamma^k_{ij} = \Gamma^k_{ji}$.

Exercise 128. Show that

$$\nabla_2 \mathbf{Z}_2 = \mathbf{0}. \tag{8.50}$$

We have just shown that all elements of $\nabla_i \mathbf{Z}_j$ in polar coordinates are zero:

$$\nabla_i \mathbf{Z}_j = \mathbf{0}. \tag{8.51}$$

Importantly, this is not a special property of polar coordinates, but true in all coordinates. This is an important feature of the covariant derivative. Note that one of the main reasons that Cartesian coordinates are so convenient in many problems is that the coordinate basis $\mathbf{i}, \mathbf{j}$ is unchanged from one point to another. Operating in curvilinear coordinates sacrifices this desirable property. However, at least from the algebraic point of view, the covariant derivative restores it!

This important property of the covariant derivative can be demonstrated in several ways. Once we prove that tensor property of the covariant derivative, we will be able to conclude that $\nabla_i \mathbf{Z}_j$ vanishes in all coordinate systems because it vanishes in a particular coordinate system. (Note, that we could have shown even more easily that $\nabla_i \mathbf{Z}_j = \mathbf{0}$ in Cartesian coordinates.) This is a common way of showing that a tensor vanishes. It was justified in Exercise 94 in Chap. 6.

This property can also be easily demonstrated by a direct algebraic calculation. By definition,

$$\nabla_i \mathbf{Z}_j = \frac{\partial \mathbf{Z}_j}{\partial Z_i} - \Gamma_{ij}^k \mathbf{Z}_k. \tag{8.52}$$

However, by the very definition of the Christoffel symbol given in equation (5.56),

$$\frac{\partial \mathbf{Z}_j}{\partial Z_i} = \Gamma_{ij}^k \mathbf{Z}_k. \tag{8.53}$$

Therefore

$$\nabla_i \mathbf{Z}_j = \mathbf{0}. \tag{8.54}$$

Exercise 129. Show similarly that

$$\nabla_i \mathbf{Z}^j = \mathbf{0}. \tag{8.55}$$

We now turn to the task of defining the covariant derivative for variants of order greater than one.

8.5 The Covariant Derivative for General Tensors

The definitions (8.11) and (8.14) can be extended to variants with arbitrary collections of indices. The general definition includes a Christoffel term for each index. We give the complete definition for a variant T_j^i with one covariant and one contravariant index. It is clear how this definition extends to variants with other combinations of indices. For T_j^i, the definition reads

$$\nabla_k T^i_j = \frac{\partial T^i_j}{\partial Z^k} + \Gamma^i_{km} T^m_j - \Gamma^m_{jk} T^i_m. \tag{8.56}$$

As you can see, for each index there is a term in which T^i_j is multiplied by the Christoffel symbol and the index is contracted with the appropriate index from the Christoffel symbol.

Let us apply the recipe encoded in equation (8.56) to the triply covariant tensor T_{ijk}. In additional to the partial derivative $\partial T_{rst}/\partial Z^k$, we have a Christoffel term with a minus sign for each covariant index:

$$\nabla_k T_{rst} = \frac{\partial T_{rst}}{\partial Z^k} - \Gamma^m_{kr} T_{mst} - \Gamma^m_{ks} T_{rmt} - \Gamma^m_{kt} T_{rsm}. \tag{8.57}$$

Exercise 130. Determine the expression for the covariant derivative $\nabla_l T^{ij}_k$.

Exercise 131. Show that

$$\nabla_k \delta^i_j = 0. \tag{8.58}$$

This is a special case of the more general metrinilic property of the covariant derivative discussed below in Sect. 8.6.7.

In the next section we discuss the key properties of the covariant derivative. These properties make the covariant derivative as simple to use as the partial derivative. However, its tensor property make it much more powerful than the partial derivative. When an analysis is carried out within the tensor framework that employs covariant differentiation, that analysis is *simultaneously valid in all coordinate systems*. Therefore, the tensor approach offers the best of both worlds: it utilizes the advantages of coordinate systems while producing results that are independent of coordinate systems.

8.6 Properties of the Covariant Derivative

8.6.1 The Tensor Property

The key property of the covariant derivative and, in fact, the very motivation for its invention, is that it produces tensors out of tensors. The implications of this property are crucial and far reaching. For example, the fact that

$$\nabla_i V_j \tag{8.59}$$

is tensor implies that

$$Z^{ij} \nabla_i V_j = \nabla^j V_j \tag{8.60}$$

is an invariant and is therefore a suitable definition for the divergence of a tensor field V_j. Similarly, the fact that $\nabla_i \nabla_j F$ leads to the definition of the Laplacian as we saw in Sect. 8.3.

In this section, we would like to present an elegant argument that proves the tensor property of the covariant derivative. We have already demonstrated the tensor property for the contravariant component V^i of an invariant vector field $\mathbf{V}$. However, any tensor or order one can be treated as the component of an invariant field and that is the basis of the proof.

Suppose that V^i is a contravariant tensor. Then $\mathbf{V} = V^i \mathbf{Z}_i$ is an invariant and its partial derivative

$$\frac{\partial \mathbf{V}}{\partial Z^j} \tag{8.61}$$

is a tensor. We repeat equation (8.13)

$$\frac{\partial \mathbf{V}}{\partial Z^j} = \left(\frac{\partial V^i}{\partial Z^j} + \Gamma^i_{jm} V^m \right) \mathbf{Z}_i, \tag{8.13}$$

and again use the argument that the combination $\partial V^i / \partial Z^j + \Gamma^i_{jm} V^m$ must be a tensor since it represents the component of a tensor field $(\partial \mathbf{V} / \partial Z^j)$ with respect to the covariant basis $\mathbf{Z}_i$. Therefore, $\nabla_j V^i = \partial V^i / \partial Z^j + \Gamma^i_{jm} V^m$ is a tensor.

Exercise 132. Prove that $\nabla_j V_i$ is a tensor by forming the invariant $V_i \mathbf{Z}^i$ and analyzing its partial derivative with respect to Z^j.

I believe that this elegant proof of the tensor property leaves little doubt that, for a tensor $\mathbf{T}^i$ with *vector* elements, the covariant derivative produces a tensor. However, the argument doesn't quite work since the contraction $\mathbf{T}^i \mathbf{Z}_i$ is meaningless. Instead, one needs to consider the scalar invariant T formed by the dot product $T = \mathbf{T}^i \cdot \mathbf{Z}_i$ and this argument is left as an exercise.

Exercise 133. For a contravariant vector $\mathbf{T}^i$, prove that $\nabla_j \mathbf{T}^i$ is a tensor by forming the invariant $T = \mathbf{T}^i \cdot \mathbf{Z}_i$.

Exercise 134. For a covariant vector $\mathbf{T}_i$, prove that $\nabla_j \mathbf{T}_i$ is a tensor by forming the invariant $T = \mathbf{T}_i \cdot \mathbf{Z}^i$.

So far, we have demonstrated the tensor property of the covariant derivative for tensors of order one with scalar and vector elements. This opens up an elegant way for proving the tensor property of the covariant derivative for tensors of order two. For example, suppose that T_{ij} is a covariant tensor of order two. Then $\mathbf{T}_j = T_{ij} \mathbf{Z}^i$ is a covariant tensor of order one for which the tensor property of the covariant derivative has already been established. Therefore,

$$\nabla_k \mathbf{T}_j \tag{8.62}$$

is a tensor. By the product rule, which is proved below in Sect. 8.6.6,

$$\nabla_k \mathbf{T}_j = \nabla_k \left(T_{ij} \mathbf{Z}^i\right) = \nabla_k T_{ij} \; \mathbf{Z}^i + T_{ij} \nabla_k \mathbf{Z}^i. \tag{8.63}$$

Since $\nabla_k \mathbf{Z}^i$ is zero, we have

$$\nabla_k \mathbf{T}_j = \nabla_k T_{ij} \; \mathbf{Z}^i$$

and if follows that $\nabla_k T_{ij}$ is a tensor since it is the covariant component of a tensor field $\nabla_k \mathbf{T}_j$.

Exercise 135. Show similarly that $\nabla_k T^i_j$ and $\nabla_k T^{ij}$ are tensors.

Exercise 136. Show, by forming dot products with the covariant and contravariant bases, that $\nabla_k \mathbf{T}_{jj}$, $\nabla_k \mathbf{T}^i_{\cdot j}$, and $\nabla_k \mathbf{T}^{ij}$ are tensors.

Exercise 137. Argue the general tensor property by induction. That is, if $T^{ijk\ldots}_{rst\ldots}$ is a tensor, then $\nabla_p T^{ijk\ldots}_{rst\ldots}$ is also a tensor.

This concludes our discussion of the paramount tensor property of the covariant derivative. The proof present in this section is quite elegant and robust. However, it is also of great value to be able to demonstrate the tensor property more directly by showing that the covariant derivative of a tensor transforms under a change of variables in an appropriate way. This is accomplished below in Sect. 8.7

8.6.2 The Covariant Derivative Applied to Invariants

The covariant derivative ∇_k coincides with partial derivative $\partial/\partial Z^k$ for invariants, or any variants of order zero. Rather than labeling this fact as a *property* of the covariant derivative, it is better viewed as a reasonable extension of definition (8.56) to variants of order zero. Indeed, definition (8.56) instructs us to include a Christoffel term of each index. Since variants of order zero do not have any indices, the right-hand side reduces to a single term with the partial derivative. This property is used quite frequently in tensor calculations when there is a need to relate partial and covariant derivatives.

8.6.3 The Covariant Derivative in Affine Coordinates

The covariant derivative ∇_k coincides with partial derivative $\partial/\partial Z^k$ in affine and, in particular, Cartesian coordinates. This is true because the Christoffel symbol vanishes in affine coordinates. This is an important property and we have already used it to a great effect in a proof of the metrinilic property. It is frequently used in this way: a certain characteristic is demonstrated in affine coordinates and subsequently extended to general coordinates by the tensor property. This way of reasoning was illustrated in Exercise 94 in Chap. 6. Another great example of a proof like this comes in the very next section where we show that covariant derivatives commute. It is a fundamental property of partial derivatives, but it proves much more subtle and geometrically insightful for covariant derivatives.

8.6.4 Commutativity

Covariant derivatives commute, that is

$$\nabla_k \nabla_l T^i_j = \nabla_l \nabla_k T^i_j \tag{8.64}$$

and this property holds for tensors with any collection of indices. The difference

$$\nabla_k \nabla_l T^i_j - \nabla_l \nabla_k T^i_j. \tag{8.65}$$

is often written as $(\nabla_k \nabla_l - \nabla_l \nabla_k)\, T^i_j$ in order to emphasize the switching of the operators. The operator $\nabla_k \nabla_l - \nabla_l \nabla_k$ is commonly referred to as the *commutator*.

The cornerstone of the upcoming argument is that the difference (8.65) is a tensor. Furthermore, in Cartesian coordinates, it reduces to the difference of partial derivatives

$$\frac{\partial^2 T^i_j}{\partial Z^k \partial Z^l} - \frac{\partial^2 T^i_j}{\partial Z^l \partial Z^k} \tag{8.66}$$

and therefore vanishes due to the commutative property of the partial derivative. Consequently, being a tensor, it must vanish in all coordinate systems.

Given how easy it is to justify, commutativity may seem to be a relatively unremarkable property. However, this couldn't be further from the truth. For example, in Sect. 8.129, we use this fact to show that the object $R^k_{\cdot mij} = \partial \Gamma^k_{im}/\partial Z^j - \partial \Gamma^k_{jm}/\partial Z^i + \Gamma^k_{jn}\Gamma^n_{im} - \Gamma^k_{in}\Gamma^n_{jm}$ is a tensor and must therefore vanish in all coordinates system. This is quite a remarkable relationship which can be used for many purposes. For example, given a variant G^i_{jk} of order three, it can be used to determine whether G^i_{jk} can serve as the Christoffel symbol for some coordinate system.

Furthermore, the commutative property of the covariant derivative and the corresponding property that $R^k_{\cdot mij} = 0$ no longer holds in non-Euclidean spaces. Non-Euclidean spaces arise in many fundamental applications that include differential geometry on curved surfaces and Einstein's general relativity theory. We study the differential geometry on curved surfaces starting with Chap. 10.

8.6.5 The Sum Rule

The covariant derivative satisfies the sum rule

$$\nabla_k \left(T^i_j + S^i_j \right) = \nabla_k T^i_j + \nabla_k S^i_j. \tag{8.67}$$

This rule holds for tensors with arbitrary collections of indices and its proof is left as an easy exercise. The sum rule is easily generalized to linear combinations

$$\nabla_k \left(a T^i_j + b S^i_j \right) = a \nabla_k T^i_j + b \nabla_k S^i_j, \tag{8.68}$$

where a and b are constants. There's very little more that can be said about the sum rule and we are therefore moving on to the product rule.

8.6.6 The Product Rule

The covariant derivative satisfies the product rule also known as *Leibniz's law*. The product rule reads

$$\nabla_k \left(T^i U_j \right) = \nabla_k T^i \; U_j + T^i \nabla_k U_j \tag{8.69}$$

and holds for tensors with arbitrary collections of indices. To demonstrate the product rule for $S^i T_j$, apply the definition of the covariant derivative

$$\nabla_k \left(T^i U_j \right) = \frac{\partial \left(T^i U_j \right)}{\partial Z^k} + \Gamma^i_{km} T^m U_j - \Gamma^m_{kj} T^i U_m. \tag{8.70}$$

Since the partial derivative $\partial / \partial Z^k$ satisfies the usual product rule, we have

$$\nabla_k \left(T^i U_j \right) = \frac{\partial T^i}{\partial Z^k} U_j + T^i \frac{\partial U_j}{\partial Z^k} + \Gamma^i_{km} T^m U_j - \Gamma^m_{kj} T^i U_m. \tag{8.71}$$

Group the first term with the third and the second with the fourth

$$\nabla_k \left(T^i U_j \right) = \left(\frac{\partial T^i}{\partial Z^k} + \Gamma^i_{km} T^m \right) U_j + T^i \left(\frac{\partial U_j}{\partial Z^k} - \Gamma^m_{kj} U_m \right), \tag{8.72}$$

and note that the parenthesized expressions are $\nabla_k T^i$ and $\nabla_k U_j$. This completes the proof.

Note that the product rule also applies in the case of the dot product:

$$\nabla_k \left(\mathbf{T}^i \cdot \mathbf{U}_j\right) = \nabla_k \mathbf{T}^i \cdot \mathbf{U}_j + \mathbf{T}^i \cdot \nabla_k \mathbf{U}_j. \tag{8.73}$$

The proof of this property is left as an exercise.

8.6.7 The Metrinilic Property

The metrics $\mathbf{Z}_i$, $\mathbf{Z}^i$, Z_{ij}, Z^{ij} and δ^i_j vanish under the covariant derivative. This is the *metrinilic* property. For the bases $\mathbf{Z}_i$ and $\mathbf{Z}^i$, we have already encountered this property in Sect. 8.4, where it was demonstrated in two different ways. Furthermore, this property was shown for the Kronecker symbol δ^i_j in Exercise 131.

For the metric tensors Z_{ij} and Z^{ij} it can be shown by a simple application of the product rule. For example, for the covariant metric tensor Z_{ij}, we have

$$\nabla_k Z_{ij} = \nabla_k \left(\mathbf{Z}_i \cdot \mathbf{Z}_j\right) = \nabla_k \mathbf{Z}_i \cdot \mathbf{Z}_j + \mathbf{Z}_i \cdot \nabla_k \mathbf{Z}_j = 0.$$

Exercise 138. Show the metrinilic property for the contravariant metric tensor Z^{ij} by the product rule.

Exercise 139. Show the metrinilic property for the Kronecker symbol δ^i_j by the product rule.

Exercise 140. Show the metrinilic property for the delta symbols δ^{ij}_{rs} and δ^{ijk}_{rst} by expressing them in terms of the Kronecker delta symbol δ^i_j.

Exercise 141. Show the metrinilic property for the delta symbol δ^{ij}_{rs} by a direct application of the definition

Exercise 142. Use the property that $\nabla_k Z_{ij} = 0$ to show that

$$\frac{\partial Z_{ij}}{\partial Z^k} = \Gamma_{i,jk} + \Gamma_{j,ik}. \tag{8.74}$$

Exercise 143. Use the property that $\nabla_k Z^{ij} = 0$ to derive the expression of $\partial Z^{ij}/\partial Z^k$.

Let us summarize the metrinilic property:

$$\nabla_j \mathbf{Z}_i,\ \nabla_j \mathbf{Z}^i = \mathbf{0} \tag{8.75}$$

$$\nabla_k Z_{ij},\ \nabla_k Z^{ij} = 0 \tag{8.76}$$

$$\nabla_p \delta^i_j,\ \nabla_p \delta^{ij}_{rs},\ \nabla_p \delta^{ijk}_{rst} = 0. \tag{8.77}$$

In Chap. 9, we encounter the Levi-Civita symbols ε^{ijk} and ε_{ijk} that also belong to the group of metrics. The metrinilic property extends to the Levi-Civita symbols as well:

$$\nabla_p \varepsilon^{ijk}, \nabla_p \varepsilon_{ijk} = 0.$$

The metrinilic property has far-reaching implications. The first impact is on the essential operation of index juggling. We have grown accustomed to raising and lowering free indices on both sides of an identity. For example, we know that

$$T_i = S_i^{\cdot j} U_j^{kl} V_{kl} \tag{8.78}$$

is equivalent to

$$S^i = T^{ij} U_j^{kl} V_{kl}. \tag{8.79}$$

Let us remind ourselves of why this true: the latter identity can be obtained from the former by contracting both sides with the contravariant metric Z^{ir} (and then renaming $r \to i$).

What happens, however, when the index that we intend to juggle occurs under the covariant derivative? For example, consider the identity

$$S_i = \nabla_k T_i^{j}\ U_j^{kl} V_l. \tag{8.80}$$

Contracting both sides with Z^{ir} yields

$$S^r = Z^{ir} \nabla_k T_i^{j}\ U_j^{kl} V_l. \tag{8.81}$$

Can Z^{ir} be passed under the covariant derivative so that

$$S^r = \nabla_k \left(Z^{ir} T_i^{j} \right) U_j^{kl} V_l\ ? \tag{8.82}$$

Thanks to the metrinilic property the answer is *yes*. By the product rule, we have

$$Z^{ir} \nabla_k T_i^{\cdot j} = \nabla_k \left(Z^{ir} T_i^{j} \right) - \nabla_k Z^{ir}\ T_i^{j} \tag{8.83}$$

and the second term vanishes by the metrinilic property, yielding

$$Z^{ir} \nabla_k T_i^{\cdot j} = \nabla_k \left(Z^{ir} T_i^{j} \right). \tag{8.84}$$

This shows that *metric tensors pass freely across the covariant derivative.* Thus,

$$S^r = \nabla_k T^{rj} \, U_j^{kl} V_l \ , \tag{8.85}$$

or, by renaming $r \to i$,

$$S^i = \nabla_k T^{ij} \, U_j^{kl} V_l \ ? \tag{8.86}$$

We have therefore shown that *free indices can be juggled across the covariant derivative.*

We have also grown accustomed to the fact that contracted indices can exchange flavors. For example,

$$S_j^i T_i^k = S_{ij} T^{ik}. \tag{8.87}$$

Is this operation still valid if the exchange takes place across the covariant derivative, as in

$$S_j^i \nabla_l T_i^k = S_{ij} \nabla_l T^{ik} \ ? \tag{8.88}$$

Once again, thanks to the metrinilic property, the answer is *yes* and the proof is left as an exercise.

Exercise 144. Show that contracted indices can exchange flavors across the covariant derivative.

In addition to being able to freely juggle indices on the arguments of the covariant derivative, we are able to freely juggle indices on the covariant derivative itself. For example, the expression

$$T^i \nabla_i S_k^j \tag{8.89}$$

is clearly equivalent to

$$T_i \nabla^i S_k^j. \tag{8.90}$$

Furthermore,

$$T_i \nabla_j \nabla^i S_l^k = T^i \nabla_j \nabla_i S_l^k. \tag{8.91}$$

Finally, the Laplacian $\Delta F = Z^{ij} \nabla_i \nabla_j F$ can be written compactly as

$$\Delta F = \nabla_i \nabla^i F. \tag{8.92}$$

Written in this way, it is quite reminiscent of its Cartesian origins as the sum of second derivatives. We do not need to state again, however, just how powerful equation (8.92) is by virtue of being valid in any coordinate system. Since the symbol $\nabla_i \nabla^i$ is so compact, from this point on we will is it instead of the more conventional symbol Δ.

Another important consequence of the metrinilic property is that we are able to study vector fields by analyzing their components. Suppose that V^i is the contravariant component of the vector field $\mathbf{V}$:

$$\mathbf{V} = V^i \mathbf{Z}_i. \tag{8.93}$$

Then

$$\frac{\partial \mathbf{V}}{\partial Z^k} = \nabla_k \mathbf{V} = \nabla_k V^i \ \mathbf{Z}_i. \tag{8.94}$$

Therefore, we see that the tensor $\nabla_k V^i$ carries all of the information regarding the rate of change of the vector field $\mathbf{V}$. By contrast, the partial derivative $\partial V^i / \partial Z^k$ fails to capture the change in the accompanying covariant basis.

8.6.8 Commutativity with Contraction

Covariant differentiation commutes with contraction. Consider the expression

$$\nabla_k T^i_{ij}. \tag{8.95}$$

It can be interpreted in two ways. First, it can be seen as the covariant derivative applied to the tensor $S_j = T^i_{ij}$. The tensor S_j is of order one. Its covariant derivative $\nabla_k S_j$ is given by

$$\nabla_k S_j = \frac{\partial S_j}{\partial Z^k} - \Gamma^m_{jk} S_m. \tag{8.96}$$

Therefore, in this interpretation, $\nabla_k T^i_{ij}$ reads

$$\nabla_k T^i_{ij} = \frac{\partial T^i_{ij}}{\partial Z^k} - \Gamma^m_{kj} T^i_{im}. \tag{8.97}$$

Alternatively, $\nabla_k T^i_{ij}$ can be interpreted as the covariant derivative applied to the tensor T^i_{rj} of order three to produce the tensor $\nabla_k T^i_{rj}$ of order four which is subsequently contracted on i and r. Following this interpretation, the tensor $\nabla_k T^i_{rj}$ is given by

$$\nabla_k T^i_{rj} = \frac{\partial T^i_{rj}}{\partial Z^k} + \Gamma^i_{km} T^m_{rj} - \Gamma^m_{kr} T^i_{mj} - \Gamma^m_{kj} T^i_{rm}. \tag{8.98}$$

Now, let us contract both sides on i and r:

$$\nabla_k T^i_{ij} = \frac{\partial T^i_{ij}}{\partial Z^k} + \Gamma^i_{km} T^m_{ij} - \Gamma^m_{ki} T^i_{mj} - \Gamma^m_{kj} T^i_{im}. \tag{8.99}$$

The terms $\Gamma^i_{km} T^m_{ij}$ are $\Gamma^m_{ki} T^i_{mj}$ correspond to the two indices in T^i_{rj} that we contracted. These terms are equal as can be seen by exchanging the names of the indices i and m in one of them. Therefore, these terms cancel and we find that the second interpretation of the expression $\nabla_k T^i_{ij}$ yields

$$\nabla_k T^i_{ij} = \frac{\partial T^i_{ij}}{\partial Z^k} - \Gamma^m_{kj} T^i_{im}, \tag{8.100}$$

which is equivalent to the contraction-first interpretation.

Thanks to the contraction property there is, in fact, no ambiguity in the expression $\nabla_k T^i_{ij}$. Had this not been the case, the Einstein notation would have been viable and there would have been a need to reintroduce the summation sign to distinguish between

$$\sum_{i=r} \nabla_k T^i_{rj} \text{ and } \nabla_k \sum_{i=r} T^i_{rj}. \tag{8.101}$$

This contraction property completes the list of the key properties of the covariant derivative. As we see, despite its apparent complexity, the covariant derivative is simply a better alternative to the partial derivative.

The next section is devoted to a very important exercise: proving that the tensor property by deriving the rule by which the covariant derivative transforms under a change of variables.

8.7 A Proof of the Tensor Property

The marquee property of the covariant derivative is that it produces tensors out of tensors. We have established this property in Sect. 8.6.1 by an elegant argument. In this section, we will give a more direct proof based on establishing the rule by which the covariant derivative transforms under a change of variables. We show this property for three types of tensors: a covariant tensor T_i, a contravariant tensor T^i, and a tensor of order two T^i_j. We do not give a proof for a tensor with an arbitrary collection of indices, but it will be apparent to the reader that the proof can be extended to those tensors, as well.

8.7.1 A Direct Proof of the Tensor Property for $\nabla_j T_i$

Consider the covariant derivative $\nabla_j T_i$ for a covariant tensor T_i. In the primed coordinates $Z^{i'}$, $\nabla_{j'} T_{i'}$ is given by

$$\nabla_{j'} T_{i'} = \frac{\partial T_{i'}}{\partial Z^{j'}} - \Gamma^{k'}_{i'j'} T_{k'}. \tag{8.102}$$

We have previously established how each of the terms in this expression transforms under the change of coordinates. We repeat those relationships here

$$\frac{\partial T_{i'}}{\partial Z^{j'}} = \frac{\partial T_i}{\partial Z^j} J^i_{i'} J^j_{j'} + J^i_{i'j'} T_i \tag{8.103}$$

$$\Gamma^{k'}_{i'j'} = \Gamma^k_{ij} J^i_{i'} J^j_{j'} J^{k'}_k + J^k_{i'j'} J^{k'}_k \tag{8.104}$$

$$T_{k'} = T_l J^l_{k'} \tag{8.105}$$

The last two relationships help establish the transformation rule for the contraction $\Gamma^{k'}_{i'j'} T_{k'}$

$$\Gamma^{k'}_{i'j'} T_{k'} = \left(\Gamma^k_{ij} J^{i'}_i J^{j'}_j J^{k'}_k + J^k_{i'j'} J^{k'}_k \right) T_l J^l_{k'} = \Gamma^k_{ij} J^{i'}_i J^{j'}_j T_k + J^k_{i'j'} T_k, \tag{8.106}$$

where the justification of the second equality is left as an exercise. Thus, combining both terms in equation (8.102), we find

$$\nabla_{j'} T_{i'} = \frac{\partial T_i}{\partial Z^j} J^i_{i'} J^j_{j'} + J^i_{i'j'} T_i - \Gamma^k_{ij} J^{i'}_i J^{j'}_j T_k - J^k_{i'j'} T_k. \tag{8.107}$$

The terms with the second-order Jacobians cancel and we find

$$\nabla_{j'} T_{i'} = \left(\frac{\partial T_i}{\partial Z^j} - \Gamma^k_{ij} T_k \right) J^i_{i'} J^j_{j'}, \tag{8.108}$$

which is precisely

$$\nabla_{j'} T_{i'} = \nabla_j T_i J^i_{i'} J^j_{j'}, \tag{8.109}$$

which shows that $\nabla_j T_i$ is a tensor.

8.7.2 A Direct Proof of the Tensor Property for $\nabla_j T^i$

The proof for the contravariant tensor T^i will require an additional step. In coordinates $Z^{i'}$, $\nabla_{j'} T^{i'}$ is

$$\nabla_{j'} T^{i'} = \frac{\partial T^{i'}}{\partial Z^{j'}} + \Gamma^{i'}_{j'k'} T^{k'}. \tag{8.110}$$

Let us catalog how each of the terms transforms under the change of coordinates:

$$\frac{\partial T^{i'}}{\partial Z^{j'}} = \frac{\partial T^{i}}{\partial Z^{j}} J^{j}_{j'} J^{i'}_{i} + T^{i} J^{i'}_{ij} J^{j}_{j'} \tag{8.111}$$

$$\Gamma^{i'}_{j'k'} = \Gamma^{i}_{jk} J^{i'}_{i} J^{j}_{j'} J^{k}_{k'} + J^{i}_{j'k'} J^{i'}_{i} \tag{8.112}$$

$$T^{k'} = T^{l} J^{k'}_{l}. \tag{8.113}$$

From the last two relationships we find

$$\Gamma^{i'}_{j'k'} T^{k'} = \left(\Gamma^{i}_{jk} J^{i'}_{i} J^{j}_{j'} J^{k}_{k'} + J^{i}_{j'k'} J^{i'}_{i} \right) T^{l} J^{k'}_{l} = \Gamma^{i}_{jk} J^{i'}_{i} J^{j}_{j'} T^{k} + J^{i}_{j'k'} J^{i'}_{i} J^{k'}_{l} T^{l}. \tag{8.114}$$

Combining all terms we have

$$\nabla_{j'} T^{i'} = \frac{\partial T^{i}}{\partial Z^{j}} J^{j}_{j'} J^{i'}_{i} + T^{i} J^{i'}_{ij} J^{j}_{j'} + \Gamma^{i}_{jk} J^{i'}_{i} J^{j}_{j'} T^{k} + J^{i}_{j'k'} J^{i'}_{i} J^{k'}_{l} T^{l}. \tag{8.115}$$

The first and third terms combine to yields $\nabla_j T^i J^{j}_{j'} J^{i'}_{i}$. However, this time, the fact that the terms containing the second-order Jacobians cancel is not immediately clear. Factoring out T^i (this requires some index renaming), we find

$$\nabla_j T^i = \nabla_j T^i J^{j}_{j'} J^{i'}_{i} + \left(J^{i'}_{ij} J^{j}_{j'} + J^{j}_{j'k'} J^{i'}_{j} J^{k'}_{i} \right) T^i. \tag{8.116}$$

To show that the expression in parentheses vanishes, start with the identity

$$J^{i'}_{j} J^{j}_{j'} = \delta^{i'}_{j'} \tag{8.117}$$

and differentiate with respect to Z^i:

$$J^{i'}_{ji} J^{j}_{j'} + J^{i'}_{j} J^{j}_{j'k'} J^{k'}_{i} = 0. \tag{8.118}$$

Thus, the nontensor term indeed drops out and we have

$$\nabla_{j'} T^{i'} = \nabla_j T^i Z^{i'}_{i} Z^{j}_{j'} \tag{8.119}$$

proving that $\nabla_j T^i$ is indeed a tensor.

8.7.3 A Direct Proof of the Tensor Property for $\nabla_k T^i_j$

We now turn to the proof of the tensor property of the covariant derivative when applied to a tensor T^i_j. In the coordinate system $Z^{i'}$, $\nabla_{k'} T^{i'}_{j'}$ reads

$$\nabla_{k'} T^{i'}_{j'} = \frac{\partial T^{i'}_{j'}}{\partial Z^{k'}} + \Gamma^{i'}_{k'm'} T^{m'}_{j'} - \Gamma^{m'}_{j'k'} T^{i'}_{m'}, \tag{8.120}$$

Let us determine how each term transforms under the change of coordinates. For $\partial T^{i'}_{j'} / \partial Z^{k'}$, we have

$$T^{i'}_{j'} = T^i_j J^{i'}_i J^j_{j'}. \tag{8.121}$$

By a combination of the product rule and the sum rule, we find

$$\frac{\partial T^{i'}_{j'}}{\partial Z^{k'}} = \frac{\partial T^i_j}{\partial Z^k} J^{i'}_i J^j_{j'} J^k_{k'} + T^i_j J^{i'}_{ik} J^j_{j'} J^k_{k'} + T^i_j J^{i'}_i J^j_{j'k'}. \tag{8.122}$$

For the second term $\Gamma^{i'}_{k'm'} T^{m'}_{j'}$, we have

$$\Gamma^{i'}_{k'm'} T^{m'}_{j'} = \left(\Gamma^i_{km} J^{i'}_i J^k_{k'} J^m_{m'} + J^i_{k'm'} J^{i'}_i \right) T^r_j J^{m'}_r J^j_{j'} \tag{8.123}$$

Multiplying out,

$$\Gamma^{i'}_{k'm'} T^{m'}_{j'} = \Gamma^i_{km} T^m_j J^{i'}_i J^k_{k'} J^s_{j'} + T^r_j J^i_{k'm'} J^{i'}_i J^{m'}_r J^j_{j'}. \tag{8.124}$$

Finally, for the last term $\Gamma^{m'}_{j'k'} T^{i'}_{m'}$, we have

$$\Gamma^{m'}_{j'k'} T^{i'}_{m'} = \left(\Gamma^m_{jk} J^{m'}_m J^j_{j'} J^k_{k'} + J^m_{j'k'} J^{m'}_m \right) T^i_s J^{i'}_{i'} J^s_{m'}. \tag{8.125}$$

Expand the expression on the right hand side:

$$\Gamma^{m'}_{j'k'} T^{i'}_{m'} = \Gamma^m_{jk} T^i_m J^j_{j'} J^k_{k'} J^{i'}_r + T^i_m J^{i'}_i J^m_{j'k'}. \tag{8.126}$$

Combining the terms in equations (8.122), (8.124) and (8.126), we recognize the terms from the proofs of the tensor property of $\nabla_i T_j$ and $\nabla_i T^j$. Namely, the leading terms from each equation combine to give precisely $\nabla_k T^i_j \; J^{i'}_i J^j_{j'} J^k_{k'}$. Among the remaining terms, the last term in equation (8.122) cancel the second term in equation (8.126). This is analogous to the cancellation that took place in the proof of the tensor property of $\nabla_i T_j$. The other two terms, $T^i_j J^{i'}_{ik} J^j_{j'} J^k_{k'}$ from equation (8.122) and $T^r_j J^i_{k'm'} J^{i'}_i J^{m'}_r J^j_{j'}$ from equation (8.124) combine to give

$$T^i_j \left(J^{i'}_{ik} J^k_{k'} + J^k_{k'm'} J^{i'}_k J^{m'}_i \right) J^j_{j'}, \tag{8.127}$$

where, as before, the last step required renaming some of the dummy indices and, as before, we conclude that this term vanishes. We have therefore shown that

$$\nabla_{k'} T^{i'}_{j'} = \nabla_k T^i_j \ J^{i'}_i J^j_{j'} J^k_{k'}, \tag{8.128}$$

which proves the tensor property of $\nabla_k T^i_j$.

8.8 The Riemann–Christoffel Tensor: A Preview

The Riemann–Christoffel tensor is a crucially important object in tensor calculus and has an interesting history. The concepts that are captured so beautifully by the Riemann–Christoffel tensor were conceived by the German mathematician Bernhard Riemann (Fig. 8.2) and first described in his masterpiece [35]. Of course, [35] predates tensor calculus by 35 years and does not use the tensor notation championed in this book. In fact, Riemann felt that his ideas were best expressed in words rather than in equations. As a result, [35] contains very few equations which makes it a challenging read to say the least. If you can read German [36] or Russian [37], you can find a detailed commentary—in tensor notation—by Hermann Weyl in the 1919 edition of Riemann's collected works.

The Riemann–Christoffel tensor arises in the analysis of the commutator $\nabla_i \nabla_j - \nabla_j \nabla_i$. It is straightforward to show that the identity

$$\nabla_i \nabla_j T^k - \nabla_j \nabla_i T^k = \left(\frac{\partial \Gamma^k_{jm}}{\partial Z^i} - \frac{\partial \Gamma^k_{im}}{\partial Z^j} + \Gamma^k_{in} \Gamma^n_{jm} - \Gamma^k_{jn} \Gamma^n_{im} \right) T^m \tag{8.129}$$

Fig. 8.2 Bernhard Riemann (1826–1866) was a pivotal figure in the development of modern geometry

holds for any tensor T^k. Thus the expression in parentheses is a tensor. (Demonstrate this.) This tensor is known as the Riemann–Christoffel tensor $R^k_{\cdot mij}$

$$R^k_{\cdot mij} = \frac{\partial \Gamma^k_{jm}}{\partial Z^i} - \frac{\partial \Gamma^k_{im}}{\partial Z^j} + \Gamma^k_{in}\Gamma^n_{jm} - \Gamma^k_{jn}\Gamma^n_{im}. \tag{8.130}$$

This object allows us to write the commutator $\nabla_i \nabla_j T^k - \nabla_j \nabla_i T^k$ very compactly:

$$\nabla_i \nabla_j T^k - \nabla_j \nabla_i T^k = R^k_{\cdot mij} T^m. \tag{8.131}$$

Since the covariant derivatives commute in a Euclidean space, the Riemann–Christoffel symbol vanishes

$$R^k_{\cdot mij} = 0. \tag{8.132}$$

This equation is a powerful statement about the relationship among the entries of the Christoffel symbol.

The Riemann–Christoffel tensor takes center stage in Part II, which is devoted to the tensor description of curved surfaces. In Chap. 12, the profound significance of the Riemann–Christoffel tensor becomes apparent. Furthermore, the Riemann–Christoffel tensor plays a central role in Einstein's general relativity in which the ambient space is assumed to be non-Euclidean.

Exercise 145. Derive equation (8.129).

Exercise 146. Show that the Riemann–Christoffel tensor vanishes in a Euclidean space by substituting equation (5.60) into equation (8.130). After an initial application of the product rule, you will need equation (5.65) to re-express the derivative of the contravariant basis.

8.9 A Particle Moving Along a Trajectory

This section presents a problem that is best solved by a developing its own mini-calculus. A situation such as this arises quite frequently in practice. The problem discussed here is usually presented in the context of *parallelism* along a curve in a Riemann space.

Consider a particle moving in a Euclidean space, referred to a general coordinate system Z^i, along the trajectory $\Gamma(t)$ given by

$$Z^i = Z^i(t). \tag{8.133}$$

Our goal is to calculate the components of the particle's velocity, acceleration, and *jolt* (rate of change in acceleration, also known as *jerk* and *surge*).

In order to determine the components of the velocity vector, note that the position vector $\boldsymbol{R}$ of the particle is given by the composition

$$\mathbf{R}(t) = \mathbf{R}(Z(t)) \tag{8.134}$$

of the functions in equations (5.1) and (8.133). Differentiating equation (8.134) with respect to t, yields

$$\mathbf{V}(t) = \frac{\partial \mathbf{R}}{\partial Z^i}\frac{dZ^i}{dt}, \tag{8.135}$$

or

$$\mathbf{V}(t) = \frac{dZ^i}{dt}\mathbf{Z}_i. \tag{8.136}$$

Thus

$$V^i(t) = \frac{dZ^i(t)}{dt}. \tag{8.137}$$

We note that $\mathbf{V}(t)$, as well as derivatives $\mathbf{R}^{(n)}(t)$ of all orders of $\mathbf{R}(t)$ with respect to t, are invariants with respect to the change of coordinates Z^i, since the differentiation of $\mathbf{R}(t)$ with respect to t does not even involve the coordinates Z^i.

Exercise 147. Argue, on the basis of $\mathbf{V}$ being a tensor, that V^i is tensor with respect to the change of coordinates Z^i.

Exercise 148. Show, by relating $V^{i'}$ to V^i, that V^i is a tensor with respect to the change of coordinates Z^i.

Exercise 149. Show that dV^i/dt is not a tensor with respect to the change of coordinates Z^i.

To obtain the components $A^i(t)$ of acceleration $\mathbf{A}(t)$, differentiate equation (8.136) once again with respect to t.

Exercise 150. Show that

$$\mathbf{A}(t) = \left(\frac{dV^i}{dt} + \Gamma^i_{jk}V^jV^k\right)\mathbf{Z}_i, \tag{8.138}$$

thus

$$A^i(t) = \frac{dV^i}{dt} + \Gamma^i_{jk}V^jV^k. \tag{8.139}$$

Exercise 151. Argue that the combination on the right-hand side of this equation is a tensor. (Hint: $\mathbf{A}$ is an invariant).

Equation (8.139) can be used as a motivation for the following definition of the $\delta/\delta t$-derivative for a general variant T^i defined along the trajectory $\gamma(t)$:

$$\frac{\delta T^i}{\delta t} = \frac{dT^i}{dt} + V^j \Gamma^i_{jk} T^k \tag{8.140}$$

This definition can be extended to tensors of arbitrary indicial signature as follows:

$$\frac{\delta T^i_j}{\delta t} = \frac{dT^i_j}{dt} + V^k \Gamma^i_{km} T^m_j - V^k \Gamma^m_{kj} T^i_m. \tag{8.141}$$

The $\delta/\delta t$-derivative can also be referred to as the *intrinsic derivative*.

Exercise 152. Suppose that T^i is the contravariant component of a vector field $\mathbf{T}$ that is constant along the trajectory γ. Show that $\delta T^i/\delta t = 0$.

This exercise gives insight into the term *parallelism*. If a vector $\mathbf{T}$ travels unchanged along the trajectory Γ then its components T^i, with respect to the changing basis $\mathbf{Z}_i$, are characterized by the equation $\delta T^i/\delta t = 0$.

The $\delta/\delta t$-derivative has all of the desirable properties discussed in chapter. The following exercises summarize these properties and invite you to use the newly developed calculus to obtain the expression for components of the jolt $\mathbf{J}$.

Exercise 153. Use the techniques developed in this chapter to show that the $\delta/\delta t$-derivative satisfies the tensor property: that is, the $\delta/\delta t$-derivative produces tensors out of tensors.

Exercise 154. Show that the $\delta/\delta t$-derivative coincides with the ordinary derivative d/dt when applied to variants of order zero.

Exercise 155. Show that the $\delta/\delta t$-derivative coincides with the ordinary derivative d/dt in affine coordinates.

Exercise 156. Show that the $\delta/\delta t$-derivative satisfies the sum and product rules.

Exercise 157. Show that the $\delta/\delta t$-derivative commutes with contraction.

Exercise 158. Suppose that T^i_j is defined in the entire Euclidean space. Show that the trajectory restriction of T^i_j satisfies the chain rule

$$\frac{\delta T^i_j}{\delta t} = \frac{\partial T^i_j(t, Z)}{\partial t} + V^k \nabla_k T^i_j. \tag{8.142}$$

Exercise 159. Show that the $\delta/\delta t$-derivative is metrinilic with respect to all metrics associated with the coordinate system A^i.

Exercise 160. Rewrite equation (8.138) in the form

$$\mathbf{A}(t) = \frac{\delta V^i}{\delta t}\mathbf{Z}_i. \tag{8.143}$$

Apply the $\delta/\delta t$-derivative to both sides of this equation to obtain the following expression for the jolt $\mathbf{J}$:

$$\mathbf{J}(t) = \frac{\delta^2 V^i}{\delta t^2}\mathbf{Z}_i. \tag{8.144}$$

Thus

$$J^i = \frac{\delta^2 V^i}{\delta t^2}. \tag{8.145}$$

8.10 Summary

In Chap. 6, we discussed the crucial importance of the tensor property in forming geometrically meaningful invariant objects. However, we discovered that partial differentiation of a tensor results in a variant that is not a tensor. This problem had to be overcome in order to preserve the tensor calculus framework. The solution presented in this chapter comes in the form of a new differential operator, the covariant derivative.

The covariant derivative truly saves the day! It preserves just about all important properties of the covariant derivative and *provides the crucial benefit that it produces tensors out tensors*. The invention of the covariant derivative allowed the development of tensor calculus to go forward. The definition developed in this chapter will serve as a prototype for other differential operators that produce tensors out of tensors, include the covariant differentiation on surfaces as well as the invariant time derivative on moving surfaces.

Chapter 9
Determinants and the Levi-Civita Symbol

9.1 Preview

I have been looking forward to writing this chapter because we now get to use the machinery we have been constructing. And to what great effect! Tensor calculus is a fantastic language for determinants. The subject of determinants is beautiful and is too often obfuscated by cumbersome notation. Even some of the most fundamental properties of determinants can be difficult to follow without tensor notation. If you ever found the proof that the determinant of a product equals the product of determinants less than satisfying, you will find your satisfaction in this chapter.

You are well aware of the importance of determinants for solving linear systems, computing volumes, changing variables in volume integrals, and evaluating vector products and the curl of a vector field. In tensor calculus, determinants also play a special role. There is a kind of symbiotic relationship between determinants and our subject: the tensor notation makes working with determinants easy while determinants give us the Levi-Civita symbols and help us approach integration from an invariant point of view.

We will begin with the *permutation symbols*. The determinant of a second-order system can then be defined in terms of the permutation symbols. We will then demonstrate the key properties of the determinant, in tensor notation. We will also introduce the delta symbols, which are nifty little objects of tremendous utility. The ultimate highlight of the chapter is the Voss–Weyl formula, which gives a convenient expression for the divergence, and therefore the Laplacian, that bypasses the covariant derivative. Finally, we will present the volume element and the Levi-Civita symbols, rounding out our collection of metrics.

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,

DOI 10.1007/978-1-4614-7867-6_9,

9.2 The Permutation Symbols

Our discussion will initially be limited to three dimensions. Consider a *skew-symmetric* system A_{ijk}. Skew-symmetry refers to the property that an exchange of any two indices changes the sign of the entry, whether it is the first two indices

$$A_{ijk} = -A_{jik}, \tag{9.1}$$

the last two

$$A_{ijk} = -A_{ikj}, \tag{9.2}$$

or the first and the third

$$A_{ijk} = -A_{kji}. \tag{9.3}$$

It immediately follows that if two of the indices have the same value. then the corresponding entry in A_{ijk} is zero. For instance, applying equation (9.1) to the values of indices $i = j = 1$ and $k = 2$, yields

$$A_{112} = -A_{112} \tag{9.4}$$

and it follows $A_{112} = 0$. We therefore conclude that the entries of A_{ijk} are not zero only for indices that form a permutation of numbers 1, 2, and 3. There are 6 entries

$$A_{123}, A_{132}, A_{213}, A_{231}, A_{312}, A_{321}, \tag{9.5}$$

but only one degree of freedom: each entry is either A_{123} or $-A_{123}$. These values are determined by the skew-symmetry conditions (9.1), (9.2), and (9.3). Let α denote A_{123}. Then

$$A_{123} = A_{312} = A_{231} = \alpha \tag{9.6}$$

$$A_{213} = A_{321} = A_{132} = -\alpha. \tag{9.7}$$

As you can see, there isn't much variety in skew-symmetric systems: once the value of A_{123} is selected, the remaining five nonzero entries follow. When $\alpha = 1$ the system is denoted by e_{ijk} and is called a *permutation symbol*. Let us catalog its values in a format that generalizes to higher (and lower) dimensions

$$e_{ijk} = \begin{cases} 1 & \text{if } ijk \text{ is an even permutation of } 1,2,3 \\ -1 & \text{if } ijk \text{ is an odd permutation of } 1,2,3 \\ 0 & \text{otherwise} \end{cases}. \tag{9.8}$$

The same system can be denoted by the symbol e^{ijk} with upper indices. The placement of the indices does not matter until we discuss transformation of coordinates.

Exercise 161. Show that equation (9.3) follows from equations (9.1) and (9.2).

Exercise 162. Evaluate the contraction $e^{ijk}e_{ijk}$.

Exercise 163. Evaluate $e^{ijk}e_{kij}$.

Exercise 164. What are the entries of the system $f_{ijk} = e_{ijk} + e_{jik}$?

Exercise 165. What are the entries of the system $f_{ijk} = e_{ijk} + e_{kij}$?

9.3 Determinants

Consider a system a^i_j of order two. In this section we are not concerned with the variant or tensor properties of a^i_j. That is, it is unimportant—for now—according to what rule a^i_j was constructed. Consequently, the placement of indices does not carry any tensorial information and we picked an object with one upper and one lower index simply as a convenience. The determinant A of a^i_j is defined as

$$A = e^{ijk}a^1_i a^2_j a^3_k. \tag{9.9}$$

You should convince yourself that this definition is consistent with the definition given in linear algebra. It is certainly more compact, but the compactness comes at the expense of including a multitude of terms that equal zero. The usual linear algebra definition includes 3! terms. Equation (9.9), on the other hand, has 3^3 terms, all but 3! of them zero. The surviving terms are the same as in the usual definition and the permutation symbol e^{ijk} assures that they are counted with the correct sign.

Our first goal is to replace equation (9.9) with an equivalent one that conforms to all the conventions of tensor calculus:

$$A = \frac{1}{3!}e^{ijk}e_{rst}a^r_i a^s_j a^t_k. \tag{9.10}$$

To show the equivalence of equations (9.9) and (9.10), note that switching the values 1 and 2 in the original definition results in the same value of opposite sign:

$$e^{ijk}a^2_i a^1_j a^3_k = e^{ijk}a^1_j a^2_i a^3_k = e^{jik}a^1_i a^2_j a^3_k = -e^{ijk}a^1_i a^2_j a^3_k = -A. \tag{9.11}$$

This observation can be applied to any pair of indices. In other words, the $e^{ijk}a^r_i a^s_j a^t_k$ is skew-symmetric in r, s, and t. As we have previously learned, a third-order skew-symmetric system has a single degree of freedom, which can be identified as the entry corresponding to $r, s, t = 1, 2, 3$. By definition, for these values of the indices, $e^{ijk}a_i^r a_j^s a_k^t$ is precisely A. This observation can be summarized as follows:

$$e^{ijk}a_i^r a_j^s a_k^t = Ae^{rst}. \tag{9.12}$$

This is already better since we have succeeded in eliminating explicit indices. What we need to do now is isolate A on the right-hand side. In Exercise 162 you discovered that $e_{rst}e^{rst} = 3!$. Therefore, contracting both sides of this equation with e_{rst}, we arrive at equation (9.10). From this point forward, we will treat equation (9.10) as the definition of the determinant.

Definitions similar to (9.9) can be given for systems a_{ij} with lower indices and a^{ij} with upper indices:

$$A = e^{ijk}a_{1i}a_{2j}a_{3k} \tag{9.13}$$

$$A = e_{ijk}a^{1i}a^{2j}a^{3k}. \tag{9.14}$$

and it can be similarly shown that

$$A = \frac{1}{3!}e^{ijk}e^{rst}a_{ir}a_{js}a_{kt} \tag{9.15}$$

$$A = \frac{1}{3!}e_{ijk}e_{rst}a^{ir}a^{js}a^{kt}. \tag{9.16}$$

Once again, the placement of the index is only a matter of convenience until the transformation of A and underlying systems is analyzed. This will be discussed later in this chapter.

Exercise 166. Explain each step in equation (9.11).

Exercise 167. Derive equations (9.15) and (9.16).

9.4 The Delta Systems

The fullest *delta system* δ_{rst}^{ijk} is defined as the tensor product of two permutation symbols

$$\delta_{rst}^{ijk} = e^{ijk}e_{rst}. \tag{9.17}$$

Using this new object, the determinant A of a_j^i can be written more compactly as

$$A = \frac{1}{3!}\delta_{rst}^{ijk}a_i^r a_j^s a_k^t. \tag{9.18}$$

Furthermore, by multiplying both sides with e_{lmn}, equation (9.12) leads to the following expression

$$\delta^{ijk}_{lmn} a^r_i a^s_j a^t_k = A\delta^{rst}_{lmn}, \tag{9.19}$$

which will prove important in showing the product property of the determinant. This equation is quite appealing, even if it doesn't say anything beyond what we have already learned. Also, in a way, it may also be the most of all tensor identities we have experienced so far. After all, it represents 27 identities, each containing 729 terms on the left-hand side.

The symbol δ^{ijk}_{rst} is only one of a whole group of delta symbols δ^i_r, δ^{ij}_{rs} and δ^{ijk}_{rst}. Of course, we have already encountered δ^i_j, the Kronecker symbol. All delta systems, including their generalizations to higher dimensions, can be captured by a single definition: *A particular entry of a delta system has value* 1 *when the upper and lower are identical sets of distinct numbers related by an even permutation,* -1 *if the sets are related by an odd permutation, and* 0 *otherwise*. In other words, the entry is 0 if any of the upper or any of the lower indices repeat, or if they do not represent identical sets of numbers. You should be able to see that this definition correctly describes the familiar objects δ^i_j and δ^{ijk}_{rst}, for which alternative definition exist.

Thus, the value is 0 if any upper or lower indices are equal and, in the case of distinct sets, if those sets are not the same. For example, $\delta^{11}_{12} = 0$ because the two upper indices are the same and $\delta^{12}_{13} = 0$ because the upper and lower indices do not consist of the same numbers. As a final example, $\delta^{12}_{12} = 1$ and $\delta^{23}_{32} = -1$ because, in each case, the upper and lower indices form identical sets of distinct numbers. For δ^{12}_{12}, the upper indices are an even (zero, actually) permutation away from the lower indices while for δ^{23}_{32} the permutation is odd.

It is evident that delta systems are skew-symmetric in its upper and its lower indices. The full delta system δ^{ijk}_{rst} can be expressed in terms of the Kronecker symbol by the determinant

$$\delta^{ijk}_{rst} = \begin{vmatrix} \delta^i_r & \delta^i_s & \delta^i_t \\ \delta^j_r & \delta^j_s & \delta^j_t \\ \delta^k_r & \delta^k_s & \delta^k_t \end{vmatrix}. \tag{9.20}$$

The most important consequence of equation (9.20) is that δ^{ijk}_{rst} is a tensor. Consequently, according to equation (9.18), the determinant of a second-order tensor with one covariant and one contravariant index is an *invariant*. The same cannot be said of tensors with two covariant or two contravariant indices. The tensor properties of determinants of such tensors are discussed later in this chapter.

Delta systems of different orders are related by the following contractions:

$$\delta^{ij}_{rs} = \delta^{ijk}_{rsk} \tag{9.21}$$

$$2\delta^i_r = \delta^{ij}_{rj}. \tag{9.22}$$

Proofs of these relationships are left as exercises. Equations (9.21) and (9.22) show that all delta systems are tensors. These relationships extend to higher dimensions. For example, in four dimensions,

$$\delta^{ijk}_{rst} = \delta^{ijkl}_{rstl} \tag{9.23}$$

$$2\delta^{ij}_{rs} = \delta^{ijk}_{rsk} \tag{9.24}$$

$$3\delta^{i}_{j} = \delta^{ij}_{rj}. \tag{9.25}$$

The symbol δ^{ij}_{rs} can be expressed in term of the Kronecker symbol by the following frequently used identity

$$\delta^{ij}_{rs} = \delta^{i}_{r}\delta^{j}_{s} - \delta^{j}_{r}\delta^{i}_{s}. \tag{9.26}$$

The right-hand side of equation (9.26) can also be captured by a determinant:

$$\delta^{ij}_{rs} = \begin{vmatrix} \delta^{i}_{r} & \delta^{i}_{s} \\ \delta^{j}_{r} & \delta^{j}_{s} \end{vmatrix}. \tag{9.27}$$

Exercise 168. Justify equation (9.20).

Exercise 169. Explain why equation (9.20) shows that δ^{ijk}_{rst} is a tensor.

Exercise 170. Explain why it follows that the determinant of a tensor with one covariant and one contravariant index is an invariant.

Exercise 171. Justify equations (9.21) and (9.25).

Exercise 172. Justify equations (9.23) and (9.25) in four dimensions.

Exercise 173. Generalize equations (9.23) and (9.25) to N dimensions.

Exercise 174. Justify equation (9.26).

Exercise 175. Evaluate δ^{ijk}_{ijk}.

Exercise 176. Evaluate δ^{ikj}_{jir}.

Exercise 177. Evaluate $\delta^{ijk}_{rst}\delta^{rst}_{lmn}$.

Exercise 178. Evaluate $\delta^{ijk}_{rst}\delta^{rst}_{lmn}\delta^{lmn}_{ijk}$.

9.5 A Proof of the Multiplication Property of Determinants

The multiplication property of determinants reads

$$|MN| = |M|\,|N|. \tag{9.28}$$

In words, the determinant of a product of two matrices equals the product of their determinants. Three different proofs of this property can be found in linear algebra textbooks including [13, 27, 44], and [22]. The tensor derivation is based on the approach in [13], and [22]. However, the proofs without the tensor notation lead to cumbersome expressions, while the repeated use of the symbol $\sum_{\text{all permutations}}$ makes those proofs difficult to follow.

The tensor notation translates the logic of permutations into simple algebraic manipulations. Suppose that the double system c^i_j is the "matrix product" of a^i_j and b^j_k:

$$c^i_j = a^i_k b^k_j. \tag{9.29}$$

Then its determinant C is given by

$$C = \frac{1}{3!}\delta^{ijk}_{rst} c^r_i c^s_j c^t_k. \tag{9.30}$$

Substitute equation (9.29) in (9.30):

$$C = \frac{1}{3!}\delta^{ijk}_{rst} a^r_l b^l_i a^s_m b^m_j a^t_n b^n_k. \tag{9.31}$$

Since, by equation (9.19), $\delta^{ijk}_{rst} a^r_l a^s_m a^t_n = A\delta^{lmn}_{ijk}$, we have

$$C = \frac{1}{3!}A\delta^{ijk}_{lmn} b^l_i b^m_j b^n_k, \tag{9.32}$$

which gives

$$C = AB. \tag{9.33}$$

The entire derivation can be combined in a single line

$$C = \frac{1}{3!}\delta^{ijk}_{rst} c^r_i c^s_j c^t_k = \frac{1}{3!}\delta^{ijk}_{rst} a^r_l b^l_i a^s_m b^m_j a^t_n b^n_k = \frac{1}{3!}A\delta^{ijk}_{lmn} b^l_i b^m_j b^n_k = AB. \tag{9.34}$$

This proof is a demonstration of the effectiveness of the tensor notation.

Exercise 179. Show $C = AB$ property for systems c^i_k, a^{ij}, and b_{jk} related by $c^i_k = a^{ij} b_{jk}$.

9.6 Determinant Cofactors

If the term a^r_i is removed from the product $\delta^{ijk}_{rst} a^r_i a^s_j a^t_k$ the result is a second-order system of tremendous utility. Define the *cofactor* A^i_r by

$$A^i_r = \frac{1}{2!}\delta^{ijk}_{rst} a^s_j a^t_k. \tag{9.35}$$

For a tensor a_i^r, the cofactor A_r^i is also a tensor. The cofactor A_r^i has two remarkable properties. First, it represents the partial derivative $\partial A / \partial a_r^i$ of the determinant A with respect to its entries. Second, the product $A^{-1} A_r^i$ is the matrix inverse of a_r^i.

To prove the first relationship

$$\frac{\partial A}{\partial a_i^r} = A_r^i, \tag{9.36}$$

compute the derivative of A with respect to a_l^u, since the indices i and r are already used in equation (9.35). We have

$$\frac{\partial A}{\partial a_l^u} = \frac{1}{3!} \delta_{rst}^{ijk} \frac{\partial \left(a_i^r a_j^s a_k^t \right)}{\partial a_l^u}. \tag{9.37}$$

By the product rule,

$$\frac{\partial A}{\partial a_l^u} = \frac{1}{3!} \delta_{rst}^{ijk} \left(\frac{\partial a_i^r}{\partial a_l^u} a_j^s a_k^t + a_i^r \frac{\partial a_j^s}{\partial a_l^u} a_k^t + a_i^r a_j^s \frac{\partial a_k^t}{\partial a_l^u} \right). \tag{9.38}$$

Each of the partial derivatives in (9.38) can be captured by the Kronecker symbol. For example,

$$\frac{\partial a_i^r}{\partial a_l^u} = \delta_u^r \delta_i^l. \tag{9.39}$$

Therefore

$$\frac{\partial A}{\partial a_l^u} = \frac{1}{3!} \delta_{rst}^{ijk} \left(\delta_u^r \delta_i^l a_j^s a_k^t + \delta_u^s \delta_j^l a_i^r a_k^t + \delta_u^t \delta_k^l a_i^r a_j^s \right). \tag{9.40}$$

Multiply out this expression and use the index-renaming property of the Kronecker symbol:

$$\frac{\partial A}{\partial a_l^u} = \frac{1}{3!} \left(\delta_{ust}^{ljk} a_j^s a_k^t + \delta_{rut}^{ilk} a_i^r a_k^t + \delta_{rsu}^{ijl} a_i^r a_j^s \right). \tag{9.41}$$

The terms in parentheses are each equivalent to $2A_u^l$. Therefore

$$\frac{\partial A}{\partial a_l^u} = A_u^l, \tag{9.42}$$

which completes the proof.

We next show that the cofactor A_r^i is proportional to the matrix inverse of the system a_i^r:

$$A_r^i a_m^r = A\delta_m^i. \tag{9.43}$$

Denote the product $A_r^i a_m^r$ by D_m^i:

$$D_m^i = A_r^i a_m^r = \frac{1}{2!}\delta_{rst}^{ijk} a_m^r a_j^s a_k^t. \tag{9.44}$$

We first show that $D_m^i = 0$ when $i \neq m$. For example, consider D_2^1

$$D_2^1 = \frac{1}{2!}\delta_{rst}^{1jk} a_2^r a_j^s a_k^t. \tag{9.45}$$

Out of the 3^5 terms in the sum on the right-hand side, only two ($j,k = 2,3$ and $j,k = 3,2$) correspond to nonzero entries of δ_{rst}^{1jk}:

$$D_2^1 = \frac{1}{2!}\left(\delta_{rst}^{123} a_2^r a_2^s a_3^t + \delta_{rst}^{132} a_2^r a_3^s a_2^t\right). \tag{9.46}$$

However, each of these terms is zero because the value 2 appears twice as the lower index. Therefore, $D_m^i = 0$ for $i \neq m$.

Next consider the case $i = m$. For example, consider D_1^1:

$$D_1^1 = \frac{1}{2!}\delta_{rst}^{1jk} a_1^r a_j^s a_k^t. \tag{9.47}$$

The same two terms correspond not nonvanishing values of δ_{rst}^{1jk}:

$$D_2^1 = \frac{1}{2!}\left(\delta_{rst}^{123} a_1^r a_2^s a_3^t + \delta_{rst}^{132} a_2^r a_3^s a_1^t\right). \tag{9.48}$$

This time, each equals A and we have therefore shown that $D_m^i = A$ for $i = m$. This completes the proof of the fact that A_r^i is A times the inverse of a_i^r.

Exercise 180. Explain equation (9.39).

Exercise 181. Show that each term in equation (9.41) equals $2A_l^u$.

Exercise 182. Generalize the proof of equation (9.36) to arbitrary dimension.

Exercise 183. Derive similar expressions for determinants of tensors a_{ij} and a^{ij}.

Exercise 184. Explain why each term vanishes on the right-hand side of equation (9.46).

Exercise 185. Explain why each term in parentheses in equation (9.48) equals A.

Exercise 186. Extend the subject of cofactors to systems with two lower indices (denote the cofactor by A^{ir}) and systems with two upper indices (denote the cofactor by A_{ir}).

9.7 The Object Z and the Volume Element

Let Z denote the determinant of the covariant metric tensor Z_{ij}:

$$Z = |Z_{\cdot\cdot}|. \tag{9.49}$$

We have used the letter Z to denote several objects: the independent variable Z^i, the bases $\mathbf{Z}_i$ and $\mathbf{Z}^i$, the metric tensors Z_{ij} and Z^{ij}, and now the determinant of Z_{ij}. However, there is no confusion since the index signatures and the font make these symbols different. Interestingly, some texts use the same letter to denote the position vector $\mathbf{R}$, the Jacobian matrices $J^i_{i'}$ and $J^{i'}_i$ and the upcoming Levi-Civita symbols ε_{ijk} and ε^{ijk}. The combination of index signature and font would still enable us to distinguish among all objects $\mathbf{Z}$, $Z^i_{i'}$, $Z^{i'}_i$, Z_{ijk}, Z^{ijk}, and allow us to label all of the most fundamental objects by the same letter. In this book, however, we limit the use of the letter Z and this is our last use for Z. Finally, note the use of dots $\cdot\cdot$ in equation (9.49) to indicate the covariant metric tensor. We prefer to use this notation to the alternative $Z = |Z_{ij}|$, because in the latter, the indices i and j are neither live nor contracted and we do not employ indices in any other way.

The variant Z is not a tensor. However it does transform according to a very interesting rule that makes it a *relative tensor*. Relative tensors are discussed in Sect. 9.9. Consider an alternative coordinate system $Z^{i'}$and the corresponding variant Z'

$$Z' = |Z_{\cdot'\cdot'}|. \tag{9.50}$$

Our goal is to establish the relationship between Z' and Z. Let us remind ourselves of the relationship between $Z_{i'j'}$ and Z_{ij}:

$$Z_{i'j'} = Z_{ij} J^i_{i'} J^j_{j'}. \tag{9.51}$$

Let the letter J denote the determinant of the Jacobian $J^i_{i'}$

$$J = \left| J^{\cdot}_{\cdot'} \right|. \tag{9.52}$$

Since $J^i_{i'}$ and $J^{i'}_i$ are the inverses of each other, the determinant of $J^{i'}_i$ is J^{-1}:

$$J^{-1} = \left| J^{\cdot'}_{\cdot} \right|. \tag{9.53}$$

From equation (9.51) we conclude, by the multiplication property of determinants, that

$$Z' = ZJ^2. \tag{9.54}$$

This equation confirms that Z is not a tensor. A variant that changes according to this rule is called a *relative tensor of weight* 2 (because of the 2 in the exponent). Despite not being a tensor, the variant Z is an object of utmost importance as will become evident in the remainder of this chapter.

The quantity $\sqrt{Z}$ is called the *volume element*. In the two-dimensional plane, it may be referred to as the *area element*. On a line, it may be called the *length element* or the *line element*. When the dimension of the space is not specified, we say *volume element*. The quantity Z is the determinant of a positive definite matrix and is therefore positive. More accurately, it is the determinant of a positive *semidefinite* matrix and is therefore *nonnegative*. It may vanish at special points, such as the origin of a polar coordinate system. In either case, the taking of the square root is justified.

You may be familiar with the term *volume element* from multivariable calculus, where it is defined with a reference to a Cartesian coordinate system. According to the multivariable calculus definition, the volume element in coordinates $Z^{i'}$ is defined as the determinant J of the Jacobian matrix $J^i_{i'}$, where Z^i are Cartesian coordinates. The tensor definition has the advantage that it is given in absolute terms without a reference to a particular secondary coordinate system.

The two definitions lead to nearly identical objects that may differ by a sign. Consider a Cartesian coordinate system Z^i and an alternative coordinate system $Z^{i'}$. Then, according to equation (9.54), the object Z' in the primed coordinates is given by

$$Z' = J^2, \tag{9.55}$$

since $Z = 1$ in Cartesian coordinates. This is clear from equation (5.37) and is catalogued in equation (9.57). Taking square roots of both sides, we find

$$\sqrt{Z'} = \text{Absolute value}\,(J)\,. \tag{9.56}$$

Therefore, the volume element $\sqrt{Z'}$ and the determinant J of the Jacobian coincide when J is positive and are of opposite signs when J is negative. In other words, the volume element $\sqrt{Z'}$ agrees with the Jacobian if the coordinate system $Z^{i'}$ is positively oriented with respect to the Cartesian coordinates Z^i. The orientation of a coordinate system was defined in Sect. 4.10.

We now catalogue the values of the volume element for the most common coordinates. In Cartesian coordinates,

$$\sqrt{Z} = 1. \tag{9.57}$$

In polar and cylindrical coordinates,

$$\sqrt{Z} = r. \tag{9.58}$$

In spherical coordinates,

$$\sqrt{Z} = r \sin\theta. \tag{9.59}$$

Exercise 187. Confirm equations (9.57), (9.58), and (9.59).

9.8 The Voss–Weyl Formula

The magical *Voss–Weyl formula* for the calculation of divergence $\nabla_i T^i$ of a tensor field T^i illustrates another use of the volume element $\sqrt{Z}$:

$$\nabla_i T^i = \frac{1}{\sqrt{Z}} \frac{\partial}{\partial Z^i} \left(\sqrt{Z} T^i \right). \tag{9.60}$$

The Voss–Weyl formula expresses the divergence of a contravariant tensor field without a reference to the Christoffel symbols. It is therefore an interesting example of an invariant expression constructed in a nontensor way. Furthermore, it is a formula that has an attractive structure that is often recognized in various expressions.

To confirm the Voss–Weyl formula, first apply the product rule:

$$\frac{1}{\sqrt{Z}} \frac{\partial}{\partial Z^i} \left(\sqrt{Z} T^i \right) = \frac{\partial T^i}{\partial Z^i} + \frac{T^i}{\sqrt{Z}} \frac{\partial \sqrt{Z}}{\partial Z^i}. \tag{9.61}$$

Thus, we are therefore faced with the intriguing task of calculating the partial derivative $\partial\sqrt{Z}/\partial Z^i$. The variant Z depends on the coordinates Z^k via the entries of the metric tensor Z_{ij}. Therefore, $\partial Z/\partial Z^k$ can be found by the chain rule

$$\frac{\partial Z}{\partial Z^k} = \frac{\partial Z}{\partial Z_{ij}} \frac{\partial Z_{ij}}{\partial Z^k}. \tag{9.62}$$

Recall equation (5.68) which gives the partial derivative $\partial Z_{ij}/\partial Z^k$:

$$\frac{\partial Z_{ij}}{\partial Z^k} = \Gamma_{i,jk} + \Gamma_{j,ik} \tag{5.68}$$

According to Exercise 186, the derivative of Z with respect to the entries of the metric tensor Z_{ij} is given by

$$\frac{\partial Z}{\partial Z_{ij}} = ZZ^{ij}. \tag{9.63}$$

Putting these equations together, yields

$$\frac{\partial Z}{\partial Z^k} = ZZ^{ij}\left(\Gamma_{i,jk} + \Gamma_{j,ik}\right), \tag{9.64}$$

or, equivalently,

$$\frac{\partial Z}{\partial Z^k} = 2Z\Gamma^i_{ik}. \tag{9.65}$$

The partial derivative $\partial\sqrt{Z}/\partial Z^k$ of the volume element $\sqrt{Z}$ is now computed by a simple application of the chain rule

$$\frac{\partial\sqrt{Z}}{\partial Z^k} = \frac{1}{2\sqrt{Z}}\frac{\partial Z}{\partial Z^k}. \tag{9.66}$$

The final expression reads

$$\frac{\partial\sqrt{Z}}{\partial Z^k} = \sqrt{Z}\Gamma^i_{ik}. \tag{9.67}$$

To complete the derivation of the Voss–Weyl formula, substitute equation (9.67) into (9.61)

$$\frac{1}{\sqrt{Z}}\frac{\partial}{\partial Z^i}\left(\sqrt{Z}T^i\right) = \frac{\partial T^i}{\partial Z^i} + T^i\Gamma^k_{ki}, \tag{9.68}$$

and note that the expression on the right-hand side coincides with the definition of $\nabla_i T^i$.

The Voss–Weyl formula provides the most effective way to calculate the Laplacian $\nabla_i\nabla^i F$ of an invariant field F. Since the Laplacian $\nabla_i\nabla^i F$ is the divergence of $T^i = \nabla^i F$ and contravariant gradient $\nabla^i F$ is given by

$$\nabla^i F = Z^{ij}\frac{\partial F}{\partial Z^j}, \tag{9.69}$$

the Voss–Weyl formula yields

$$\nabla_i\nabla^i F = \frac{1}{\sqrt{Z}}\frac{\partial}{\partial Z^i}\left(\sqrt{Z}Z^{ij}\frac{\partial F}{\partial Z^j}\right). \tag{9.70}$$

Exercise 188. Use the Voss–Weyl formula to derive the Laplacian in spherical coordinates

$$\nabla_i \nabla^i F = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial F}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial F}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 F}{\partial \phi^2}. \tag{9.71}$$

This expression is equivalent to equation (8.39) which was obtained in a much more laborious fashion with the help of the Christoffel symbols. It is common not to expand any of the terms in this expression and to present the Laplacian in this appealing form. When looking at this expression, one can see upon it the characteristic imprint of the Voss–Weyl formula.

Exercise 189. Use the Voss–Weyl formula to derive the expression (8.37) for the Laplacian in cylindrical coordinates.

9.9 Relative Tensors

The object Z is an example of a variant that, while not quite a tensor, transforms according to a very interesting rule given in equation (9.54). The transformation rule is nearly that of an invariant except for the additional factor of J^2. There are other important objects that are nearly tensors in this sense. A variant T^i_j is a called *relative tensor* of weight M if it changes according to the rule

$$T^{i'}_{j'} = J^M T^i_j J^{i'}_i J^j_{j'}. \tag{9.72}$$

Thus, tensors are special cases of relative tensors corresponding to $M = 0$. To make the distinction clearer, tensors are sometimes referred to as *absolute tensors*. A relative tensor of order zero of weight M is called a *relative invariant* of weight M. As illustrated by the following exercises, relative tensors satisfy all of the appropriately adjusted properties of regular tensors.

Exercise 190. Show that if S^i_{jk} is a relative tensor of weight M and T^{rs}_t is a relative tensor of weight N, then $S^i_{jk} T^{rs}_t$ is a relative tensor of weight $M + N$. In particular, if $M = -N$, then $S^i_{jk} T^{rs}_t$ is a tensor.

Exercise 191. Show that the result of contraction for a relative tensor of weight M is also a relative tensor of weight M.

Exercise 192. Conclude that δ^{ijk}_{rst} is an absolute tensor on the basis of equation (9.17).

We now turn to the most interesting relative tensors e_{ijk} and e^{ijk}. Let us first look at the symbol e_{ijk} and see what we get when we transform it as if it were a covariant tensor:

$$e_{ijk} J^i_{i'} J^j_{j'} J^k_{k'}. \tag{9.73}$$

By equation analogous to (9.12), this expression equals $Je_{i'j'k'}$:

$$e_{ijk}J_{i'}^{i}J_{j'}^{j}J_{k'}^{k}=Je_{i'j'k'}. \tag{9.74}$$

Therefore,

$$e_{i'j'k'}=J^{-1}e_{ijk}J_{i'}^{i}J_{j'}^{j}J_{k'}^{k} \tag{9.75}$$

which leads to the conclusion that e_{ijk} is a relative covariant tensor of weight -1.

What about e^{ijk}? Interestingly, since the entries of e^{ijk} are identical to those of e_{ijk}, it can be viewed as the same kind of relative tensor as e_{ijk}. On the other hand, let us try to interpret e^{ijk} as some kind of contravariant tensor. Let us therefore transform e^{ijk} contravariantly:

$$e^{ijk}J_{i}^{i'}J_{j}^{j'}J_{k}^{k'}. \tag{9.76}$$

According to equation (9.12), we find

$$e^{ijk}J_{i}^{i'}J_{j}^{j'}J_{k}^{k'}=J^{-1}e^{i'j'k'}, \tag{9.77}$$

or

$$e^{i'j'k'}=Je^{ijk}J_{i}^{i'}J_{j}^{j'}J_{k}^{k'}, \tag{9.78}$$

and we can therefore interpret e^{ijk} as a relative contravariant tensor of weight 1.

These properties of e_{ijk} and e^{ijk} can help us determine the similar properties for determinants of second-order tensors. Suppose that a_{ij} is an absolute covariant tensor. Its determinant A is given by equation (9.15) and is therefore a *relative invariant or weight* 2. Similarly, the determinant A of an absolute contravariant tensor is a *relative invariant of weight* -2. Finally, as we have seen previously, the determinant of an absolute tensor with one covariant and one contravariant index is also an absolute tensor.

Exercise 193. Show that the determinant of a relative covariant tensor a_{ij} of weight M is a relative invariant of weight $2+3M$ and that in n dimensions the expression generalizes to $2+nM$.

Exercise 194. Show that the volume element $\sqrt{Z}$ is a relative invariant of weight 1 with respect to orientation-preserving coordinate changes.

We now have all the machinery to introduce the Levi-Civita tensors—this author's favorite objects.

9.10 The Levi-Civita Symbols

Our goal is to adjust the relative tensors e^{ijk} and e_{ijk} to produce absolute tensors. This can be accomplished by scaling these systems by the volume element $\sqrt{Z}$ which is a relative invariant (with respect to orientation-preserving coordinate changes) of weight 1. The result is the *Levi-Civita symbols* ε^{ijk} and ε_{ijk}:

$$\varepsilon^{ijk} = \frac{e^{ijk}}{\sqrt{Z}} \tag{9.79}$$

$$\varepsilon_{ijk} = \sqrt{Z} e_{ijk}. \tag{9.80}$$

The Levi-Civita symbols are *absolute tensors* with respect to orientation-preserving coordinate changes. As absolute tensors, the Levi-Civita symbols can be used effectively in defining the curl operator and the cross product of vectors. The Levi-Civita symbols are considered to be *metrics*—a special group of objects derived from the position vector **R**. The metrics now include the covariant and contravariant bases, the metric tensor, and now the volume element and the Levi-Civita symbols. The introduction of the Levi-Civita symbols completes the list of metrics.

In Chap. 8, we discovered the metrinilic property of the covariant derivative. That is the fact that each metric vanishes under the covariant derivative. Does this property extend to the Levi-Civita symbols? The answer is *yes* and proving this property is the subject of the next section.

Exercise 195. Confirm that the Levi-Civita symbols are absolute tensors with respect to orientation preserving coordinate changes.

Exercise 196. Show that

$$\delta^{ijk}_{rst} = \varepsilon^{ijk}\varepsilon_{rst}, \tag{9.81}$$

which offers yet another proof that δ^{ijk}_{rst} is an absolute tensor.

Before we discuss the application of the covariant derivative to the Levi-Civita symbols, we point out an important notational point. Note that the symbol ε_{ijk} admits two interpretations. On the one hand, it is defined by an explicit expression in equation (9.80). On the other hand, it can be interpreted as the result of lowering each index of the contravariant Levi-Civita symbol ε^{rst}:

$$\varepsilon_{ijk} = \varepsilon^{rst} Z_{ir} Z_{js} Z_{kt}. \tag{9.82}$$

Do these two distinct interpretation result in the same object? Fortunately, they do. Let us take the expression $\varepsilon^{rst} Z_{ir} Z_{js} Z_{kt}$ as the starting point and substitute the definition of the Levi-Civita symbol ε^{rst}:

$$\varepsilon^{rst} Z_{ir} Z_{js} Z_{kt} = \frac{e^{rst} Z_{ir} Z_{js} Z_{kt}}{\sqrt{Z}} \tag{9.83}$$

By equation (9.12), the expression in the numerator equals Ze_{ijk}, thus

$$\varepsilon^{rst} Z_{ir} Z_{js} Z_{kt} = \frac{Ze_{ijk}}{\sqrt{Z}} = \sqrt{Z} e_{ijk}. \tag{9.84}$$

We have therefore arrived at the definition of ε_{ijk} showing that the two interpretations are equivalent.

Exercise 197. Show that the permutation e^{ijk} and e_{ijk} are **not** related by index juggling.

9.11 The Metrinilic Property with Respect to the Levi-Civita Symbol

Does the metrinilic property extend to our newest metrics, the Levi-Civita symbols? That is, do the Levi-Civita symbols vanish under the covariant derivative? The answer is *yes* and this section is devoted to proving this important property—the *long* way. Of course, there is the standard short argument that relies on the Euclidean nature of our space. Simply observe that $\nabla_i \varepsilon_{rst}$ is a tensor that vanishes in Cartesian coordinates. Therefore, it vanishes in all coordinates, Q.E.D.

The long way is a valuable technical exercise. It entails a direct application of the covariant derivative to the definition of the Levi-Civita symbol. It has the important advantage that it will continue to work on curved surfaces.

We first evaluate $\nabla_i e_{rst}$, that is, the covariant derivative applied to the permutation symbol. Note that, despite the fact that the entries of e_{rst} are spatial constants, there is no reason to expect that the covariant derivative of e_{rst} vanishes. As we discussed in Chap. 8, this property of the partial derivative does not extend to the covariant derivative (nor should it).

Introduce the symbol, T_{irst}:

$$T_{irst} = \nabla_i e_{rst}. \tag{9.85}$$

Since their partial derivative $\partial e_{rst}/\partial Z^i$ vanishes, T_{irst} consists of the following three terms

$$T_{irst} = -\Gamma^m_{ir} e_{mst} - \Gamma^m_{is} e_{rmt} - \Gamma^m_{it} e_{rsm}. \tag{9.86}$$

From this expression we can conclude that T_{irst} is skew-symmetric in the indices rst. Swapping r and s, we find

$$T_{isrt} = \nabla_i e_{srt} = -\nabla_i e_{rst} = -T_{irst}, \tag{9.87}$$

confirming the skew-symmetry. Therefore, we need to consider only T_{i123} which is given by

$$T_{i123} = -\Gamma^m_{i1} e_{m23} - \Gamma^m_{i2} e_{1m3} - \Gamma^m_{i3} e_{12m}. \tag{9.88}$$

In each contraction, only one term survives:

$$T_{i123} = -\Gamma^1_{i1} e_{123} - \Gamma^2_{i2} e_{123} - \Gamma^3_{i3} e_{123}. \tag{9.89}$$

Factor out e_{123} and summarize this identity using a contraction:

$$T_{i123} = -\Gamma^m_{mi} e_{123}. \tag{9.90}$$

This argument can be generalized for other permutations of 123, yielding

$$T_{irst} = -\Gamma^m_{mi} e_{rst}. \tag{9.91}$$

The metrinilic property of ε_{rst} can now be demonstrated. By the product rule,

$$\nabla_i \varepsilon_{rst} = \nabla_i \left(\sqrt{Z} e_{rst} \right) = \frac{\partial \sqrt{Z}}{\partial Z^i} e_{rst} + \sqrt{Z} \nabla_i e_{rst}. \tag{9.92}$$

We have calculated the partial derivative $\partial \sqrt{Z} / \partial Z^i$ previously and the result be found in equation (9.67). Combining equations (9.67) and (9.91), we find

$$\nabla_i \varepsilon_{rst} = \Gamma^m_{mi} e_{rst} - \Gamma^m_{mi} e_{rst} = 0, \tag{9.93}$$

and the proof is complete.

Exercise 198. Show that the metrinilic property applies to the contravariant Levi-Civita symbol ε^{rst}.

Exercise 199. Use the metrinilic property of the Levi-Civita symbols to show the same for the delta systems:

$$\nabla_m \delta^{ijk}_{rst}, \quad \nabla_m \delta^{ij}_{rs}, \quad \nabla_m \delta^i_r = 0. \tag{9.94}$$

9.12 The Cross Product

The Levi-Civita symbols, being tensors, are key to the analytical definitions of the *cross product* and *the curl.* The cross product W^i of two vectors U^i and V^i is defined by

$$W^i = \varepsilon^{ijk} U_j V_k. \tag{9.95}$$

Since all the elements on the right-hand side of (9.95) are tensors, W^i is a tensor as well. To form an invariant, contract W^i with the covariant basis $\mathbf{Z}_i$ to form $\mathbf{W} = W^i \mathbf{Z}_i$. Then

$$\mathbf{W} = \varepsilon^{ijk} U_j V_k \mathbf{Z}_i. \tag{9.96}$$

In dyadic notation, the relationship among $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$ is denoted by

$$\mathbf{W} = \mathbf{U} \times \mathbf{V}. \tag{9.97}$$

We are now thoroughly enjoying the power of tensor calculus! We wrote down equation (9.95) with great ease, confident that we produced the right equation. Although this definition is algebraic, we are assured of its geometric meaningfulness. Why?—because the result is a tensor! Furthermore, equation (9.95) gives us a whole lot more, including an explicit algorithm for constructing the gradient in any coordinate system. (The definition given in most calculus books certainly does not give us that.)

Furthermore, the definition (9.95) gives us an expression that can be further manipulated. For example, let us evaluate the divergence $\nabla_i W^i$ of the cross product. By the product rule and metrinilic property, the Levi-Civita symbol ε^{ijk} *passes through* the covariant derivative ∇_i

$$\nabla_i W^i = \varepsilon^{ijk} \nabla_i \left(U_j V_k \right). \tag{9.98}$$

Subsequently, by the product rule,

$$\nabla_i W^i = \varepsilon^{ijk} \nabla_i U_j \; V_k + \varepsilon^{ijk} U_j \nabla_i V_k. \tag{9.99}$$

This expression does not admit any further analysis and constitutes the final answer. The combination $\varepsilon^{ijk} \nabla_i U_j$ is known as the curl of U_j and is the subject of the following section.

Just as easily, all the other fundamental properties follow from the definition (9.95). First, let us show the antisymmetric property of the cross product. It follows instantly from the skew-symmetric property of the Levi-Civita symbol. Let

$$W_1^i = \varepsilon^{ijk} U_j V_k \tag{9.100a}$$

$$W_2^i = \varepsilon^{ijk} V_j U_k. \tag{9.100b}$$

Then, by a combination of steps, that we may call switch-rename-reshuffle, we have

$$W_1^i = \varepsilon^{ijk} U_j V_k = -\varepsilon^{ikj} U_j V_k = -\varepsilon^{ijk} U_k V_j = -\varepsilon^{ijk} V_j U_k = -W_2^i, \text{ Q.E.D.} \tag{9.101}$$

In dyadic notation

$$\mathbf{U} \times \mathbf{V} = -\mathbf{V} \times \mathbf{U}. \tag{9.102}$$

The antisymmetric property can be expressed by the attractive equation (9.102) in dyadic notation. It may appear as a disadvantage of the tensor notation that there is no such expression in tensor notation. In fact, in tensor notation there is not even a symbol for the cross product. However, we see this as an advantage rather than as a disadvantage. On the one hand, we save yet another symbol that helps prevent our framework from being overwhelmed by the multitude of operators (not that we are in any danger) and we are also able to view the anti-symmetry of the cross product as a near triviality, given the skew-symmetry of the Levi-Civita symbol, rather than as a stand-alone property of a novel operator.

The cross product of U^i and V^j is orthogonal to each of these vectors. Indeed, dotting W^i and V^i, we find

$$W^i V_i = \varepsilon^{ijk} U_i U_j V_k = 0, \tag{9.103}$$

which also follows from the skew-symmetric property of the Levi-Civita symbol and can be shown by a switch-rename-reshuffle. This is a fundamental and familiar property of the cross product. We have obtained it as a consequence of the algebraic definition (9.95). Historically, it arose geometrically, where orthogonality was the key element of the definition. This is an instance, where tensor calculus inverts the traditional perspective and gives preference to the algebraic viewpoint.

Exercise 200. Show that $\varepsilon^{ijk} U_i U_j V_k = 0$ by a rename-switch-reshuffle.

Next, let us calculate the length of the vector W^i. It is an excellent technical exercise for working with Levi-Civita symbols. The length squared of the vector W^i is given by $W_i W^i$:

$$W_i W^i = \varepsilon_{ijk} U^j V^k \varepsilon^{irs} U_r V_s. \tag{9.104}$$

Combine the two Levi-Civita symbols into a single δ-symbol:

$$W_i W^i = \delta^{irs}_{ijk} U^j V^k U_r V_s. \tag{9.105}$$

By equation (9.21),

$$W_i W^i = \delta^{rs}_{jk} U^j V^k U_r V_s. \tag{9.106}$$

and by equation (9.26),

$$W_i W^i = \left(\delta^r_j \delta^s_k - \delta^s_j \delta^r_k\right) U^j V^k U_r V_s. \tag{9.107}$$

Continuing, we find

$$W_i W^i = U^j V^k U_j V_k - U^j V^k U_k V_j. \tag{9.108}$$

Rewriting this equation in a more attractive form yields

$$W_i W^i = U_j U^j V_k V^k - U_j V^j V_k U^k. \tag{9.109}$$

This is the final answer. In dyadic form, it reads

$$|\mathbf{W}|^2 = |\mathbf{U}|^2 |\mathbf{V}|^2 - (\mathbf{U} \cdot \mathbf{V})^2 . \tag{9.110}$$

Exercise 201. Show that $|\mathbf{W}| = |\mathbf{U}| |\mathbf{V}| \sin \alpha$ where $\sin \alpha$ is the angle between $\mathbf{U}$ and $\mathbf{V}$.

Exercise 202. As practice with working with Levi-Civita symbols, derive the expression for the cross product of three vectors. You should obtain an expression with a dyadic form that reads

$$(\mathbf{U} \times \mathbf{V}) \times \mathbf{W} = (\mathbf{U} \cdot \mathbf{W}) \mathbf{V} - (\mathbf{V} \cdot \mathbf{W}) \mathbf{U}. \tag{9.111}$$

Conclude that the cross product is not associative.

Exercise 203. Show that the indices can be rotated in the expression for $\mathbf{W}$ to yield

$$\mathbf{W} = \varepsilon^{ijk} U_i V_j \mathbf{Z}_k . \tag{9.112}$$

Exercise 204. Show that

$$\mathbf{U} \cdot (\mathbf{V} \times \mathbf{W}) = (\mathbf{U} \times \mathbf{V}) \cdot \mathbf{W}. \tag{9.113}$$

Exercise 205. Show that in Cartesian coordinates, the cross product is given by

$$\mathbf{W} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ U^1 & U^2 & U^3 \\ V^1 & V^2 & V^3 \end{vmatrix} \tag{9.114}$$

Exercise 206. Show that in cylindrical coordinates, the cross product is given by

$$\mathbf{W} = \begin{vmatrix} r\mathbf{Z}^1 & r\mathbf{Z}^2 & r\mathbf{Z}^3 \\ U^1 & U^2 & U^3 \\ V^1 & V^2 & V^3 \end{vmatrix} . \tag{9.115}$$

Exercise 207. Derive the expression for the cross product in spherical coordinates.

9.13 The Curl

The *curl* is an invariant differential operator which can be defined with the help of the Levi-Civita symbol. The term curl was coined by J. Clerk Maxwell in his celebrated *Treatise on Electricity and Magnetism [30]*. Given the tensor field U_k, its curl V^i is defined as

$$V^i = \varepsilon^{ijk} \nabla_j U_k. \tag{9.116}$$

In dyadic notation, the curl is denoted by a new symbol $\nabla\times$:

$$\mathbf{V} = \nabla \times \mathbf{U}. \tag{9.117}$$

As an illustration, let us analyze the curl of a curl of a vector field. In dyadic notation, let $\mathbf{V} = \nabla \times \nabla \times \mathbf{U}$. Then

$$V_r = \varepsilon_{rsi} \nabla^s \left(\varepsilon^{ijk} \nabla_j U_k\right). \tag{9.118}$$

By the metrinilic property of the Levi-Civita symbol,

$$V_r = \varepsilon_{rsi} \varepsilon^{ijk} \nabla^s \nabla_j U_k. \tag{9.119}$$

Combine the Levi-Civita symbols in to a single δ-symbol

$$V_r = \delta^{ijk}_{rsi} \nabla^s \nabla_j U_k, \tag{9.120}$$

and cycle the contravariant indices of the δ-symbol to put the index i in the last position:

$$V_r = \delta^{jki}_{rsi} \nabla^s \nabla_j U_k. \tag{9.121}$$

The rest of the analysis parallels that of the example in preceding section. By equations (9.21) and (9.26) we have

$$V_r = \delta^j_r \delta^k_s \nabla^s \nabla_j U_k - \delta^k_r \delta^j_s \nabla^s \nabla_j U_k, \tag{9.122}$$

which yields the near-final expression

$$V_r = \nabla^k \nabla_r U_k - \nabla^s \nabla_s U_r. \tag{9.123}$$

Renaming indices and switching the order of the covariant derivatives, we arrive at the final identity

$$V_i = \nabla_i \nabla_j U^j - \nabla_j \nabla^j U_i. \tag{9.124}$$

In dyadic notation, this identity reads

$$\nabla \times \nabla \times \mathbf{U} = \nabla \left(\nabla \cdot \mathbf{U}\right) - \nabla \cdot \nabla \mathbf{U}. \tag{9.125}$$

Exercise 208. Analyze the curl applied to a cross product. Show that $\mathbf{W} = \nabla \times (\mathbf{U} \times \mathbf{V})$ is given by

$$W^i = -\nabla_j U^j \, V^i + \nabla_j U^i \, V^j - U^j \nabla_j V^i + U^i \nabla_j V^j. \tag{9.126}$$

Although there is no good way to do so express this relationship in dyadic notation.

Exercise 209. Show that divergence of curl vanishes.

Exercise 210. Show that curl of a gradient vanishes.

9.14 Generalization to Other Dimensions

The ideas presented so far in this chapter naturally generalize to any number of dimensions. We spell out the generalization to two dimensions and outline the generalization to higher dimensions.

Suppose that all indices can now have values 1 or 2. The permutation symbols e_{ij} and e^{ij} are defined analogously to equation (9.8) and therefore have the values captured by the following matrix

$$e_{ij}, e^{ij} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{9.127}$$

The full delta system δ^{ij}_{rs} is defined by

$$\delta^{ij}_{rs} = e^{ij} e_{rs} \tag{9.128}$$

and satisfies the properties

$$\delta^{ij}_{rs} = \delta^i_r \delta^j_s - \delta^j_r \delta^i_s \tag{9.129}$$

and

$$\delta^i_j = \delta^{ir}_{jr}, \tag{9.130}$$

where the Kronecker symbol δ^i_j has the usual definition.

The determinant A of a ~~9~~2×2 system a^i_j is given by

$$A = \frac{1}{2}\delta^{ij}_{rs} a^r_i a^s_j, \tag{9.131}$$

and similarly satisfies the product property.

The definition (9.49) of the object Z is valid in all dimensions. Its square root $\sqrt{Z}$ is the volume element which, in two dimensions, could be referred to as the *area element*. However, we will stick to the term *volume element* and reserve the term *area element* for embedded surfaces in Part II.

The Levi-Civita symbols ε_{ij} and ε^{ij} are defined according to the formulas

$$\varepsilon_{ij} = \sqrt{Z} e_{ij} \tag{9.132}$$

$$\varepsilon^{ij} = \frac{e^{ij}}{\sqrt{Z}}. \tag{9.133}$$

The cross product involves a single tensor U^i:

$$V_i = \varepsilon_{ij} U^j, \tag{9.134}$$

and produces a vector that is orthogonal to U^j, since

$$V_i U^i = \varepsilon_{ij} U^i U^j = 0. \tag{9.135}$$

Exercise 211. Explain why $\varepsilon_{ij} U^i U^j = 0$ in equation (9.135).

Exercise 212. Show that the length of the vector V^i equals the length of the vector U^i.

The divergence and the Laplacian operators are defined identically in all dimensions, and the Voss–Weyl formula is universal, as well. The curl of a vector field U^i is the scalar field V given by

$$V = \varepsilon_{ij} \nabla^i U^j. \tag{9.136}$$

Next, let us discuss generalization to a general n-dimensional case. The permutation symbols $e_{i_1 \cdots i_n}$ and $e^{i_1 \cdots i_n}$ have n indices and yield the value 1 if all indices are different and form an even permutation, -1 if odd permutation, and 0 otherwise. The full delta system $\delta^{i_1 \cdots i_n}_{j_1 \cdots j_n}$ is defined by

$$\delta^{i_1 \cdots i_n}_{j_1 \cdots j_n} = e^{i_1 \cdots i_n} e_{j_1 \cdots j_n}, \tag{9.137}$$

and can be expressed by the determinant

$$\delta^{i_1\cdots i_n}_{j_1\cdots j_n} = \begin{vmatrix} \delta^{i_1}_{j_1} & \delta^{i_1}_{j_n} \\ & \\ \delta^{i_n}_{j_1} & \delta^{i_n}_{j_n} \end{vmatrix}. \tag{9.138}$$

The delta systems are related by contractions and a coefficient that equals the number of missing indices:

$$\delta^{i_1\cdots i_{k-1}}_{j_1\cdots j_{k-1}} = (n-k)\,\delta^{i_1\cdots i_{k-1}i_k}_{j_1\cdots j_{k-1}i_k} \text{ (note the contraction on } i_k\text{)}. \tag{9.139}$$

In particular,

$$\delta^{i_1\cdots i_k}_{i_1\cdots i_k} = k! \tag{9.140}$$

The determinant A of an $n \times n$ matrix a^i_j is given by the equation

$$A = \frac{1}{n!}\delta^{i_1\cdots i_n}_{j_1\cdots j_n} a^{j_1}_{i_1} \cdots a^{j_n}_{i_n}.$$

The cross product involves $n-1$ vectors $U^i_{(1)}, \cdots U^i_{(n-1)}$

$$V_i = \varepsilon_{ij_1\cdots j_{n-1}} U^{j_1}_{(1)}, \cdots U^{j_{n-1}}_{(n-1)} \tag{9.141}$$

and the curl is not defined.

9.15 Summary

This chapter offered us a welcome opportunity to use the tensor framework developed over the preceding chapters. The object of our study here was the determinant. In the course of our discussion, we introduced the δ-symbols, which proved to be tensors of tremendous utility. We proceeded to define the determinant in tensor notation. The tensor notation proved to be a natural turn to expresses the determinant and enabled us to offer simple proofs of the key properties, including the one-line proof of the product property $|AB| = |A|\,|B|$ in equation (9.34).

The Levi-Civita symbols ε^{ijk} and ε_{ijk} were the stars of the second half of this chapter. Along with the volume element $\sqrt{Z}$, the Levi-Civita symbols complete the list of metrics. We demonstrated that the Levi-Civita symbols are tensors and that the metrinilic property applies to them. Finally, we used the Levi-Civita symbols to introduce the vector product and the curl.

Part II
Tensors on Surfaces

Chapter 10
The Tensor Description of Embedded Surfaces

10.1 Preview

We have finally arrived at the subject of surfaces. This topic is extraordinarily rich and it lets tensor calculus shine at its brightest. We will focus on two-dimensional surfaces in three-dimensional Euclidean spaces. The Euclidean space in which the surface is embedded is called the *ambient* space. The analysis presented in this chapter is easily extended to *hypersurfaces* of any dimension. A *hypersurface* is a differentiable $(n-1)$-dimensional subspace of an n-dimensional space. That is, a hypersurface is characterized by a codimension of 1, codimension being the difference between the dimensions of the ambient and embedded spaces. Curves, the subject of Chap. 13, embedded in the three-dimensional Euclidean space have codimension two.

A useful term that applies to subspaces of any codimension is *manifold*. The meaning of this term, originally introduced by Henri Poincaré in *Analysis Situs*, has undergone an evolution in recent decades. Its currently accepted definition is given in the context of topology. We adopt a narrower meaning that is closer to Poincaré's original definition: a manifold is a subspace of the Euclidean space that can be parameterized by smooth functions defined on a region of $\mathbb{R}^m$, as in equations (10.1a)–(10.1c). For $m = 1$, the manifold is a curve. For $m = n - 1$, the manifold is a hypersurface.

This chapter focuses on the tangent and normal spaces that exist at all points on the surface where the surface is sufficiently smooth. For a hypersurface, the tangent space is $(n-1)$-dimensional and is spanned by a set of $n-1$ vectors that form surface covariant basis. The normal space is one-dimensional and is spanned by a single vector $\mathbf{N}$ called the *normal*.

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_10,

10.2 Parametric Description of Surfaces

There are various ways to specify a surface. We describe surfaces *parametrically*. That is, we choose two variables S^1 and S^2 (collectively, S^α) and let each of the *ambient coordinates* Z^1, Z^2, and Z^3 depend on S^α:

$$Z^1 = Z^1\left(S^1, S^2\right) \tag{10.1a}$$

$$Z^2 = Z^2\left(S^1, S^2\right) \tag{10.1b}$$

$$Z^3 = Z^3\left(S^1, S^2\right). \tag{10.1c}$$

For example, if the ambient Euclidean space is referred to Cartesian coordinates, then the equations

$$x(\theta, \phi) = A \sin\theta \cos\phi \tag{10.2a}$$

$$y(\theta, \phi) = B \sin\theta \sin\phi \tag{10.2b}$$

$$z(\theta, \phi) = C \sin\theta \tag{10.2c}$$

describe an ellipsoid with semiaxes A, B, and C. The coordinate lines for this parameterization are shown in Fig. 10.1.

The variables S^α are called *surface coordinates*. We use letters from the Greek alphabet for surface indices, which assume values from 1 to m. We continue to use the Latin alphabet for ambient indices which change from 1 to n. Most of the equations in this chapter apply for arbitrary n and m and will point out when that is not the case. Our primary interest is two-dimensional surfaces ($m = 2$) embedded in three-dimensional Euclidean spaces ($n = 3$).

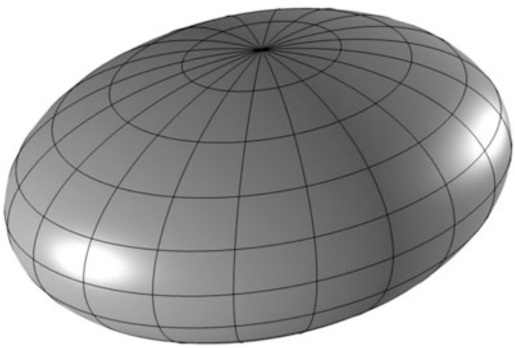


Fig. 10.1 The surface of an ellipsoid referred to in spherical coordinates, a natural choice

10.3 The Fundamental Differential Objects on the Surface

Rewrite equations (10.1a)–(10.1c) in tensor notation

$$Z^i = Z^i(S), \tag{10.3}$$

where we enumerated the ambient variables and suppressed the surface index of the function argument.

The important role of the shift tensor becomes evident when we introduce the *covariant basis* $\mathbf{S}_\alpha$. It is constructed in a way that is analogous to the *ambient covariant basis* $\mathbf{Z}_i$, the definition of which is given in equation (5.2). Consider the position vector $\mathbf{R}$ as a function of surface coordinates S^α:

$$\mathbf{R} = \mathbf{R}(S). \tag{10.4}$$

The covariant basis $\mathbf{S}_\alpha$ is defined by partial differentiation

$$\mathbf{S}_\alpha = \frac{\partial \mathbf{R}}{\partial S^\alpha}. \tag{10.5}$$

The geometric interpretation of $\mathbf{S}_\alpha$ is straightforward: these vectors are tangential to the surface S; more specifically, they are tangential to the coordinate lines corresponding to S^1 and S^2. The plane spanned by the vectors $\mathbf{S}_\alpha$ is called the *tangential plane* (Fig. 10.2).

There is a natural connection between the surface and ambient bases. It comes from the identity

$$\mathbf{R}(S) = \mathbf{R}(Z(S)), \tag{10.6}$$

which states this obvious fact: if we consider $\mathbf{R}$ as a function of the ambient coordinates Z^i and substitute into that function the equation of the surface (10.3), then the result of the composition is the position vector $\mathbf{R}$ as a function of the surface coordinates S^α. Differentiate the identity in equation (10.6) with respect to S^α:

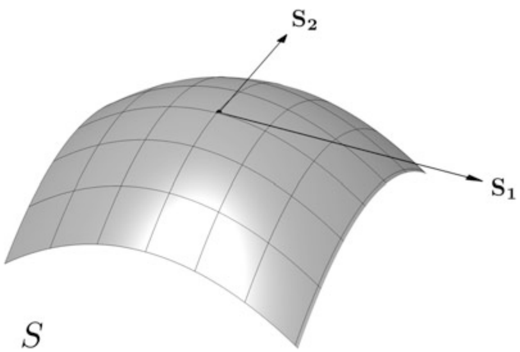


Fig. 10.2 A curved surface, its coordinate lines, and a covariant basis at one of the points

$$\frac{\partial \mathbf{R}}{\partial S^\alpha} = \frac{\partial \mathbf{R}}{\partial Z^i}\frac{\partial Z^i}{\partial S^\alpha}, \tag{10.7}$$

where $\partial Z^i/\partial S^\alpha$ are the partial derivatives of the equation of the surface (10.3). These derivatives are denoted by the symbol Z^i_α known as the *shift tensor*:

$$Z^i_\alpha = \frac{\partial Z^i}{\partial S^\alpha}. \tag{10.8}$$

With this new symbol, equation (10.7) can be rewritten as

$$\mathbf{S}_\alpha = Z^i_\alpha \mathbf{Z}_i. \tag{10.9}$$

Thus, the shift tensor relates the surface and the ambient basis. Equation (10.9) is only one of out many important relationships in which the shift tensor plays a key role. The shift tensor is prominent is all identities that relate ambient and surface objects.

We see from equation (10.9) that the entries of the shift tensor Z^i_α are the components of the surface covariant basis vectors $\mathbf{S}_\alpha$ with respect to the ambient basis $\mathbf{Z}_i$. Dotting both sides of this equation with $\mathbf{Z}^j$ (and the renaming $j \to i$), we obtain an equation that is equivalent to (10.9) and gives an explicit expression for Z^i_α:

$$Z^i_\alpha = \mathbf{S}_\alpha \cdot \mathbf{Z}^i. \tag{10.10}$$

It may be said that *the shift tensor represents the tangent space*. Indeed, suppose that a vector $\mathbf{T}$ lies in the tangent space and has components T^i:

$$\mathbf{T} = T^i \mathbf{Z}_i. \tag{10.11a}$$

The tensor T^i is called the *ambient components* of vector $\mathbf{T}$. Then the system T^i can be represented as a linear combinations of the systems Z^i_1 and Z^i_2, and that is the sense in which Z^i_α represents the tangent space.

Let us demonstrate why the components T^i of any tangent vector $\mathbf{T}$ can be represented by a linear combination of Z^i_1 and Z^i_2. By definition, $\mathbf{T}$ can be represented by a linear combination of the surface covariant basis

$$\mathbf{T} = T^\alpha \mathbf{S}_\alpha. \tag{10.12}$$

The tensor T^α represents the *surface components* of the vector $\mathbf{T}$. Substitute equation (10.9) into equation (10.12):

$$\mathbf{T} = T^\alpha Z^i_\alpha \mathbf{Z}_i. \tag{10.13}$$

In combination with (10.11a), this equation leads to the conclusion that

$$T^i = T^\alpha Z_\alpha^i \tag{10.14}$$

which shows T^i as a linear combination of Z_1^i and Z_2^i and indicates that the coefficients of this linear combination are the surface components of $\mathbf{T}$. Equation (10.14) also justifies the name of the shift tensor: it translates or *shifts* components of tangent vector from surface to ambient.

This shifting works in both directions. We show below that the following identity holds

$$T^\alpha = T^i Z_i^\alpha. \tag{10.15}$$

However, we have not yet developed all the machinery to use the shift tensor Z_i^α with an upper surface index because we have not yet introduced the surface metric tensor. We now turn to the task of developing the metric tensors.

The *surface covariant metric tensor* $S_{\alpha\beta}$ is defined analogously to the ambient metric tensor:

$$S_{\alpha\beta} = \mathbf{S}_\alpha \cdot \mathbf{S}_\beta. \tag{10.16}$$

In words, the entries of the metric tensor are the pairwise dot products of the covariant basis elements. Compare equation (10.16) to the analogous definition of the ambient covariant metric tensor in equation (5.7). Therefore, by similar arguments, the covariant metric tensor is symmetric

$$S_{\alpha\beta} = S_{\beta\alpha} \tag{10.17}$$

and positive definite.

Naturally, the ambient and the surface covariant bases are related by the shift tensor. Equation (10.9) helps derive that relationship:

$$S_{\alpha\beta} = \mathbf{S}_\alpha \cdot \mathbf{S}_\beta = \mathbf{Z}_i Z_\alpha^i \cdot \mathbf{Z}_j Z_\beta^j = Z_{ij} Z_\alpha^i Z_\beta^j. \tag{10.18}$$

In brief,

$$S_{\alpha\beta} = Z_{ij} Z_\alpha^i Z_\beta^j. \tag{10.19}$$

By letting the metric tensor lower one of the ambient indices, we may rewrite equation (10.19) as

$$S_{\alpha\beta} = Z_{i\alpha} Z_\beta^i. \tag{10.20}$$

The contravariant metric tensor $S^{\alpha\beta}$ is defined as the matrix inverse of $S_{\alpha\beta}$:

$$S^{\alpha\beta} S_{\beta\gamma} = \delta^{\alpha}_{\gamma}, \tag{10.21}$$

where δ^{α}_{γ} satisfies the usual definition of the Kronecker symbol. The contravariant metric tensor is symmetric and positive, definite and much of what could be said of the ambient metric tensor can also be said of the surface metric tensor. Now that we have the covariant and the contravariant metric tensors, we have rules for juggling surface indices. For instance,

$$Z^{i\beta} = Z^{i}_{\beta} S^{\alpha\beta}. \tag{10.22}$$

Further, by raising the index α, we can rewrite equation (10.20) in the following beautiful way

$$Z^{\alpha}_{i} Z^{i}_{\beta} = \delta^{\alpha}_{\beta} \tag{10.23}$$

In other words, the objects Z^{α}_{i} and Z^{i}_{β} are matrix inverses of each other in the sense of contraction by the longer dimension.

Exercise 213. Explain why it is not possible that $Z^{i}_{\alpha} Z^{\alpha}_{j} = \delta^{i}_{j}$.

Exercise 214. Derive equation (10.15) from equation (10.14).

The rest of the surface metrics are defined analogously to their ambient counterparts. The object S is defined as the determinant of the covariant metric tensor

$$S = |S_{\cdot\cdot}| \tag{10.24}$$

and its square root $\sqrt{S}$ is the *area element*, by analogy with the volume element $\sqrt{Z}$ in the ambient space. The Levi-Civita symbols are defined according to the formulas

$$\varepsilon_{\alpha\beta} = \sqrt{S} e_{\alpha\beta} \tag{10.25a}$$

$$\varepsilon^{\alpha\beta} = \frac{e^{\alpha\beta}}{\sqrt{S}}, \tag{10.25b}$$

where $e_{\alpha\beta}$ is the skew-symmetric system with nonzero entries

$$e_{12} = -e_{21} = 1. \tag{10.26}$$

The surface δ-symbols are also defined analogously to their ambient counterparts.

10.4 Surface Tensors

In what sense is Z_α^i a tensor? We are working with two simultaneous systems of coordinates, ambient coordinates Z^i, and surface coordinates S^α. Both choices are arbitrary and we must consider the transformations of variants under changes of both coordinate systems. Suppose that $Z^{i'}$ and $S^{\alpha'}$ are two alternative coordinate systems. The ambient coordinates Z^i and $Z^{i'}$ are related by equations (4.63) and (4.64) and the surface coordinates S^α and $S^{\alpha'}$ are related by the following mutually inverse relationships:

$$S^{\alpha'} = S^{\alpha'}(S) \tag{10.27}$$

$$S^\alpha = S^\alpha\left(S'\right). \tag{10.28}$$

Define the Jacobian objects $J_{\alpha'}^\alpha$ and $J_\alpha^{\alpha'}$ by

$$J_\alpha^{\alpha'} = \frac{\partial S^{\alpha'}(S)}{\partial S^\alpha} \tag{10.29}$$

$$J_{\alpha'}^\alpha = \frac{\partial S^\alpha(S')}{\partial S^{\alpha'}}. \tag{10.30}$$

We now define the tensor property for surface variants. Consider a variant $T_{j\beta}^{i\alpha}$ with a fully representative collection of indices. Suppose that its values in the alternative coordinate systems is $T_{j'\beta'}^{i'\alpha'}$. Then $T_{j\beta}^{i\alpha}$ is a tensor if

$$T_{j'\beta'}^{i'\alpha'} = T_{j\beta}^{i\alpha} J_i^{i'} J_{j'}^j J_\alpha^{\alpha'} J_{\beta'}^\beta, \tag{10.31}$$

and it is clear how this definition should be modified for other index signatures.

Let us demonstrate that the shift tensor Z_α^i is indeed a tensor. Its entries $Z_{\alpha'}^{i'}$ in the alternative coordinates are given by the partial derivative

$$Z_{\alpha'}^{i'} = \frac{\partial Z^{i'}(S')}{\partial S^{\alpha'}}, \tag{10.32}$$

where $Z^{i'}\left(S'\right)$ is the parametric representation of the surface in the alternative coordinates. The three functions $Z^{i'}(S')$ can be constructed by the following double composition

$$Z^{i'}\left(S'\right) = Z^{i'}\left(Z\left(S\left(S'\right)\right)\right), \tag{10.33}$$

where $Z^{i'}(Z)$ gives the new ambient coordinates in terms of the old ambient coordinates, $Z(S)$ represents the surface in old coordinates and $S(S')$ gives the old surface coordinates in terms of the new surface coordinates. By a repeated application the chain rule, we have

$$\frac{\partial Z^{i'}}{\partial Z^{\alpha'}} = \frac{\partial Z^{i'}}{\partial Z^{i}} \frac{\partial Z^{i}}{\partial S^{\alpha}} \frac{\partial S^{\alpha}}{\partial S^{\alpha'}}, \tag{10.34}$$

or

$$Z^{i'}_{\alpha'} = Z^{i}_{\alpha} J^{i'}_{i} J^{\alpha}_{\alpha'}, \tag{10.35}$$

and the proof of the tensor property of the shift tensor is complete.

At this point, we only need to mention that the entire contents of Chap. 6 carry over naturally to surface tensors. We conclude that $\mathbf{S}_\alpha$, $S_{\alpha\beta}$, $S^{\alpha\beta}$, $\varepsilon^{\alpha\beta}$, $\varepsilon_{\alpha\beta}$ are tensors and that S, $e_{\alpha\beta}$ and $e^{\alpha\beta}$ are relative tensors.

10.5 The Normal N

The covariant basis $\mathbf{S}_\alpha$ and its ambient components captured by the shift tensor Z^i_α are analytical representations of the tangent plane. The normal direction is captured by the normal $\mathbf{N}$ with components N^i. The normal vector $\mathbf{N}$ is defined by the identities

$$\mathbf{N} \cdot \mathbf{S}_\alpha = 0 \tag{10.36a}$$

$$\mathbf{N} \cdot \mathbf{N} = 1, \tag{10.36b}$$

where the first equation states that the normal is orthogonal to the tangent plane and the second equation states that the normal is unit length. These equations are satisfied by two equal and opposite vectors and the choice of normal depends on the particular problem. For closed surfaces, the exterior normal is often chosen.

Denote the contravariant components of the normal vector by N^i:

$$\mathbf{N} = N^i \mathbf{Z}_i. \tag{10.37}$$

We will later establish the following explicit expression for N^i

$$N^i = \frac{1}{2} \varepsilon^{ijk} \varepsilon_{\alpha\beta} Z^\alpha_j Z^\beta_k. \tag{10.38}$$

However, this expression is not needed to derive most of the relationships involving N^i. For example, equation (10.36a) implies

$$N_i Z^i_\alpha = 0. \tag{10.39}$$

To show this, we write

$$0 = \mathbf{N} \cdot \mathbf{S}_\alpha = N^j \mathbf{Z}_j \cdot Z^i_\alpha \mathbf{Z}_i = N^j Z^i_\alpha Z_{ij} = N_i Z^i_\alpha. \tag{10.40}$$

Similarly, equation (10.36b) implies

$$N_i N^i = 1. \tag{10.41}$$

10.6 The Normal and Orthogonal Projections

The normal projection of a vector $\mathbf{V}$ is a vector $\mathbf{P}$ that point along the normal $\mathbf{N}$ and has the length such that the difference $\mathbf{V} - \mathbf{P}$ is orthogonal to $\mathbf{N}$. From geometric considerations, it is evident that the length of $\mathbf{P}$ equals $\mathbf{V} \cdot \mathbf{N}$ and therefore

$$\mathbf{P} = (\mathbf{V} \cdot \mathbf{N})\, \mathbf{N}. \tag{10.42}$$

If the geometric reasoning does not convince you, the validity of equation (10.42) can be established algebraically.

Exercise 215. Show that $(\mathbf{V} - \mathbf{P}) \cdot \mathbf{N}$, where $\mathbf{P}$ is defined by equation (10.42), is zero.

In component form, equation (10.42) reads

$$\mathbf{P} = V^j N_j N^i \mathbf{Z}_i. \tag{10.43}$$

Therefore, the components of P^i of $\mathbf{P}$ are given by

$$P^i = N^i N_j V^j. \tag{10.44}$$

The tensor

$$N^i N_j \tag{10.45}$$

is therefore the projection operator onto the normal $\mathbf{N}$. In alternative terminology, it is the projection *away from the surface* S.

Exercise 216. Denote the tensor $N^i N_j$ by P^i_j. Show that

$$P^i_j P^j_k = P^i_k. \tag{10.46}$$

In other words, a repeated application of this projection operator leaves the projection unchanged. In matrix notation, this property reads $P^2 = P$. From this form, it is evident that all eigenvalues of P or either 0 or 1.

We now turn to the orthogonal projection *onto the surface* S. The orthogonal projection of a vector $\mathbf{V}$ is the vector $\mathbf{T}$ in the tangent plane such that $\mathbf{V} - \mathbf{T}$ is orthogonal to the tangent plane. The orthogonal projection $\mathbf{T}$ is given by the remarkably simple expression

$$\mathbf{T} = (\mathbf{V} \cdot \mathbf{S}^\alpha)\, \mathbf{S}_\alpha. \tag{10.47}$$

It is evident that $\mathbf{T}$ lies in the tangent plane.

Exercise 217. To show $\mathbf{V} - \mathbf{T}$ that is orthogonal to the tangent plane, show that $(\mathbf{V} - \mathbf{T}) \cdot \mathbf{S}^{\beta} = 0$.

In component form, equation (10.47) reads

$$\mathbf{T} = V^j Z_j^{\alpha} Z_{\alpha}^i \mathbf{Z}_i, \tag{10.48}$$

which yields the component T^i of the vector $\mathbf{T}$

$$T^i = Z_{\alpha}^i Z_j^{\alpha} V^j. \tag{10.49}$$

Therefore, the tensor

$$Z_{\alpha}^i Z_j^{\alpha} \tag{10.50}$$

is interpreted as the operator of projection onto the surface S.

The compactness of the expression $Z_{\alpha}^i Z_j^{\alpha}$ is indeed remarkable. If you will recall from linear algebra, the problem of finding the orthogonal projection onto a linear subspace is solved by the technique of least squares [44]. If A is the rectangular matrix that has columns spanning the linear subspace (i.e., A corresponds to Z_{α}^i), then the projection operator is given by the combination

$$A\left(A^T A\right)^{-1} A^T. \tag{10.51}$$

Expression (10.50) includes all of the same elements; however, the cumbersome part of equation (10.51) $\left(A^T A\right)^{-1}$ is captured by the raised surface index of the second shift tensor. This is because $A^T A$ is the matrix expression for $S_{\alpha\beta}$ and $\left(A^T A\right)^{-1}$ is the matrix expression for $S^{\alpha\beta}$.

Exercise 218. Denote the tensor $N^i N_j$ by T_j^i. Show that

$$T_j^i T_k^j = T_k^i. \tag{10.52}$$

The vector $\mathbf{V}$ equals the sum of its normal and orthogonal projections

$$\mathbf{V} = \mathbf{P} + \mathbf{T}. \tag{10.53}$$

This identity is evident geometrically, but can also be shown algebraically. For an algebraic proof, note that the set of vectors $\{\mathbf{N}, \mathbf{S}_1, \mathbf{S}_2\}$ forms a basis for the three-dimensional linear space at any point on the surface. Since dotting equation (10.53) with each of these vectors produce the same result on both sides, the validity of this equation is confirmed.

In component form, equation (10.53) reads

$$V^i = \left(N^i N_j + Z_{\alpha}^i Z_j^{\alpha}\right) V^j. \tag{10.54}$$

Since this identity is valid for any V^j, we arrive at the fundamental relationship

$$N^i N_j + Z^i_\alpha Z^\alpha_j = \delta^i_j. \tag{10.55}$$

This beautiful relationship is used frequently in tensor calculus, and you will encounter it often in the remainder of this book.

Exercise 219. Derive equation (10.41) from equation (10.55).

10.7 Working with the Object N^i

We derived the key properties of the normal **N** and its components N^i without the explicit formula (10.38):

$$N^i = \frac{1}{2}\varepsilon^{ijk}\varepsilon_{\alpha\beta}Z^\alpha_j Z^\beta_k. \tag{10.38}$$

In applications, the formula (10.38) is often used to calculate the actual components N^i. We note that equation (10.38) applies to two-dimensional surfaces. For other dimensions it needs to be modified. For a two-dimensional ambient space, the equation becomes

$$N^i = \varepsilon^{ij}\varepsilon_\alpha Z^\alpha_j. \tag{10.56}$$

For an n-dimensional ambient space, the definition of N^i reads

$$N^i = \frac{1}{(n-1)!}\varepsilon^{ij_1\cdots j_{n-1}}\varepsilon_{\alpha_1\cdots\alpha_{n-1}}Z^{\alpha_1}_{j_1}\dots Z^{\alpha_{n-1}}_{j_{n-1}}. \tag{10.57}$$

The following discussion applies to the three-dimensional ambient space, but can be easily extended to hypersurfaces in any dimension.

We first establish that formula (10.38) is indeed a correct expression for the components of the normal. In other words, we must verify equations (10.39) and (10.41). To verify (10.39), we write

$$N^i Z^\gamma_i = \frac{1}{2}\varepsilon^{ijk}\varepsilon_{\alpha\beta}Z^\alpha_j Z^\beta_k Z^\gamma_i. \tag{10.58}$$

For any system T^α_i, the combination $\varepsilon^{ijk}T^\alpha_i T^\beta_j T^\gamma_k$ is fully skew-symmetric on the indices α, β, and γ. For instance, switching α and β, we find by the swap–reshuffle–rename strategy

$$\varepsilon^{ijk}T^\beta_i T^\alpha_j T^\gamma_k = -\varepsilon^{jik}T^\beta_i T^\alpha_j T^\gamma_k = -\varepsilon^{jik}T^\alpha_j T^\beta_i T^\gamma_k = -\varepsilon^{ijk}T^\alpha_i T^\beta_j T^\gamma_k. \tag{10.59}$$

Being fully-skew symmetric, $\varepsilon^{ijk} T_i^\alpha T_j^\beta T_k^\gamma$ vanishes if any of the indices α, β, or γ have the same value, which is unavoidable since these three indices can assume only two different values. Therefore, $\varepsilon^{ijk} Z_j^\alpha Z_k^\beta Z_i^\gamma$ vanishes identically and equation (10.39) is confirmed.

Equation (10.41) presents a more interesting opportunity of working with the Levi-Civita symbols. Let us begin by setting up the expression $N_i N^i$

$$N_i N^i = \frac{1}{4} \varepsilon_{ijk} \varepsilon^{\alpha\beta} Z_\alpha^j Z_\beta^k \varepsilon^{irs} \varepsilon_{\rho\sigma} Z_r^\rho Z_s^\sigma. \tag{10.60}$$

First, combine the Levi-Civita symbols into δ-symbols:

$$N_i N^i = \frac{1}{4} \delta_{ijk}^{irs} \delta_{\rho\sigma}^{\alpha\beta} Z_\alpha^j Z_\beta^k Z_r^\rho Z_s^\sigma. \tag{10.61}$$

Recall that

$$\delta_{\rho\sigma}^{\alpha\beta} = \delta_\rho^\alpha \delta_\sigma^\beta - \delta_\rho^\beta \delta_\sigma^\alpha. \tag{10.62}$$

Thus, the right-hand side in equation (10.61) breaks up into two terms. In the first term, ρ becomes α and σ becomes β, and visa versa in the second term:

$$N_i N^i = \frac{1}{4} \delta_{ijk}^{irs} Z_\alpha^j Z_\beta^k Z_r^\alpha Z_s^\beta - \frac{1}{4} \delta_{ijk}^{irs} Z_\alpha^j Z_\beta^k Z_r^\beta Z_s^\alpha. \tag{10.63}$$

Next, recall that

$$\delta_{ijk}^{irs} = \delta_{jk}^{rs} \tag{10.64}$$

and

$$\delta_{jk}^{rs} = \delta_j^r \delta_k^s - \delta_j^s \delta_k^r. \tag{10.65}$$

We therefore have four terms

$$N_i N^i = \frac{1}{4} Z_\alpha^j Z_\beta^k Z_j^\alpha Z_k^\beta - \frac{1}{4} Z_\alpha^j Z_\beta^k Z_k^\alpha Z_j^\beta - \frac{1}{4} Z_\alpha^j Z_\beta^k Z_j^\beta Z_k^\alpha + \frac{1}{4} Z_\alpha^j Z_\beta^k Z_k^\beta Z_j^\alpha. \tag{10.66}$$

The key to the next step is equation (10.23). Using that identity twice in each term yields

$$N_i N^i = \frac{1}{4} \delta_\alpha^\alpha \delta_\beta^\beta - \frac{1}{4} \delta_\alpha^\beta \delta_\beta^\alpha - \frac{1}{4} \delta_\alpha^\beta \delta_\beta^\alpha + \frac{1}{4} \delta_\alpha^\alpha \delta_\beta^\beta. \tag{10.67}$$

It is left as an exercises to confirm that the sum of the four terms on the right-hand side is indeed 1.

The final relationship to be verified is equation (10.55). The derivation essentially follows that of equation (10.41). Form the combination $N^i N_r$

$$N^i N_r = \frac{1}{4}\varepsilon^{ijk}\varepsilon_{\alpha\beta}Z_j^\alpha Z_k^\beta \varepsilon_{rst}\varepsilon^{\sigma\tau}Z_\sigma^s Z_\tau^t, \tag{10.68}$$

and combine the Levi-Civita symbols into δ-symbols:

$$N^i N_r = \frac{1}{4}\delta_{rst}^{ijk}\delta_{\alpha\beta}^{\sigma\tau}Z_j^\alpha Z_k^\beta Z_\sigma^s Z_\tau^t, \tag{10.69}$$

By the identity $\delta_{\alpha\beta}^{\sigma\tau} = \delta_\alpha^\sigma \delta_\beta^\tau - \delta_\alpha^\tau \delta_\beta^\alpha$, we have

$$N^i N_r = \frac{1}{4}\delta_{rst}^{ijk}Z_j^\alpha Z_k^\beta Z_\alpha^s Z_\beta^t - \frac{1}{4}\delta_{rst}^{ijk}Z_j^\alpha Z_k^\beta Z_\beta^s Z_\alpha^t, \tag{10.70}$$

For brevity, let

$$T_j^i = Z_\alpha^i Z_j^\alpha \tag{10.71}$$

and recall from an earlier exercise that

$$T_j^i T_k^j = T_k^i. \tag{10.72}$$

Further,

$$T_i^i = 2 \tag{10.73}$$

Then

$$N^i N_r = \frac{1}{4}\delta_{rst}^{ijk}T_j^s T_k^t - \frac{1}{4}\delta_{rst}^{ijk}T_j^t T_k^s. \tag{10.74}$$

Exercise 220. Show that the two terms in equation (10.74) are equal, thus

$$N^i N_r = \frac{1}{2}\delta_{rst}^{ijk}T_j^s T_k^t \tag{10.75}$$

According to equation (9.20), we find

$$N^i N_r = \frac{1}{2}\left(\delta_r^i\delta_s^j\delta_t^k - \delta_r^j\delta_s^i\delta_t^k + \delta_r^j\delta_s^k\delta_t^i - \delta_r^k\delta_s^j\delta_t^i + \delta_r^k\delta_s^i\delta_t^j - \delta_r^i\delta_s^k\delta_t^j\right)T_j^s T_k^t, \tag{10.76}$$

which yields the following six terms

$$N^i N_r = \frac{1}{2}\left(\delta_r^i T_j^j T_k^k - \delta_r^j T_j^i T_k^k + \delta_r^j T_j^k T_k^i - \delta_r^k T_j^j T_k^i + \delta_r^k T_j^i T_k^j - \delta_r^i T_j^k T_k^j\right). \tag{10.77}$$

With the help of equations (10.72) and (10.73), we simplify each term

$$N^i N_r = \frac{1}{2}\left(4\delta_r^i - 2T_r^i + T_r^i - 2T_r^i + T_r^i - 2\delta_r^i\right). \tag{10.78}$$

We have therefore obtained

$$N^i N_r = \delta_r^i - T_r^i, \tag{10.79}$$

which is precisely the relationship that we set out to prove.

10.8 The Christoffel Symbol $\Gamma_{\beta\gamma}^{\alpha}$

The Christoffel symbol measures the rate of change of the covariant basis with respect the coordinate variables. Recall the definition of the ambient Christoffel symbol Γ_{ij}^k in Chap. 5:

$$\frac{\partial \mathbf{Z}_i}{\partial Z^j} = \Gamma_{ij}^k \mathbf{Z}_k. \tag{5.56}$$

The analogous definition $\partial \mathbf{S}_\alpha / \partial S^\beta = \Gamma_{\alpha\beta}^{\gamma} \mathbf{S}_\gamma$ is not possible on an embedded surface. This is because the covariant basis $\mathbf{S}_\alpha$ is capable of representing vectors only in the tangent plane. On the other hand, there is no reason to expect the vectors $\partial \mathbf{S}_\alpha / \partial S^\beta$ to lie in the tangent plane. If the surface curves, at least some of the four vectors $\partial \mathbf{S}_\alpha / \partial S^\beta$ will have a component along the normal direction.

We will therefore define the Christoffel symbol $\Gamma_{\alpha\beta}^{\gamma}$ by analogy with equation (5.60):

$$\Gamma_{ij}^k = \mathbf{Z}^k \cdot \frac{\partial \mathbf{Z}_i}{\partial Z^j}. \tag{5.60}$$

The definition reads

$$\Gamma_{\beta\gamma}^{\alpha} = \mathbf{S}^\alpha \cdot \frac{\partial \mathbf{S}_\beta}{\partial S^\gamma}. \tag{10.80}$$

Since the central function of the Christoffel symbol is to facilitate covariant differentiation, our primary concern is that the Christoffel symbol transforms according to a certain rule and that is certainly achieved by definition (10.80). That rule is

$$\Gamma_{\beta'\gamma'}^{\alpha'} = \Gamma_{\beta\gamma}^{\alpha} J_\alpha^{\alpha'} J_{\beta'}^{\beta} J_{\gamma'}^{\gamma} + J_{\beta'\gamma'}^{\alpha} J_\alpha^{\alpha'}. \tag{10.81}$$

Exercise 221. Derive equation (10.81) from the definition (10.80).

The Christoffel tensor can also be defined bypassing the covariant basis $\mathbf{S}_\alpha$ and by referencing the covariant basis $S_{\alpha\beta}$

$$\Gamma^\alpha_{\beta\gamma} = \frac{1}{2} Z^{\alpha\omega} \left(\frac{\partial S_{\omega\beta}}{\partial S^\gamma} + \frac{\partial S_{\omega\gamma}}{\partial S^\beta} - \frac{\partial S_{\beta\gamma}}{\partial S^\omega} \right), \tag{10.82}$$

which is analogous to equation (5.66). The possibility of this definition takes on great significance in the discussion of Riemann spaces.

Exercise 222. Derive equation (10.82).

Finally, we determine the relationship between the ambient and surface Christoffel symbols. Into equation (10.80) substitute the expression for the surface basis in terms of the ambient basis

$$\Gamma^\alpha_{\beta\gamma} = Z^\alpha_i \mathbf{Z}^i \cdot \frac{\partial \left(Z^j_\beta \mathbf{Z}_j \right)}{\partial S^\gamma}. \tag{10.83}$$

By the product rule

$$\Gamma^\alpha_{\beta\gamma} = Z^\alpha_i \mathbf{Z}^i \cdot \frac{\partial Z^j_\beta}{\partial S^\gamma} \mathbf{Z}_j + Z^\alpha_i Z^j_\beta \mathbf{Z}^i \cdot \frac{\partial \mathbf{Z}_j}{\partial S^\gamma}. \tag{10.84}$$

In the first term, the dot product of the basis vectors gives δ^i_j. In the second term, the partial derivative is analyzed by the chain rule. The final relationship is

$$\Gamma^\alpha_{\beta\gamma} = Z^\alpha_i \frac{\partial Z^i_\beta}{\partial S^\gamma} + \Gamma^i_{jk} Z^\alpha_i Z^j_\beta \cancel{Z^k_\beta}. \tag{10.85}$$

This relationship offers an effective way of calculating the surface Christoffel symbol.

Exercise 223. Show that when the ambient space is referred to affine coordinates, the relationship between the surface and the ambient Christoffel symbols simplifies to

$$\Gamma^\alpha_{\beta\gamma} = Z^\alpha_i \frac{\partial Z^i_\beta}{\partial S^\gamma}. \tag{10.86}$$

10.9 The Length of an Embedded Curve

Suppose that a curve that lies in the surface S is given parametrically by $\left(S^1(t), S^2(t) \right)$. Collectively,

$$S^{\alpha} = S^{\alpha}(t). \tag{10.87}$$

This curve can also be considered as a curve in the ambient space where its parametric equations are obtained by composing the surface equations (10.3) with curve equations

$$Z^{i}(t) = Z^{i}(S(t)). \tag{10.88}$$

Differentiating with respect to t, we find

$$\frac{dZ^{i}}{dt} = Z_{\alpha}^{i} \frac{dS^{\alpha}}{dt}. \tag{10.89}$$

Since the length of the curve is given by

$$L = \int_{a}^{b} \sqrt{Z_{ij} \frac{dZ^{i}}{dt} \frac{dZ^{j}}{dt}} dt, \tag{5.20}$$

by substituting (10.89) in (5.20), we find

$$L = \int_{a}^{b} \sqrt{Z_{ij} Z_{\alpha}^{i} \frac{dS^{\alpha}}{dt} Z_{\beta}^{j} \frac{dS^{\beta}}{dt}} dt, \tag{10.90}$$

which, by equation (10.19), yields the beautiful, albeit expected, result

$$L = \int_{a}^{b} \sqrt{S_{\alpha\beta} \frac{dS^{\alpha}}{dt} \frac{dS^{\beta}}{dt}} dt. \tag{10.91}$$

Therefore, much of what was said in Sect. 5.8 regarding reconstructing the ambient metric tensor from curves lengths can be carried over to surfaces. The surface metric tensor can likewise be reconstructed from the lengths of curves that lie in the surface. Therefore, the metric tensor is *intrinsic*, despite the fact that definition (10.16) is extrinsic since it relies on the covariant basis.

Intrinsic geometry on surfaces is much more exciting than in the ambient space. The interest comes from the fact that different surfaces can have identical intrinsic geometries. For example, a section of a plane can be deformed into a cylinder or a cone without changing the lengths of any of the curves contained within it. The word *deformation* refers to a pointwise mapping between two surfaces and the coordinate systems are such that corresponding points have identical coordinates. A deformation that preserves the lengths of curves is called an *isometric deformation* (or *transformation*) or *an isometry*. Two surfaces related by an isometric deformation are said to be *isometric*. For example, the catenoid and the helicoid shown if Fig. 10.3 are isometric. Isometries are often described informally as distance-preserving transformations. This description captures the spirit of the definition;

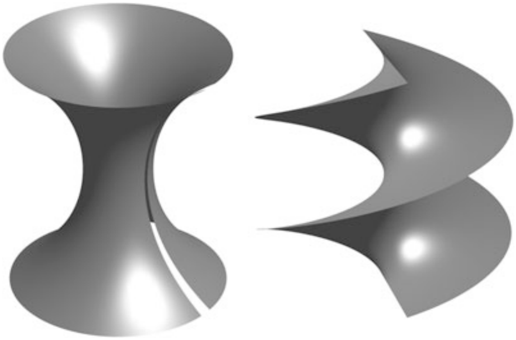


Fig. 10.3 The catenoid and the helicoid are famous for being minimal surfaces discovered by Euler and Meussnier. These surfaces are also famous for being isometric and therefore intrinsically identical. The catenoid is shown with a cut through and along which a discontinuity in the isometric transformation may occur

however, the term *distance*, when used in the context of a curved surface, is vague. We therefore rely on the clear notion of length of curves.

Since the metric tensor can be reconstructed from lengths of curves, isometric surfaces have identical metric tensors at corresponding points. This means that all objects derived from the metric tensor—most notably the upcoming Christoffel symbol and the Riemann–Christoffel tensor—are identical as well. As in the ambient case, these objects are called *intrinsic* and it can be said that isometric surfaces have identical *intrinsic geometries*.

10.10 The Impossibility of Affine Coordinates

Until this chapter, we have been discussing Euclidean spaces. In Chap. 2, while we did not give a formal definition of a Euclidean space, we agreed that we intuitively understood Euclidean spaces as being straight. A related property of a Euclidean space is the possibility of referring it to an affine coordinate system. An affine coordinate system is characterized by a constant metric tensor and a vanishing Christoffel symbol.

On most curved surfaces, it is not possible to introduce an affine coordinate system (i.e., a coordinate system characterized by a vanishing Christoffel symbol), not even on a small patch. In other words, a curved surface is generally not a Euclidean field. So far, we have not developed the necessary analytical machinery to demonstrate this idea. This machinery will be developed over the next two chapters and the impossibility of affine coordinates will be convincingly presented in Chap. 12 on curvature, where a nonvanishing Gaussian curvature will provide irrefutable analytical evidence of this impossibility.

There are several types of surfaces that are exceptions to this rule, namely, surfaces isometric to a plane. Two such surfaces are the cylinder and the cone. What these two surfaces have in common is that we can form them from a flat sheet of paper without sheering or stretching. In other words, we can form them by *isometric deformations*, that is. deformations that do not change distances between points. The discussion of these ideas is the highlight of Chap. 12 and possibly of this entire book.

10.11 Examples of Surfaces

In all examples, the ambient space is referred to Cartesian coordinates x, y, z or x_1, y_1, z_1.

10.11.1 A Sphere of Radius R

Consider a sphere of radius R (Fig. 10.4) given by

$$x(\theta,\phi) = R\sin\theta\cos\phi \tag{10.92}$$

$$y(\theta,\phi) = R\sin\theta\sin\phi \tag{10.93}$$

$$z(\theta,\phi) = R\cos\theta. \tag{10.94}$$

Shift tensor

$$Z_\alpha^i = \begin{bmatrix} R\cos\theta\cos\phi & -R\sin\theta\sin\phi \\ R\cos\theta\sin\phi & R\sin\theta\cos\phi \\ -R\sin\theta & 0 \end{bmatrix} \tag{10.95}$$

Exterior normal

$$N^i = \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}. \tag{10.96}$$

The normal N^i can be obtained as the cross product of the "columns" of Z_α^i. However, in this simple case, we simply guessed the values of N^i from geometric considerations.

Metric tensors

$$S_{\alpha\beta} = \begin{bmatrix} R^2 & 0 \\ 0 & R^2\sin^2\theta \end{bmatrix}; \quad S^{\alpha\beta} = \begin{bmatrix} R^{-2} & 0 \\ 0 & R^{-2}\sin^{-2}\theta \end{bmatrix}. \tag{10.97}$$

Area element

$$\sqrt{S} = R^2\sin\theta. \tag{10.98}$$

Christoffel symbol. Letting $\Theta = 1$ and $\Phi = 2$, the nonzero entries are

$$\Gamma^\Theta_{\Phi\Phi} = -\sin\theta\cos\theta \tag{10.99}$$

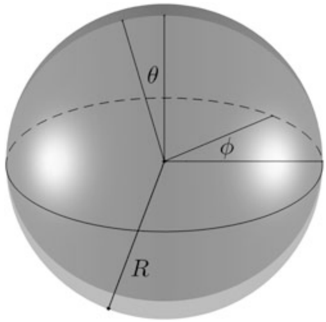


Fig. 10.4 A sphere of radius R referred to a spherical coordinate system

$$\Gamma^{\Phi}_{\Theta\Phi} = \Gamma^{\Phi}_{\Phi\Theta} = \cot\theta. \tag{10.100}$$

10.11.2 A Cylinder of Radius R

An infinite cylinder of radius R (Fig. 10.5) is given by

$$x_1(\theta, z) = R\cos\theta \tag{10.101}$$

$$y_1(\theta, z) = R\sin\theta \tag{10.102}$$

$$z_1(\theta, z) = z. \tag{10.103}$$

Shift tensor and exterior normal

$$Z^i_\alpha = \begin{bmatrix} -R\sin\theta & 0 \\ R\cos\theta & 0 \\ 0 & 1 \end{bmatrix}; \quad N^i = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix} \tag{10.104}$$

Metric tensors

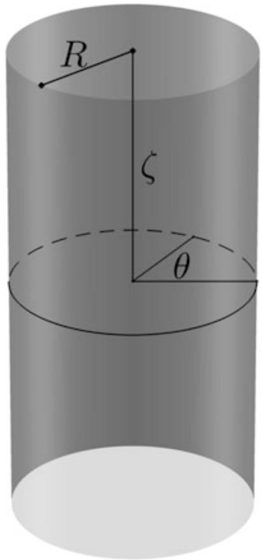
$$S_{\alpha\beta} = \begin{bmatrix} R^2 & 0 \\ 0 & 1 \end{bmatrix}; \quad S^{\alpha\beta} = \begin{bmatrix} R^{-2} & 0 \\ 0 & 1 \end{bmatrix} \tag{10.105}$$

Area element

$$\sqrt{S} = R \tag{10.106}$$

The Christoffel symbol vanishes indicated that the chosen coordinate system is affine.

Fig. 10.5 Infinite cylinder of radius R referred to coordinates $Z^1 = \theta$, $Z^2 = \zeta$



10.11.3 A Torus with Radii R and r

Suppose that the torus (Fig. 10.6) is given by the equations

$$x(\theta,\phi) = (R + r\cos\phi)\cos\theta \tag{10.107}$$

$$y(\theta,\phi) = (R + r\cos\phi)\sin\theta \tag{10.108}$$

$$z(\theta,\phi) = R\sin\phi, \tag{10.109}$$

where $\theta = S^1$ and $\phi = S^2$.

Shift tensor and exterior normal

$$Z_\alpha^i = \begin{bmatrix} -(R + r\cos\phi)\sin\theta & -r\cos\theta\sin\phi \\ (R + r\cos\phi)\cos\theta & -r\sin\theta\sin\phi \\ 0 & r\cos\phi \end{bmatrix}; \quad N^i = \begin{bmatrix} \cos\theta\cos\phi \\ \sin\theta\cos\phi \\ \sin\phi \end{bmatrix} \tag{10.110}$$

Metric tensors

$$S_{\alpha\beta} = \begin{bmatrix} (R + r\cos\phi)^2 & 0 \\ 0 & r^2 \end{bmatrix}; \quad S^{\alpha\beta} = \begin{bmatrix} (R + r\cos\phi)^{-2} & 0 \\ 0 & r^{-2} \end{bmatrix} \tag{10.111}$$

Area element

$$\sqrt{S} = r(R + r\cos\phi) \tag{10.112}$$

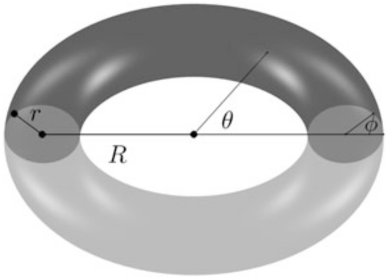


Fig. 10.6 Torus characterized by radii R and r. The surface is referred to coordinates $S^1 = \theta$ and $S^2 = \phi$

Christoffel symbol. Letting $\Theta = 1$ and $\Phi = 2$, the nonzero entries are

$$\Gamma^{\Phi}_{\Theta\Theta} = \frac{R + r\cos\phi}{r} \tag{10.113}$$

$$\Gamma^{\Theta}_{\Theta\Phi} = \Gamma^{\Theta}_{\Phi\Theta} = -\frac{r\sin\phi}{R + r\cos\phi}. \tag{10.114}$$

10.11.4 A Surface of Revolution

Consider a surface of revolution in Cartesian coordinates x_1, y_1, z_1 is given by

$$x_1(\theta, z) = r(z)\cos\theta \tag{10.115}$$

$$y_1(\theta, z) = r(z)\sin\theta \tag{10.116}$$

$$z_1(\theta, z) = z. \tag{10.117}$$

Note that the cylinder (10.101)–(10.103) is a surface of revolution with $r(z) = R$.

Shift tensor and exterior normal

$$Z^i_\alpha = \begin{bmatrix} -r(z)\sin\theta & r'(z)\cos\theta \\ r(z)\cos\theta & r'(z)\sin\theta \\ 0 & 1 \end{bmatrix}; \quad N^i = \begin{bmatrix} \frac{\cos\theta}{\sqrt{1+r'(z)^2}} \\ \frac{\sin\theta}{\sqrt{1+r'(z)^2}} \\ \frac{r'(z)}{\sqrt{1+r'(z)}} \end{bmatrix} \tag{10.118}$$

Metric tensors

$$S_{\alpha\beta} = \begin{bmatrix} r(z)^2 & 0 \\ 0 & 1 + r'(z)^2 \end{bmatrix}; \quad S^{\alpha\beta} = \begin{bmatrix} r(z)^{-2} & 0 \\ 0 & \frac{1}{1+r'(z)^2} \end{bmatrix} \tag{10.119}$$

Area element

$$\sqrt{S} = r(z)\sqrt{1 + r'(z)^2}. \qquad (10.120)$$

Christoffel symbols. Letting $\Theta = 1$ and $Z = 2$, the nonzero entries are

$$\Gamma^{\Theta}_{\Theta Z} = \Gamma^{\Theta}_{Z\Theta} = \frac{r'(z)}{r(z)} \qquad (10.121)$$

$$\Gamma^{Z}_{\Theta\Theta} = \frac{r(z)\,r'(z)}{1 + r'(z)^2} \qquad (10.122)$$

$$\Gamma^{Z}_{ZZ} = \frac{r'(z)\,r''(z)}{1 + r'(z)^2}. \qquad (10.123)$$

Exercise 224. Derive the objects given in this section.

Exercise 225. Rederive the objects on the sphere by referring the ambient space to spherical coordinates r_1, θ_1, ϕ_1 in which the sphere is given by

$$r_1(\theta, \phi) = R \qquad (10.124)$$

$$\theta_1(\theta, \phi) = \theta \qquad (10.125)$$

$$\phi_1(\theta, \phi) = \phi. \qquad (10.126)$$

10.11.5 A Planar Curve in Cartesian Coordinates

A curve embedded in a two-dimensional plane can be viewed as a hypersurface. Suppose that the plane is referred to Cartesian coordinates x, y and let the curve be referred to $S^1 = t$.

Shift tensor and normal

$$Z^i_\alpha = \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix}; \quad N^i = \begin{bmatrix} \frac{y'(t)}{\sqrt{x'(t)^2+y'(t)^2}} \\ -\frac{x'(t)}{\sqrt{x'(t)^2+y'(t)^2}} \end{bmatrix} \qquad (10.127)$$

Metric tensors

$$S_{\alpha\beta} = x'(t)^2 + y'(t)^2; \quad S^{\alpha\beta} = \frac{1}{x'(t)^2 + y'(t)^2} \qquad (10.128)$$

Volume (or length) element

$$\sqrt{S} = \sqrt{x'(t)^2 + y'(t)^2} \tag{10.129}$$

The Christoffel symbol $\Gamma^{\alpha}_{\beta\gamma}$ has a single entry $\tilde{\Gamma}$ with the value

$$\tilde{\Gamma} = \frac{x'(t)\,x''(t) + y'(t)\,y''(t)}{x'(t)^2 + y'(t)^2}. \tag{10.130}$$

Exercise 226. Show that when the curve is referred to the arc length s, the above expression simplifies to the following

$$N^i, = \begin{bmatrix} y'(s) \\ x'(s) \end{bmatrix}; \;\; S_{\alpha\beta} = 1; \;\; S^{\alpha\beta} = 1; \;\; \sqrt{S} = 1; \;\; \tilde{\Gamma} = 0 \tag{10.131}$$

Note that $\tilde{\Gamma}$ simplifies to $x'(s)\,x''(s) + y'(s)\,y''(s)$ and you need to show that this expression vanishes identically.

Exercise 227. Show that when the curve is given as a graph $y = y(x)$, the same differential objects are given by

$$Z^i_\alpha = \begin{bmatrix} 1 \\ g'(x) \end{bmatrix}; \;\; N^i = \begin{bmatrix} \frac{y'(x)}{\sqrt{1+y'(x)^2}} \\ -\frac{1}{\sqrt{1+y'(x)^2}} \end{bmatrix} \tag{10.132}$$

$$S_{\alpha\beta} = 1 + y'(x)^2; \;\; S^{\alpha\beta} = \frac{1}{1 + y'(x)^2} \tag{10.133}$$

$$S = \sqrt{1 + y'(x)^2}; \;\; \tilde{\Gamma} = \frac{y'(x)\,y''(x)}{1 + y'(x)^2} \tag{10.134}$$

10.12 A Planar Curve in Polar Coordinates

Suppose that the plane is referred to Cartesian coordinates (r, θ) and let the curve be referred to $S^1 = t$. Let the curve be given by $(r(t), \theta(t))$.

Shift tensor and normal

$$Z_\alpha^i = \begin{bmatrix} r'(t) \\ \theta'(t) \end{bmatrix}; \quad N^i = \begin{bmatrix} \frac{r(t)\theta'(t)}{\sqrt{r'(t)^2+r(t)^2\theta'(t)^2}} \\ -\frac{r(t)^{-1}r'(t)}{\sqrt{r'(t)^2+r(t)^2\theta'(t)^2}} \end{bmatrix} \tag{10.135}$$

Metric tensors

$$S_{\alpha\beta} = r'(t)^2 + r(t)^2\,\theta'(t)^2; \quad S^{\alpha\beta} = \frac{1}{r'(t)^2 + r(t)^2\,\theta'(t)^2} \tag{10.136}$$

Area element

$$\sqrt{S} = \sqrt{r'(t)^2 + r(t)^2\,\theta'(t)^2} \tag{10.137}$$

The Christoffel symbol $\Gamma^\alpha_{\beta\gamma}$ has a single entry with the value

$$\Gamma^1_{11} = \frac{r(t)\,r'(t)\,\theta'(t)^2 + r'(t)\,r''(t) + r(t)^2\,\theta'(t)\,\theta''(t)}{r'(t)^2 + r(t)^2\,\theta'(t)^2}. \tag{10.138}$$

Exercise 228. Show that when the curve is referred to the arc length s, the above expression simplify to the following

$$Z_\alpha^i = \begin{bmatrix} r'(s) \\ \theta'(s) \end{bmatrix}; \quad N^i = \begin{bmatrix} r(s)\,\theta'(s) \\ -r(s)^{-1}\,r'(s) \end{bmatrix} \tag{10.139}$$

$$S_{\alpha\beta} = 1; \quad S^{\alpha\beta} = 1 \tag{10.140}$$

$$\sqrt{S} = 1; \quad \tilde{\Gamma} = 0 \tag{10.141}$$

Again, note that $\tilde{\Gamma}$ simplifies to $r(s)\,r'(s)\,\theta'(s)^2 + r'(s)\,r''(s) + r(s)^2\,\theta'(s)\,\theta''(s)$ and it remains to show that this expression vanishes identically.

10.13 Summary

This chapter, devoted to the fundamental differential objects on surfaces, parallels Chap. 5 concerned with Euclidean spaces. Curved surfaces add new richness to tensor calculus, which proves time and again that the tensor framework is ideally suited for numerous aspects of differential geometry. This chapter was focused on the tangent and normal spaces as well as on the surface metrics. We discovered a number of important relationships involving the shift tensor and the normal. The most vivid relationships were captured by equations (10.23) and (10.55).

Chapter 11
The Covariant Surface Derivative

11.1 Preview

In the preceding chapter, we gave an overview of embedded surfaces. We now turn to the important question of covariant differentiation on the surface. We will divide the construction of the covariant derivative into two parts. We will first define this operator for objects with surface indices. The definition will be completely analogous to that of the covariant derivative in the ambient space. While the definition will be identical, some of the important characteristics of the surface covariant derivative will be quite different. In particular, surface covariant derivatives do not commute. Our proof of commutativity for the ambient derivative was based on the existence of affine coordinates in Euclidean spaces. Since affine coordinates may not be possible on a curved surface, that argument is no longer available. We will also discover that the surface covariant derivative is not metrinilic with respect to the covariant basis $\mathbf{S}_\alpha$. This will prove fundamental and will give rise to the curvature tensor, which will be further developed in Chap. 12.

Having defined the covariant derivatives for objects with surface indices, we will extend the definition to surface objects with ambient indices, such as the shift tensor Z^i_α and normal N^i. We will discover a chain rule that applies to surface restrictions of objects defined in the ambient space. The chain rule will show that the surface covariant derivative is metrinilic with respect to the ambient metrics. Finally, we will derive formulas for differentiating the shift tensor and the normal.

11.2 The Covariant Derivative for Objects with Surface Indices

We begin our discussion by constructing the covariant derivative as applied to objects with surface components only. These objects include the covariant and the contravariant bases S_α and S^α, metric tensors $S_{\alpha\beta}$ and $S^{\alpha\beta}$, and Levi-Civita symbols

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_11,

$\varepsilon_{\alpha\beta}$ and $\varepsilon^{\alpha\beta}$. For the time being, the covariant derivative will remain undefined for objects with ambient indices, such as the shift tensor Z_α^i and the normal N^i.

Having introduced the Christoffel symbol $\Gamma_{\beta\gamma}^\alpha$, we can define the covariant derivative ∇_γ for a variant T_β^α (with a representative collection of surface indices) by analogy with the ambient case:

$$\nabla_\gamma T_\beta^\alpha = \frac{\partial T_\beta^\alpha}{\partial S^\gamma} + \Gamma_{\gamma\omega}^\alpha T_\beta^\omega - \Gamma_{\gamma\beta}^\omega T_\omega^\alpha. \tag{11.1}$$

On a curved surface, the Christoffel symbol is defined by equation (10.80)

$$\Gamma_{\beta\gamma}^\alpha = \mathbf{S}^\alpha \cdot \frac{\partial \mathbf{S}_\beta}{\partial S^\gamma} \tag{10.80}$$

by analogy with the ambient equation (5.60)

$$\Gamma_{ij}^k = \mathbf{Z}^k \cdot \frac{\partial \mathbf{Z}_i}{\partial Z^j}, \tag{5.60}$$

since the analogy with the ambient equation (5.56)

$$\frac{\partial \mathbf{Z}_i}{\partial Z^j} = \Gamma_{ij}^k \mathbf{Z}_k \tag{5.56}$$

may not be possible: the vector $\partial \mathbf{S}_\beta / \partial S^\gamma$ may very well lie out of the tangent plane due to curvature. This may be seen as the source of all analytical differences between the surface and the ambient covariant derivatives.

The two main differences are (1) surface covariant derivatives do not commute, and (2) the metrinilic property of the surface covariant derivative does not hold for the surface covariant basis. The cause for both of these facts will be seen to be curvature. The replacement identities offer tremendous analytical and geometric insight into the properties of curvature. The discussion of these ideas is saved for Chap. 12, an entire chapter devoted to curvature. In this chapter, curvature and the modified commutativity equation are given but not discussed.

11.3 Properties of the Surface Covariant Derivative

In this section, we give an enumeration of the surviving properties of the covariant derivative. These properties can be demonstrated in the same way as those for the ambient covariant derivative.

1. The covariant derivative produces tensors out of tensors.
2. The covariant derivative ∇_γ coincides with the partial derivative $\partial/\partial S^\gamma$ when applied to variants of order zero.

3. The covariant derivative satisfies the sum and product rules.
4. The metrinilic property of the covariant derivative applies to all metrics, except for $\mathbf{S}_\alpha$ and $\mathbf{S}^\alpha$:

$$\nabla_\gamma S_{\alpha\beta},\ \nabla_\gamma S^{\alpha\beta} = 0 \tag{11.2}$$

$$\nabla_\gamma \varepsilon_{\alpha\beta},\ \nabla_\gamma \varepsilon^{\alpha\beta} = 0 \tag{11.3}$$

$$\nabla_\gamma \delta^{\alpha\beta}_{\sigma\rho},\ \nabla_\gamma \delta^\alpha_\beta = 0 \tag{11.4}$$

5. The covariant derivative commutes with contraction.

Exercise 229. Show the tensor property of the surface covariant derivative.

Exercise 230. Show the sum and product rules for the surface covariant derivative.

Exercise 231. Show the metrinilic property of the surface covariant derivative.

Exercise 232. Show commutativity with contraction of the surface covariant derivative.

11.4 The Surface Divergence and Laplacian

The covariant derivative gives immediate rise to the surface Laplacian $\nabla_\alpha \nabla^\alpha$. This operator is often referred to as the *Laplace–Beltrami operator* or as the *Beltrami operator*. It is particularly interesting that its geometric interpretation is not easy to see. On the other hand, its analytical definition provided by the tensor framework is entirely straightforward.

The Voss–Weyl formula continues to apply. The surface divergence $\nabla_\alpha T^\alpha$ is given by

$$\nabla_\alpha T^\alpha = \frac{1}{\sqrt{S}} \frac{\partial}{\partial S^\alpha} \left(\sqrt{S} T^\alpha \right). \tag{11.5}$$

The Laplacian of an invariant field F is calculated by the formula

$$\nabla_\alpha \nabla^\alpha F = \frac{1}{\sqrt{S}} \frac{\partial}{\partial S^\alpha} \left(\sqrt{S} S^{\alpha\beta} \frac{\partial F}{\partial S^\beta} \right). \tag{11.6}$$

Exercise 233. Show that surface Laplacian on the surface of a sphere (10.92)–(10.94) is given by

$$\nabla_\alpha \nabla^\alpha F = \frac{1}{R^2 \sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial F}{\partial \theta} \right) + \frac{1}{R^2 \sin^2\theta} \frac{\partial^2 F}{\partial \phi^2}. \tag{11.7}$$

Exercise 234. Show that surface Laplacian on the surface of a cylinder (10.101)–(10.103) is given by

$$\nabla_\alpha \nabla^\alpha F = \frac{1}{R^2}\frac{\partial^2 F}{\partial \theta^2} + \frac{\partial^2 F}{\partial z^2}. \tag{11.8}$$

Exercise 235. Show that surface Laplacian on the surface of a torus (10.107)–(10.109) is given by

$$\nabla_\alpha \nabla^\alpha F = \frac{1}{(R + r\cos\phi)^2}\frac{\partial^2 F}{\partial \theta^2} + \frac{1}{r^2 (R + r\cos\phi)}\frac{\partial}{\partial \phi}\left((R + r\cos\phi)\frac{\partial F}{\partial \phi}\right). \tag{11.9}$$

Exercise 236. Show that surface Laplacian on the surface of revolution (10.115)–(10.117) is given by

$$\nabla_\alpha \nabla^\alpha F = \frac{1}{r(z)\sqrt{1 + r'(z)^2}}\frac{\partial}{\partial z}\left(\frac{r(z)}{\sqrt{1 + r'(z)^2}}\frac{\partial F}{\partial z}\right) + \frac{1}{r(z)^2}\frac{\partial^2 F}{\partial \theta^2} \tag{11.10}$$

11.5 The Curvature Tensor

The curvature tensor arises out of the failure of the metrinilic property with respect to the covariant basis. Expand $\nabla_\alpha \mathbf{S}_\beta$ according to definition (11.1):

$$\nabla_\alpha \mathbf{S}_\beta = \frac{\partial \mathbf{S}_\beta}{\partial S^\alpha} - \Gamma^\gamma_{\alpha\beta}\mathbf{S}_\gamma. \tag{11.11}$$

Note that $\nabla_\alpha \mathbf{S}_\beta$ is symmetric

$$\nabla_\alpha \mathbf{S}_\beta = \nabla_\beta \mathbf{S}_\alpha, \tag{11.12}$$

since the first term in (11.11) equals $\partial^2 \mathbf{R}/\partial S^\alpha \partial S^\beta$ and the second term is proportional to the Christoffel symbol $\Gamma^\gamma_{\alpha\beta}$ which is symmetric in its lower indices.

Unlike the quantity $\nabla_i \mathbf{Z}_j$, which vanished by definition (5.56), the quantity $\nabla_\alpha \mathbf{S}_\beta$ does not vanish. After all, a direct analogue of equation (5.56) does not exist due to the curvature of the surface. Nevertheless, a geometric observation of great value can be made of $\nabla_\alpha \mathbf{S}_\beta$: each of these four vectors is orthogonal to the tangent plane. To put it another way, while the vector $\nabla_\alpha \mathbf{S}_\beta$ does not vanish, its tangential components do. To show this, dot the identity (11.11) with $\mathbf{S}^\delta$:

$$\mathbf{S}^\delta \cdot \nabla_\alpha \mathbf{S}_\beta = \mathbf{S}^\delta \cdot \frac{\partial \mathbf{S}_\beta}{\partial S^\alpha} - \Gamma^\gamma_{\alpha\beta}\mathbf{S}_\gamma \cdot \mathbf{S}^\delta. \tag{11.13}$$

Since $\mathbf{S}_\gamma \cdot \mathbf{S}^\delta = \delta^\delta_\gamma$, we find

$$\mathbf{S}^\delta \cdot \nabla_\alpha \mathbf{S}_\beta = \mathbf{S}^\delta \cdot \frac{\partial \mathbf{S}_\beta}{\partial S^\alpha} \tag{11.14}$$

and the right-hand side vanishes by the definition (10.80). Thus,

$$\mathbf{S}^\delta \cdot \nabla_\alpha \mathbf{S}_\beta = 0 \tag{11.15}$$

and the orthogonality of $\nabla_\alpha \mathbf{S}_\beta$ to the tangent plane is established.

This geometric observation leads to the definition of the curvature tensor. Since $\nabla_\alpha \mathbf{S}_\beta$ is normal to the plane, it must be a multiple of the normal vector $\mathbf{N}$. Let that multiple be $B_{\alpha\beta}$:

$$\nabla_\alpha \mathbf{S}_\beta = \mathbf{N} B_{\alpha\beta}. \tag{11.16}$$

The variant $B_{\alpha\beta}$ is called the *curvature tensor*. Dotting both sides of this identity with $\mathbf{N}$ yields an explicit expression for $B_{\alpha\beta}$

$$B_{\alpha\beta} = \mathbf{N} \cdot \nabla_\alpha \mathbf{S}_\beta \tag{11.17}$$

and confirms its tensor property. Note that the sign of $B_{\alpha\beta}$ depends on the choice of the normal.

The curvature tensor plays a central role in the analysis of embedded surfaces. So important is its role that we have devoted an entire chapter to its study. Having shaken hands with it in this chapter, we part with the curvature tensor until our in-depth investigation in Chap. 12.

11.6 Loss of Commutativity

Recall the argument by which we proved that ambient covariant derivatives commute. We evaluated the commutator

$$\left(\nabla_i \nabla_j - \nabla_j \nabla_i\right) T^k \tag{11.18}$$

in affine coordinates where the covariant derivatives coincide with partial derivatives and therefore commute. As a result, the commutator vanishes in affine coordinates and, being a tensor, in all coordinates.

This argument is not available on curved surfaces since there may not exist an affine coordinate system. As a result, commutativity is lost. Instead, the commutator is governed by the Riemann–Christoffel tensor $R^{\gamma}_{\cdot\delta\alpha\beta}$, introduced in the ambient space in Chap. 8. It is straightforward to show that for a contravariant tensor T^γ, the commutator takes the following form

$$\left(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha\right) T^\gamma = R^\gamma_{\cdot\delta\alpha\beta} T^\delta, \tag{11.19}$$

where

$$R^\gamma_{\cdot\delta\alpha\beta} = \frac{\partial \Gamma^\gamma_{\beta\delta}}{\partial S^\alpha} - \frac{\partial \Gamma^\gamma_{\alpha\delta}}{\partial S^\beta} + \Gamma^\gamma_{\alpha\omega}\Gamma^\omega_{\beta\delta} - \Gamma^\gamma_{\beta\omega}\Gamma^\omega_{\alpha\delta}. \tag{11.20}$$

Example 237. Derive equations (11.19) and (11.20).

An in-depth discussion of the Riemann–Christoffel tensor takes place in Chap. 12.

11.7 The Covariant Derivative for Objects with Ambient Indices

11.7.1 Motivation

The clue to the covariant derivative for objects with ambient indices comes from differentiating an invariant field $\mathbf{T}$ (such as the surface normal $\mathbf{N}$) and its components T^i. We are looking for an operator that satisfies the product rule and coincides with the partial derivative when applied to variants of order zero. For an operator ∇_γ that satisfies these properties, the following chain of identities would be valid:

$$\nabla_\gamma \mathbf{T} = \frac{\partial \mathbf{T}}{\partial S_\gamma} = \frac{\partial\left(T^i \mathbf{Z}_i\right)}{\partial S_\gamma} = \frac{\partial T^i}{\partial S_\gamma}\mathbf{Z}_i + T^i \frac{\partial \mathbf{Z}_i}{\partial S_\gamma} \tag{11.21}$$

Apply the chain rule to the second term

$$\nabla_\gamma \mathbf{T} = \frac{\partial T^i}{\partial S_\gamma}\mathbf{Z}_i + T^i \frac{\partial \mathbf{Z}_i}{\partial Z^j}\frac{\partial Z^j}{\partial S^\gamma} \tag{11.22}$$

and recognize that $\partial \mathbf{Z}_i / \partial Z^j = \Gamma^k_{ij}\mathbf{Z}_k$ and $\partial Z^j / \partial S^\gamma = Z^j_\gamma$:

$$\nabla_\gamma \mathbf{T} = \frac{\partial T^i}{\partial S_\gamma}\mathbf{Z}_i + T^i Z^j_\gamma \Gamma^k_{ij}\mathbf{Z}_k. \tag{11.23}$$

The right-hand side can be rewritten as

$$\nabla_\gamma \mathbf{T} = \left(\frac{\partial T^i}{\partial S_\gamma} + Z^k_\gamma \Gamma^i_{km} T^m\right)\mathbf{Z}_i. \tag{11.24}$$

Finally, if we demand that ∇_γ is metrinilic with respect to $\mathbf{Z}_i$, the left-hand side can be transformed by the product rule

$$\nabla_\gamma T^i \, \mathbf{Z}_i = \left(\frac{\partial T^i}{\partial S_\gamma} + Z_\gamma^k \Gamma_{km}^i T^m \right) \mathbf{Z}_i. \tag{11.25}$$

This identity suggest the definition

$$\nabla_\gamma T^i = \frac{\partial T^i}{\partial S_\gamma} + Z_\gamma^k \Gamma_{km}^i T^m \tag{11.26}$$

and indeed this good guess leads to a definition that has all of the desired properties.

11.7.2 *The Covariant Surface Derivative in Full Generality*

Motivated by equation (11.26), we give the following definition for the covariant surface derivative applied to a variant T_j^i with ambient indices:

$$\nabla_\gamma T_j^i = \frac{\partial T_j^i}{\partial S_\gamma} + Z_\gamma^k \Gamma_{km}^i T_j^m - Z_\gamma^k \Gamma_{kj}^m T_m^i. \tag{11.27}$$

To give the ultimate definition, consider a variant $T_{j\beta}^{i\alpha}$ with a fully representative collection of indices. That definition reads

$$\nabla_\gamma T_{j\beta}^{i\alpha} = \frac{\partial T_{j\beta}^{i\alpha}}{\partial S_\gamma} + Z_\gamma^k \Gamma_{km}^i T_{j\beta}^{m\alpha} - Z_\gamma^k \Gamma_{kj}^m T_{m\beta}^{i\alpha} + \Gamma_{\Gamma\omega}^{\alpha} T_{j\beta}^{i\omega} - \Gamma_{\gamma\beta}^{\omega} T_{j\omega}^{i\alpha}. \tag{11.28}$$

We leave it as an exercise to prove each of the following fundamental properties.

1. The covariant derivative produces tensors out of tensors.
2. The covariant derivative ∇_γ coincides with the partial derivative $\partial/\partial S^\gamma$ when applied to invariants (or any variants of order zero).
3. The covariant derivative satisfies the sum and product rules.
4. The metrinilic property of the covariant derivative applies to all metrics, except for $\mathbf{S}_\alpha$ and $\mathbf{S}^\alpha$:

$$\nabla_\gamma S_{\alpha\beta}, \ \nabla_\gamma S^{\alpha\beta} = 0 \tag{11.29}$$

$$\nabla_\gamma \varepsilon_{\alpha\beta}, \ \nabla_\gamma \varepsilon^{\alpha\beta} = 0 \tag{11.30}$$

$$\nabla_\gamma \delta_{\sigma\rho}^{\alpha\beta}, \ \nabla_\gamma \delta_\beta^\alpha = 0 \tag{11.31}$$

5. The covariant derivative commutes with contraction.

11.8 The Chain Rule

Consider a variant field T^i_j defined in the ambient space. We may also consider the restriction of T^i_j to the surface S and apply the covariant derivative to this restriction by definition (11.27)

$$\nabla_\gamma T^i_j = \frac{\partial T^i_j}{\partial S_\gamma} + Z^k_\gamma \Gamma^i_{km} T^m_j - Z^k_\gamma \Gamma^m_{kj} T^i_m. \tag{11.32}$$

Being a restriction of an ambient field, the function $T^i_j(S)$ can be represented by the composition $T^i_j(S) = T^i_j(Z(S))$. Therefore, the partial derivative term can be transformed by the chain rule

$$\nabla_\gamma T^i_j = \frac{\partial T^i_j}{\partial Z^k}\frac{\partial Z^k}{\partial S^\gamma} + Z^k_\gamma \Gamma^i_{km} T^m_j - Z^k_\gamma \Gamma^m_{kj} T^i_m. \tag{11.33}$$

The object $\partial Z^m / \partial S^\gamma$ is recognized as the shift tensor Z^m_γ, which can now be factored out from each term

$$\nabla_\gamma T^i_j = Z^k_\gamma \left(\frac{\partial T^i_j}{\partial Z^k} + \Gamma^i_{km} T^m_j - \Gamma^m_{kj} T^i_m \right). \tag{11.34}$$

The quantity in parentheses is recognized as $\nabla_k T^i_j$ and we conclude that

$$\nabla_\gamma T^i_j = Z^k_\gamma \nabla_k T^i_j. \tag{11.35}$$

This elegant relationship is known as the *chain rule*. It states that the covariant surface derivative of the restriction of an ambient field is the projection of the ambient covariant derivative.

The chain rule is ubiquitous in tensor calculations and has the following immediate implication: the covariant surface derivative is metrinilic with respect to all ambient metrics:

$$\nabla_\gamma \mathbf{Z}_i, \ \nabla_\gamma \mathbf{Z}^i = \mathbf{0} \tag{11.36}$$

$$\nabla_\gamma Z_{ij}, \ \nabla_\gamma Z^{ij} = 0 \tag{11.37}$$

$$\nabla_\gamma \varepsilon_{ijk}, \ \nabla_\gamma \varepsilon^{ijk} = 0 \tag{11.38}$$

$$\nabla_\gamma \delta^i_j, \ \nabla_\gamma \delta^{ij}_{rs}, \ \nabla_\gamma \delta^{ijk}_{rst} = 0. \tag{11.39}$$

One of the important implications of the metrinilic property (11.37) is that *ambient* indices can be juggled “across” the covariant derivative. For example, $S_i \nabla_\gamma T^i = S^i \nabla_\gamma T_i$, and $S_{i\gamma} = \nabla_\gamma T_i$ implies $S^i_\gamma = \nabla_\gamma T^i$.

Exercise 238. Justify equations (11.36)–(11.39).

11.9 The Formulas for $\nabla_\alpha Z^i_\beta$ and $\nabla_\alpha N^i$

There are two outstanding elements on the surface, the shift tensor Z^i_β and the normal N^i, for which we have not yet established differentiation rules. The metrinilic property (11.36) of the covariant derivative yields the rule for the shift tensor. Recall the definition (11.16) of the curvature tensor:

$$\nabla_\alpha \mathbf{S}_\beta = \mathbf{N} B_{\alpha\beta}. \tag{11.16}$$

By the product rule and the metrinilic property, we find:

$$\nabla_\alpha \mathbf{S}_\beta = \nabla_\alpha \left(Z^i_\beta \mathbf{Z}_i \right) = \nabla_\alpha Z^i_\beta \, \mathbf{Z}_i. \tag{11.40}$$

Decompose the normal $\mathbf{N}$ on the right-hand side of equation (11.16) with respect to $\mathbf{Z}_i$

$$\nabla_\alpha Z^i_\beta \, \mathbf{Z}_i = N^i B_{\alpha\beta} \mathbf{Z}_i. \tag{11.41}$$

Equating the vector components gives the desired formula

$$\nabla_\alpha Z^i_\beta = N^i B_{\alpha\beta}. \tag{11.42}$$

Contracting both sides of this equation with N_i gives an explicit expression for the curvature tensor:

$$B_{\alpha\beta} = N_i \nabla_\alpha Z^i_\beta. \tag{11.43}$$

We now turn to the normal N^i. One way to derive the expression for $\nabla_\alpha N^i$ is by appealing to the explicit expression (10.38) for N^i. Alternatively, we can appeal to the two equations (10.39) and (10.41) that define N^i implicitly. We pursue the latter approach first.

Applying the covariant derivative to equation (10.41)

$$N_i N^i = 1, \tag{10.41}$$

we find

$$N_i \nabla_\alpha N^i = 0. \tag{11.44}$$

Next, juggle the index i in equation (10.39)

$$N^i Z_{i\alpha} = 0 \tag{11.45}$$

and apply the covariant derivative. By the product rule,

$$\nabla_\beta N^i \; Z_{i\alpha} + N^i \nabla_\beta Z_{i\alpha} = 0. \tag{11.46}$$

By equation (11.43), the second term is the curvature tensor $B_{\alpha\beta}$. Thus,

$$\nabla_\beta N^i \; Z_{i\alpha} = -B_{\alpha\beta}. \tag{11.47}$$

To isolate $\nabla_\beta N^i$, contract both sides with $Z^{\alpha j}$

$$\nabla_\beta N^i \; Z_{i\alpha} Z^{\alpha j} = -Z^{\alpha j} B_{\alpha\beta}. \tag{11.48}$$

By the projection equation (10.55), the left-hand side is

$$\nabla_\beta N^j - N_i N^j \nabla_\beta N^i = -Z^\alpha_j B_{\alpha\beta} \tag{11.49}$$

Since the second term on the left-hand side vanishes by equation (11.44), we arrive at the ultimate expression for $\nabla_\alpha N^i$

$$\nabla_\alpha N^i = -Z^i_\beta B^\beta_\alpha. \tag{11.50}$$

Equation (11.50) is also known as *Weingarten's formula*.

Exercise 239. Derive equation (11.44) from (10.41).

We now derive Weingarten's formula (11.50) from the explicit expression (10.38) for the normal. We leave as an exercise to derive the following identity

$$\varepsilon^{ijk} \varepsilon_{\alpha\beta} Z^\beta_j N_k = Z^i_\alpha. \tag{11.51}$$

The geometric interpretation of this identity is straightforward: the cross-product-like combination of the normal and a covariant basis element yields a tensor orthogonal to both, that is, a corresponding contravariant basis element. This fundamental relationship is, surprisingly, not needed often, but finds a direct application in the following derivation.

Exercise 240. Confirm equation (11.51) by substituting an explicit expression for N_k.

An application of the covariant derivative to equation (10.38)

$$N^i = \frac{1}{2} \varepsilon^{ijk} \varepsilon_{\alpha\beta} Z^\alpha_j Z^\beta_k \tag{10.38}$$

yields

$$\nabla_\gamma N^i = \frac{1}{2}\varepsilon^{ijk}\varepsilon_{\alpha\beta}\left(\nabla_\gamma Z_j^\alpha \ Z_k^\beta + Z_j^\alpha \nabla_\gamma Z_k^\beta\right). \tag{11.52}$$

The terms in parentheses yield curvature tensors by equation (11.42):

$$\nabla_\gamma N^i = \frac{1}{2}\varepsilon^{ijk}\varepsilon_{\alpha\beta}\left(N_j B_\gamma^\alpha Z_k^\beta + Z_j^\alpha N_k B_\gamma^\beta\right). \tag{11.53}$$

Multiply out the right-hand side

$$\nabla_\gamma N^i = \frac{1}{2}\varepsilon^{ijk}\varepsilon_{\alpha\beta} N_j B_\gamma^\alpha Z_k^\beta + \frac{1}{2}\varepsilon^{ijk}\varepsilon_{\alpha\beta} Z_j^\alpha N_k B_\gamma^\beta. \tag{11.54}$$

The two terms are actually the same, which can be seen after a few index-renaming steps. Therefore,

$$\nabla_\gamma N^i = \varepsilon^{ijk}\varepsilon_{\alpha\beta} N_j B_\gamma^\alpha Z_k^\beta \tag{11.55}$$

and Weingarten's formula (11.50) follows by an application of equation (11.51).

11.10 The Normal Derivative

Consider an invariant function u defined in the ambient space. At the points on an embedded surface S, we may be interested in the directional derivative of u along the normal N^k. The directional derivative in this special direction is called the *normal derivative* and is often denoted by $\partial u/\partial n$. According to equation (2.11), the normal derivative is given by

$$\frac{\partial u}{\partial n} = N^i \nabla_i u. \tag{11.56}$$

The normal derivative is a ubiquitous concept in applied mathematics, physics, and engineering.

According to equation (11.56), the normal derivative $\partial u/\partial n$ can be thought of as the normal component of the ambient gradient $\nabla_i u$. Suppose that normal derivative $\partial u/\partial n$ and the covariant surface derivative $\nabla_\alpha u$ are known at a particular point on S. Then the full ambient derivative $\nabla_i u$ can be reconstructed by the formula

$$\nabla_i u = Z_i^\alpha \nabla_\alpha u + N_i \frac{\partial u}{\partial n}. \tag{11.57}$$

To demonstrate equation (11.57), rename the index i in definition (11.56) and multiply both sides by N_i,

$$N_i \frac{\partial u}{\partial n} = N_i N^j \nabla_j u. \tag{11.58}$$

By the projection rule (10.55),

$$N_i \frac{\partial u}{\partial n} = \left(\delta_i^j - Z_i^\alpha Z_\alpha^j\right) \nabla_j u. \tag{11.59}$$

Expanding, we find by the chain rule

$$N_i \frac{\partial u}{\partial n} = \nabla_i u - Z_i^\alpha \nabla_\alpha u. \tag{11.60}$$

which is equivalent to equation (11.57).

A number of beautiful relationships exist that relate the second-order ambient and surface covariant derivatives of an ambient variant. Perhaps the most commonly encountered one is

$$N^i N^j \nabla_i \nabla_j u = \nabla_i \nabla^i u - \nabla_\alpha \nabla^\alpha u + B_\alpha^\alpha N^i \nabla_i u. \tag{11.61}$$

The derivation of this equation is left as an exercise. Equation (11.61) has a number of valuable interpretations. One of them is the relationship of the surface Laplacian $\nabla_\alpha \nabla^\alpha u$ to the ambient Laplacian $\nabla_i \nabla^i u$. This perspective can be brought to the forefront by writing equation (11.61) in the form

$$\nabla_\alpha \nabla^\alpha u = \nabla_i \nabla^i u - N^i N^j \nabla_i \nabla_j u + B_\alpha^\alpha N^i \nabla_i u. \tag{11.62}$$

Exercise 241. Derive equation (11.61). Hint: Transform the surface Laplacian $\nabla_\alpha \nabla^\alpha u$ by the chain rule (11.35) starting with the innermost covariant derivative ∇^α.

Exercise 242. Show that

$$\frac{\partial^2}{\partial n^2} = N^i N^j \nabla_i \nabla_j. \tag{11.63}$$

You may find it helpful to refer the ambient space to an affine coordinate system.

11.11 Summary

In this chapter, we constructed the covariant derivative on embedded surfaces. When applied to objects with surface indices, the definition of the covariant derivative was analogous to that in the ambient Euclidean space. However, since a complete analogy in the definition of the Christoffel symbol is not possible, two properties—the metrinilic property and commutativity—had to be modified. Both alterations are associated with curvature. The loss of the metrinilic property with respect to the covariant basis led to the introduction of the curvature basis. Loss of commutativity led to the nontrivial Riemann–Christoffel tensor and an alternative introduction of Gaussian curvature.

Chapter 12
Curvature

12.1 Preview

This chapter is devoted to the study of curvature. This is an exciting and beautiful topic. The highlight of this chapter, and perhaps the entire book is Gauss's *Theorema Egregium,* or the *Remarkable Theorem*.

The first half of this chapter is devoted to the intrinsic analysis of surfaces. Intrinsic analysis deals with quantities derived from the metric tensor and its partial derivatives. The central object in the intrinsic perspective in the Riemann–Christoffel tensor $R_{\alpha\beta\gamma\delta}$ that arises in the analysis of the commutator $\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha$. The second half of the chapter focuses on the curvature tensor $B_{\alpha\beta}$, an extrinsic object because it depends on the way in which the surface is embedded in the ambient space. Theorema Egregium relates the Riemann–Christoffel and the curvature tensor and thus links the intrinsic and the extrinsic perspectives.

It is exciting to have the opportunity to come face to face with Carl Friedrich Gauss (Fig. 12.1) and his protege Bernhard Riemann, two of the greatest mathematicians in history. These amazing scientists spent their entire lives pondering the concept of space and curvature and attained a depth of understanding unmatched to this day. The equations contained in this chapter flow effortlessly from one to the next and their beauty is a tribute to the remarkable achievements of those great men and their predecessors.

12.2 The Riemann–Christoffel Tensor

We begin with a discussion of the intrinsic elements of the surface. In the previous chapter, we undertook the critical question of whether covariant surface derivatives commute. In this chapter, we discover that the general answer is no.

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_12,

Fig. 12.1 Carl Friedrich Gauss (1777–1855) was a pivotal figure in differential geometry and mathematics in general

The commutator $\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha$ applied to a variant T^γ with a surface superscript is governed by the rule

$$\left(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha\right) T^\gamma = R^{\gamma}_{\cdot\delta\alpha\beta} T^\delta, \tag{11.19}$$

where the *Riemann-Christoffel tensor* $R^{\gamma}_{\cdot\delta\alpha\beta}$ is given by

$$R^{\gamma}_{\cdot\delta\alpha\beta} = \frac{\partial \Gamma^{\gamma}_{\beta\delta}}{\partial S^\alpha} - \frac{\partial \Gamma^{\gamma}_{\alpha\delta}}{\partial S^\beta} + \Gamma^{\gamma}_{\alpha\omega}\Gamma^{\omega}_{\beta\delta} - \Gamma^{\gamma}_{\beta\omega}\Gamma^{\omega}_{\alpha\delta}. \tag{11.20}$$

Introduce the Riemann–Christoffel tensor $R_{\gamma\delta\alpha\beta}$ with the lowered first index

$$R_{\gamma\delta\alpha\beta} = S_{\gamma\omega} R^{\omega}_{\cdot\delta\alpha\beta}. \tag{12.1}$$

Exercise 243. Show that

$$R_{\gamma\delta\alpha\beta} = \frac{\partial \Gamma_{\gamma,\beta\delta}}{\partial S^\alpha} - \frac{\partial \Gamma_{\gamma,\alpha\delta}}{\partial S^\beta} + \Gamma_{\omega,\gamma\beta}\Gamma^{\omega}_{\alpha\delta} - \Gamma_{\omega,\gamma\alpha}\Gamma^{\omega}_{\beta\delta}. \tag{12.2}$$

In the ambient Euclidean space, the analogous expressions are found in equations (8.131) and (8.130). However, due to the Euclidean nature of the ambient space, the ambient Riemann–Christoffel tensor vanishes and therefore covariant derivatives commute. The vanishing of the Riemann–Christoffel tensor is shown in two ways. The first approach relies on the evaluation of the Riemann–Christoffel tensor in Cartesian coordinates. The second approach, outlined in Exercise 146, relies on substituting an explicit expression for the Christoffel symbol into the definition of the Riemann–Christoffel tensor. Neither approach carries over to curved surfaces: the surface may not admit a Cartesian coordinate system and the partial derivative of the contravariant basis may have a normal component and therefore may not be re-expressed in the contravariant basis.

From the definition (11.19) and equation (11.20), it can be easily seen that $R_{\gamma\delta\alpha\beta}$ is antisymmetric in α and β

$$R_{\gamma\delta\beta\alpha} = -R_{\gamma\delta\alpha\beta}, \tag{12.3}$$

as well as γ and δ

$$R_{\delta\gamma\alpha\beta} = -R_{\gamma\delta\alpha\beta}. \tag{12.4}$$

A little bit more work is required to show the $R_{\alpha\beta\gamma\delta}$ is symmetric with respect to switching the pairs of indices $\alpha\beta$ and $\gamma\delta$:

$$R_{\gamma\delta\alpha\beta} = R_{\alpha\beta\gamma\delta}. \tag{12.5}$$

Exercise 244. Show that

$$R^{\alpha}_{\cdot\alpha\gamma\delta} = 0. \tag{12.6}$$

Exercise 245. Show relationships (12.3)–(12.5).

Exercise 246. Show that equation (12.4) follows from equations (12.3) and (12.5).

Exercise 247. The Ricci tensor $R_{\alpha\beta}$ is defined as

$$R_{\alpha\beta} = R^{\gamma}_{\cdot\alpha\gamma\beta}. \tag{12.7}$$

Its trace

$$R = R^{\alpha}_{\alpha} \tag{12.8}$$

is called the scalar curvature. Show that the Ricci tensor is symmetric.

Exercise 248. The Einstein tensor $G_{\alpha\beta}$ is defined by

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2} R S_{\alpha\beta}. \tag{12.9}$$

Show that the Einstein tensor is symmetric.

Exercise 249. Show that, for two-dimensional manifolds, the trace G^{α}_{α} of the Einstein tensor $G_{\alpha\beta}$ vanishes.

Exercise 250. Show that for a covariant tensor T_{γ}, the commutator relationship reads

$$\left(\nabla_{\alpha}\nabla_{\beta} - \nabla_{\beta}\nabla_{\alpha}\right) T_{\gamma} = -R^{\delta}_{\cdot\gamma\alpha\beta} T_{\delta}. \tag{12.10}$$

Exercise 251. Show that covariant derivatives commute when applied to invariants and variants with ambient indices

$$\left(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha\right) T = 0 \tag{12.11}$$

$$\left(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha\right) T^i = 0 \tag{12.12}$$

$$\left(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha\right) T_i = 0. \tag{12.13}$$

Exercise 252. Show the first Bianchi identity

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\gamma\delta\beta} + R_{\alpha\delta\beta\gamma} = 0. \tag{12.14}$$

Exercise 253. Show the second Bianchi identity

$$\nabla_\varepsilon R_{\alpha\beta\gamma\delta} + \nabla_\gamma R_{\alpha\beta\delta\varepsilon} + \nabla_\delta R_{\alpha\beta\varepsilon\gamma} = 0. \tag{12.15}$$

Equations (12.12) and (12.13) can be shown rather easily by a direct application of the definition of the covariant derivative. However, there is a more elegant way of showing this. For a variant T^i, form the vector $\mathbf{T} = T^i \mathbf{Z}_i$. By equation (12.11),

$$\nabla_\alpha \nabla_\beta \mathbf{T} - \nabla_\beta \nabla_\alpha \mathbf{T} = \mathbf{0}. \tag{12.16}$$

Since the surface covariant derivative is metrinilic with respect the ambient basis $\mathbf{Z}_i$, we have

$$\left(\nabla_\alpha \nabla_\beta T^i - \nabla_\beta \nabla_\alpha T^i\right) \mathbf{Z}_i = \mathbf{0} \tag{12.17}$$

and equation (12.12) follows.

So far we have analyzed the commutator $\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha$ for variants of order one. Applying the commutator to variants of order greater than one results in a sum of appropriate Riemann–Christoffel terms for each surface index. For example, for a tensor $T^{i\gamma}_{j\delta}$ with a representative collection of indices, the commutator is given by

$$\left(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha\right) T^{i\gamma}_{j\delta} = R^{\gamma}_{\cdot\omega\alpha\beta} T^{i\omega}_{j\delta} - R^{\omega}_{\cdot\delta\alpha\beta} T^{i\gamma}_{j\omega}. \tag{12.18}$$

Equation (12.18) shows that, from the index manipulation point of view, the role of the Riemann–Christoffel tensor in the commutator is similar to that of the Christoffel symbol in the covariant derivative: superscripts are contracted with the *second* index of $+R^{\gamma}_{\cdot\delta\alpha\beta}$ and subscripts are contracted with the *first* index of $-R^{\delta}_{\cdot\gamma\alpha\beta}$.

To show equation (12.18) for a variant T^{γ}_{δ} (dropping the less interesting ambient indices), define

$$T^{\gamma} = T^{\gamma}_{\delta} S^{\delta}, \tag{12.19}$$

where S^{ω} is an *arbitrary* variant. Since

$$\nabla_\alpha \nabla_\beta T^\gamma = \nabla_\alpha \nabla_\beta T^\gamma_\delta S^\delta + \nabla_\beta T^\gamma_\delta \nabla_\alpha S^\delta + \nabla_\alpha T^\gamma_\delta \nabla_\beta S^\delta + T^\gamma_\delta \nabla_\alpha \nabla_\beta S^\delta \tag{12.20}$$

and the sum of the middle two terms on the right-hand side is symmetric in α and β, we find

$$\nabla_\alpha \nabla_\beta T^\gamma - \nabla_\beta \nabla_\alpha T^\gamma = \left(\nabla_\alpha \nabla_\beta T^\gamma_\delta - \nabla_\beta \nabla_\alpha T^\gamma_\delta\right) S^\delta + T^\gamma_\delta \left(\nabla_\alpha \nabla_\beta S^\delta - \nabla_\beta \nabla_\alpha S^\delta\right), \tag{12.21}$$

or

$$\left(\nabla_\alpha \nabla_\beta T^\gamma_\delta - \nabla_\beta \nabla_\alpha T^\gamma_\delta\right) S^\delta = \Gamma^{\gamma}_{\cdot\omega\alpha\beta} T^\omega - \Gamma^{\delta}_{\cdot\omega\alpha\beta} S^\omega T^\gamma_\delta. \tag{12.22}$$

Switching the names of the indices γ and ω in the last term, we find

$$\left(\nabla_\alpha \nabla_\beta T^\gamma_\delta - \nabla_\beta \nabla_\alpha T^\gamma_\delta\right) S^\delta = \Gamma^{\gamma}_{\cdot\omega\alpha\beta} T^\omega_\delta S^\delta - \Gamma^{\omega}_{\cdot\delta\alpha\beta} T^\gamma_\omega S^\delta. \tag{12.23}$$

Since S^{ω} is arbitrary, equation (12.18) is confirmed.

12.3 The Gaussian Curvature

For a two-dimensional surface, equations (12.3)–(12.5) show that the Riemann–Christoffel symbol has a single degree of freedom R_{1212} and the four nonzero entries are $\pm R_{1212}$. Thus, the Riemann–Christoffel tensor can be captured by the formula

$$R_{\gamma\delta\alpha\beta} = R_{1212} e_{\gamma\delta} e_{\alpha\beta}. \tag{12.24}$$

Switching from permutation systems to Levi-Civita symbols, we find

$$R_{\gamma\delta\alpha\beta} = \frac{R_{1212}}{S} \varepsilon_{\gamma\delta} \varepsilon_{\alpha\beta}. \tag{12.25}$$

The invariant quantity

$$K = \frac{R_{1212}}{S} \tag{12.26}$$

is known as the *Gaussian curvature* K. It is one of the most fundamental and interesting characteristics of a surface. In terms of the Gaussian curvature K, the Riemann–Christoffel tensor is given by

$$R_{\gamma\delta\alpha\beta} = K \varepsilon_{\gamma\delta} \varepsilon_{\alpha\beta}. \tag{12.27}$$

Exercise 254. Derive the following explicit expression for K:

$$K = \frac{1}{4}\varepsilon^{\gamma\delta}\varepsilon^{\alpha\beta}R_{\gamma\delta\alpha\beta}. \tag{12.28}$$

In particular, equation (12.28) shows that K is a tensor.

Exercise 255. Show that

$$R_{\gamma\delta\alpha\beta} = K\left(S_{\alpha\gamma}S_{\beta\delta} - S_{\alpha\delta}S_{\beta\gamma}\right). \tag{12.29}$$

Exercise 256. Show that the Gaussian curvature K is also given by

$$K = \frac{1}{2}R^{\alpha\beta}_{\cdot\cdot\alpha\beta}. \tag{12.30}$$

Equivalently,

$$K = \frac{1}{2}R^{\alpha}_{\alpha}, \tag{12.31}$$

where $R_{\alpha\beta}$ is the Ricci tensor. In other words, for a two-dimensional surface, the Gaussian curvature is half of the scalar curvature.

12.4 The Curvature Tensor

We now turn to the extrinsic elements on the surface. The primary extrinsic object is the curvature tensor $B_{\alpha\beta}$ (also known as the *extrinsic curvature tensor*) introduced in Chap. 11 in equation (11.16):

$$\nabla_\alpha \mathbf{S}_\beta = \mathbf{N}B_{\alpha\beta}. \tag{11.16}$$

In component form, this equation reads

$$\nabla_\alpha Z^i_\beta = N^i B_{\alpha\beta}. \tag{12.32}$$

The combination $N^i B^\alpha_\alpha$ is called the *curvature normal*. The vector invariant $\mathbf{N}B^\alpha_\alpha$ may be referred to as the *vector curvature normal*.

Equation (11.16) can be written in the following elegant way

$$\nabla_\alpha \nabla_\beta \mathbf{R} = \mathbf{N}B_{\alpha\beta}. \tag{12.33}$$

Thus, the surface Laplacian of the position vector $\mathbf{R}$ yields the vector curvature normal of the surface:

$$\nabla_\alpha \nabla^\alpha \mathbf{R} = \mathbf{N}B^\alpha_\alpha. \tag{12.34}$$

Contract both sides of equation (12.32) with N_i to obtain an explicit expression for $B_{\alpha\beta}$:

$$B_{\alpha\beta} = N_i \nabla_\alpha Z^i_\beta. \tag{12.35}$$

Recall that $B_{\alpha\beta}$ depends on the choice of normal. The combination $N^i B_{\alpha\beta}$ is independent of that choice.

Exercise 257. Show that $B_{\alpha\beta}$ is symmetric

$$B_{\alpha\beta} = B_{\beta\alpha}. \tag{12.36}$$

The trace B^α_α of the curvature tensor B^α_β is the *mean curvature*. Another important invariant associated with the curvature tensor is the determinant $|B^{\cdot}_{\cdot}|$ of B^α_β. The trace B^α_α and the determinant $|B^{\cdot}_{\cdot}|$ are the two invariants associated with the tensor B^α_β in the linear algebra sense of similarity transformations in a two-dimensional space. These invariants can be expressed in terms of the eigenvalues κ_1 and κ_2 of B^α_β which are real since $B_{\alpha\beta}$ is symmetric. The quantities κ_1 and κ_2 are called the *principal curvatures* of the surface. Their geometric interpretation is discussed in Chap. 13 on curves. In terms of the principal curvatures, B^α_α and $|B^{\cdot}_{\cdot}|$ are given by

$$B^\alpha_\alpha = \kappa_1 + \kappa_2 \tag{12.37}$$

$$|B^{\cdot}_{\cdot}| = \kappa_1 \kappa_2. \tag{12.38}$$

Equation (12.37) is the reason behind the word *mean* in *mean curvature*. Some texts define the mean curvature as the average $(\kappa_1 + \kappa_2)/2$. This definition certain advantages over our definition in a few situations. For example the mean curvature of a sphere according to our definition is $-2/R$. In the alternative definition, the value is $-1/R$, which may be more consistent with our intuition of curvature. However, we prefer our definition since in numerous applications, and particularly in the calculus of moving surfaces, it is the quantity B^α_α (rather than $B^\alpha_\alpha/2$) that is most frequently encountered.

The metric tensor $S_{\alpha\beta}$ is sometimes referred to as the *first groundform* of the surface and the curvature tensor $B_{\alpha\beta}$ is the *second groundform*. There is also the *third groundform* C^α_β defined as the "matrix square" of the curvature tensor

$$C^\alpha_\beta = B^\alpha_\gamma B^\gamma_\beta. \tag{12.39}$$

The third groundform is encountered frequently in applications.

The extrinsic nature of the curvature tensor can be illustrated by forming a cylinder out of a sheet of paper. When the sheet of paper is flat, the curvature tensor $B_{\alpha\beta}$ vanishes since $\nabla_\alpha \mathbf{S}_\beta$ does not have a normal component. When the sheet is curved into a cylinder without stretching and each material particle is allowed to keep its coordinates, *the metric tensor and its derivatives are unchanged*. This is because the metric tensor can be calculated from the unchanged lengths of curves embedded in the surface. Meanwhile, the curvature tensor *does* change: the mean curvature of the cylinder with respect to the exterior normal is $-1/R$. Therefore, the curvature tensor depends on the way in which the surface is embedded in the ambient Euclidean space.

Exercise 258. Show that the following relationship holds at points of zero mean curvature:

$$B_\beta^\alpha B_\alpha^\beta = -2\left|B_\cdot^\cdot\right|. \tag{12.40}$$

12.5 The Calculation of the Curvature Tensor for a Sphere

As an illustration, we calculate the curvature tensor for a sphere of radius R. In Chap. 10, we calculated the shift tensor, the metric tensors and the Christoffel symbols for a sphere referred to Cartesian coordinates in the ambient space and the standard spherical coordinates on the surface. The identity

$$B_{\alpha\beta} = -Z_{i\alpha}\nabla_\beta N^i \tag{12.41}$$

is a form Weingarten's formula (11.50) and offers a convenient way for evaluating the curvature tensor. The advantage of equation (12.41) is that it does not require the use of the surface Christoffel symbol and, additionally, does not produce any "three-dimensional" objects such as $\nabla_\alpha Z_\beta^i$ which cannot be represented by matrices. Recall that $Z_{i\alpha}$ is given by the matrix

$$M_1 = \begin{bmatrix} R\cos\theta\cos\phi & -R\sin\theta\sin\phi \\ R\cos\theta\sin\phi & R\sin\theta\cos\phi \\ -R\sin\theta & 0 \end{bmatrix}. \tag{12.42}$$

The object $\nabla_\beta N^i = \partial N^i/\partial S^\beta$ is given by the matrix

$$M_2 = \begin{bmatrix} \cos\theta\cos\phi & -\sin\theta\sin\phi \\ \cos\theta\sin\phi & \sin\theta\cos\phi \\ -\sin\theta & 0 \end{bmatrix}. \tag{12.43}$$

Therefore, the curvature tensor $B_{\alpha\beta}$ is represented by the matrix product $-M_1^T M_2$

$$B_{\alpha\beta} = \begin{bmatrix} -R & 0 \\ 0 & -R\sin^2\theta \end{bmatrix}. \tag{12.44}$$

The curvature tensor B_β^α is obtained by contraction with the contravariant metric tensor $S^{\alpha\beta}$ in equation (10.97)

$$B_\beta^\alpha = \begin{bmatrix} -\frac{1}{R} & 0 \\ 0 & -\frac{1}{R} \end{bmatrix}. \tag{12.45}$$

Thus, the mean curvature B_α^α is

$$B_\alpha^\alpha = -\frac{2}{R}. \tag{12.46}$$

Finally, the contravariant curvature tensor $B^{\alpha\beta}$ is given by

$$B^{\alpha\beta} = \begin{bmatrix} -\frac{1}{R^3} & 0 \\ 0 & -\frac{1}{R^3 \sin^2\theta} \end{bmatrix}. \tag{12.47}$$

12.6 The Curvature Tensor for Other Common Surfaces

The cylinder (10.101)–(10.103)

$$B_{\alpha\beta} = \begin{bmatrix} -R & 0 \\ 0 & 0 \end{bmatrix}; \quad B_\beta^\alpha = \begin{bmatrix} -\frac{1}{R} & 0 \\ 0 & 0 \end{bmatrix}; \quad B^{\alpha\beta} = \begin{bmatrix} -\frac{1}{R^3} & 0 \\ 0 & 0 \end{bmatrix} \tag{12.48}$$

$$B_\alpha^\alpha = -\frac{1}{R} \tag{12.49}$$

Cone with angle Θ:

$$B_{\alpha\beta} = \begin{bmatrix} r\cos\Theta\sin\Theta & 0 \\ 0 & 0 \end{bmatrix}; \quad B_\beta^\alpha = \begin{bmatrix} -\frac{\cot\Theta}{r} & 0 \\ 0 & 0 \end{bmatrix}; \quad B^{\alpha\beta} = \begin{bmatrix} -\frac{\cos\Theta}{r^3\sin^3\Theta} & 0 \\ 0 & 0 \end{bmatrix} \tag{12.50}$$

$$B_\alpha^\alpha = -\frac{\cot\Theta}{r} \tag{12.51}$$

The torus (10.107)–(10.109):

$$B_{\alpha\beta} = \begin{bmatrix} -(R + r\cos\phi)\cos\phi & 0 \\ 0 & -r \end{bmatrix} \tag{12.52}$$

$$B_\beta^\alpha = \begin{bmatrix} -\frac{\cos\phi}{R+r\cos\phi} & 0 \\ 0 & -r^{-1} \end{bmatrix} \tag{12.53}$$

$$B^{\alpha\beta} = \begin{bmatrix} -\frac{\cos\phi}{(R+r\cos\phi)^3} & 0 \\ 0 & -r^{-3} \end{bmatrix} \tag{12.54}$$

$$B_\alpha^\alpha = -\frac{R + 2r\cos\phi}{r\,(R + r\cos\phi)} \tag{12.55}$$

The surface of revolution (10.115)–(10.117):

$$B_{\alpha\beta} = \begin{bmatrix} -\frac{r(z)}{\sqrt{1+r'(z)^2}} & 0 \\ 0 & \frac{r''(z)}{\sqrt{1+r'(z)^2}} \end{bmatrix} \tag{12.56}$$

$$B^{\alpha}_{\beta} = \begin{bmatrix} -\frac{1}{r(z)\sqrt{1+r'(z)^2}} & 0 \\ 0 & \frac{r''(z)}{\left(1+r'(z)^2\right)^{3/2}} \end{bmatrix} \tag{12.57}$$

$$B^{\alpha\beta} = \begin{bmatrix} -\frac{1}{r^3(z)\sqrt{1+r'(z)^2}} & 0 \\ 0 & \frac{r''(z)}{\left(1+r'(z)^2\right)^{5/2}} \end{bmatrix} \tag{12.58}$$

$$B^{\alpha}_{\alpha} = \frac{r''(z)\,r(z) - r'(z)^2 - 1}{r(z)\sqrt{1 + r'(z)^2}} \tag{12.59}$$

Exercise 259. Show that the catenoid given by $r(z) = a\cosh(z - b)/a$ has zero mean curvature.

12.7 A Particle Moving Along a Trajectory Confined to a Surface

This section parallels Sect. 8.9, in which we analyzed the motion of a particle along a trajectory in a Euclidean space. The present problem is substantially richer since the acceleration of the particle depends on the curvature tensor of the surface. The content of this section is presented in the form of exercises.

Exercise 260. Suppose that a particle confined to the surface moves along the trajectory $\gamma(t)$ given by

$$S^{\alpha} = S^{\alpha}(t). \tag{12.60}$$

Show that its velocity $\mathbf{V}$ is given by

$$\mathbf{V} = V^{\alpha}\mathbf{S}_{\alpha} \tag{12.61}$$

where

$$V^{\alpha} = \frac{dS^{\alpha}(t)}{dt}. \tag{12.62}$$

Hint: $\mathbf{R}(t) = \mathbf{R}(S(t))$.

Exercise 261. Show that the acceleration $\mathbf{A}$ of the particle is given by

$$\mathbf{A} = \frac{\delta V^\alpha}{\delta t}\mathbf{S}_\alpha + \mathbf{N}B_{\alpha\beta}V^\alpha V^\beta, \tag{12.63}$$

where

$$\frac{\delta V^\alpha}{\delta t} = \frac{dV^\alpha}{dt} + \Gamma^\alpha_{\beta\gamma}V^\beta V^\gamma. \tag{12.64}$$

Conclude that $\delta V^\alpha / \delta t$ is a tensor with respect to coordinate transformations on the surface. Also note the term $\mathbf{N}B_{\alpha\beta}V^\alpha V^\beta$ known as the *centripetal acceleration*.

Exercise 262. Use this problem as an opportunity to develop a new $\delta/\delta t$-calculus motivated by the definition (12.64). For a general tensor T^α_β, define the $\delta/\delta t$-derivative along the trajectory as follows

$$\frac{\delta T^\alpha_\beta}{\delta t} = \frac{dT^\alpha_\beta}{dt} + V^\gamma \Gamma^\alpha_{\gamma\omega}T^\omega_\beta - V^\gamma \Gamma^\omega_{\gamma\beta}T^\alpha_\omega. \tag{12.65}$$

This derivative is also known as the *intrinsic derivative*.

Exercise 263. Show that the $\delta/\delta t$-derivative satisfies the tensor property.

Exercise 264. Show that the $\delta/\delta t$-derivative satisfies the sum and product rules.

Exercise 265. Show that the $\delta/\delta t$-derivative satisfies the chain rule

$$\frac{\delta T^\alpha_\beta}{\delta t} = \frac{\partial T^\alpha_\beta(t,S)}{\partial t} + V^\gamma \nabla_\gamma T^\alpha_\beta \tag{12.66}$$

for trajectory restrictions of time-dependent surface tensors T^α_β.

Exercise 266. Conclude that the $\delta/\delta t$-derivative is metrinilic with respect to all the surface metrics, except $\mathbf{S}_\alpha$ and $\mathbf{S}^\alpha$.

Exercise 267. Show that

$$\frac{\delta \mathbf{S}_\alpha}{\delta t} = \mathbf{N}V^\gamma B_{\gamma\alpha}. \tag{12.67}$$

Exercise 268. Show that $\delta/\delta t$ commutes with contraction.

Exercise 269. Use the newly developed calculus to rederive equation (12.63) and to derive the following equation for the jolt $\mathbf{J}$:

$$\mathbf{J} = \left(\frac{\delta^2 V^\alpha}{\delta t^2} - B^\alpha_\beta B_{\gamma\delta}V^\beta V^\gamma V^\delta\right)\mathbf{S}_\alpha + \left(3B_{\alpha\beta}\frac{\delta V^\alpha}{\delta t}V^\beta + \nabla_\alpha B_{\beta\gamma}V^\alpha V^\beta V^\gamma\right)\mathbf{N}. \tag{12.68}$$

12.8 The Gauss–Codazzi Equation

Apply the general commutator equation (12.18) to the shift tensor Z_γ^i

$$\nabla_\alpha \nabla_\beta Z_\gamma^i - \nabla_\beta \nabla_\alpha Z_\gamma^i = -R^\delta_{\cdot\gamma\alpha\beta} Z_\delta^i. \tag{12.69}$$

By equation (12.32), we find

$$\nabla_\alpha \left(N^i B_{\beta\gamma}\right) - \nabla_\beta \left(N^i B_{\alpha\gamma}\right) = -R^\delta_{\cdot\gamma\alpha\beta} Z_\delta^i. \tag{12.70}$$

Applications of the product rule and Weingarten's formula (11.50) yield

$$-Z_\delta^i B_\alpha^\delta B_{\beta\gamma} + Z_\delta^i B_\beta^\delta B_{\alpha\gamma} + N^i \nabla_\alpha B_{\beta\gamma} - N^i \nabla_\beta B_{\alpha\gamma} = -R^\delta_{\cdot\gamma\alpha\beta} Z_\delta^i. \tag{12.71}$$

This equation is called the *Gauss–Codazzi equation*. Its normal projection is known as the Codazzi equation and its tangential projection is the celebrated Theorema Egregium discussed in the following section.

Contract both sides of equation (12.71) with the normal N_i to obtain the Codazzi equation:

$$\nabla_\alpha B_{\beta\gamma} - \nabla_\beta B_{\alpha\gamma} = 0. \tag{12.72}$$

The Codazzi equation states that $\nabla_\alpha B_{\beta\gamma}$ is symmetric with respect to α and β

$$\nabla_\alpha B_{\beta\gamma} = \nabla_\beta B_{\alpha\gamma}. \tag{12.73}$$

Since the curvature tensor is symmetric, the Codazzi equation tells us that $\nabla_\alpha B_{\beta\gamma}$ is fully symmetric with respect to its indices.

12.9 Gauss's Theorema Egregium

We have arrived at one of the most beautiful relationships in differential geometry: Gauss's *Theorema Egregium*, translated from Latin as the *Remarkable Theorem*. It is contained as the tangential component in equation (12.71) and can be revealed by contracting both sides with Z_i^ω (and subsequently renaming $\omega \to \delta$):

$$B_\alpha^\delta B_{\beta\gamma} - B_\beta^\delta B_{\alpha\gamma} = R^\delta_{\cdot\gamma\alpha\beta} \tag{12.74}$$

This celebrated equation is also known as the *Gauss equation of the surface*. It usually appears in the form with lowered indices:

$$R_{\gamma\delta\alpha\beta} = B_{\alpha\gamma} B_{\beta\delta} - B_{\beta\gamma} B_{\alpha\delta}. \tag{12.75}$$

In Gauss's own words: "Thus, the formula of the preceding article leads itself to the remarkable Theorem. If a curved surface is developed upon any other surface whatever, the measure of curvature in each point remains unchanged."

Theorema Egregium makes the following powerful statement: the combination $B_{\alpha\gamma}B_{\beta\delta} - B_{\beta\gamma}B_{\alpha\delta}$, being the Riemann–Christoffel tensor, is *intrinsic*. That is, in can be expressed in terms of the covariant metric $S_{\alpha\beta}$ and its derivatives. In effect, it means that the object $B_{\alpha\gamma}B_{\beta\delta} - B_{\beta\gamma}B_{\alpha\delta}$ can be calculated by measuring distances on the surfaces.

We will now derive several elegant equivalents of Theorema Egregium. Recall equation (12.28), in which the Riemann–Christoffel tensor is expressed in terms of the Gaussian curvature K

$$R_{\gamma\delta\alpha\beta} = K\varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta}. \tag{12.28}$$

Equation (12.28) yields

$$B_{\alpha\gamma}B_{\beta\delta} - B_{\beta\gamma}B_{\alpha\delta} = K\varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta}, \tag{12.76}$$

as well as an explicit formula for the Gaussian curvature in terms of the curvature tensor

$$K = \frac{1}{4}\varepsilon^{\alpha\beta}\varepsilon^{\gamma\delta}\left(B_{\alpha\gamma}B_{\beta\delta} - B_{\beta\gamma}B_{\alpha\delta}\right). \tag{12.77}$$

Multiply the right-hand side of (12.77) and express the Levi-Civita symbols in terms of the permutation systems $e^{\alpha\beta}$ and $e^{\gamma\delta}$:

$$K = \frac{1}{4S}e^{\alpha\beta}e^{\gamma\delta}B_{\alpha\gamma}B_{\beta\delta} - \frac{1}{4S}e^{\alpha\beta}e^{\gamma\delta}B_{\beta\gamma}B_{\alpha\delta}. \tag{12.78}$$

One easily recognizes that each of the two terms on the right-hand side equals $(2S)^{-1}|B_{\cdot\cdot}|$, where $|B_{\cdot\cdot}|$ denotes the determinant of the curvature tensor $B_{\alpha\beta}$. Therefore,

$$K = S^{-1}|B_{\cdot\cdot}|. \tag{12.79}$$

Since S^{-1} is the determinant of the contravariant metric tensor $S^{\alpha\beta}$, we find

$$K = |S^{\cdot\cdot}||B_{\cdot\cdot}|. \tag{12.80}$$

By the multiplication property of the determinant (9.28), we arrive at the ultimate expression of Theorema Egregium:

$$K = |B^{\cdot}_{\cdot}|. \tag{12.81}$$

In words, Gaussian curvature equals the determinant of the curvature tensor B^{α}_{β}. In particular, Gaussian curvature equals the product of the principal curvatures

$$K = \kappa_1 \kappa_2. \tag{12.82}$$

Now, consider another way of manipulating equation (12.76)

$$B_{\alpha\gamma} B_{\beta\delta} - B_{\beta\gamma} B_{\alpha\delta} = K \varepsilon_{\alpha\beta} \varepsilon_{\gamma\delta}. \tag{12.76}$$

Raise the indices α and β and contract α with γ

$$B^{\alpha}_{\alpha} B^{\beta}_{\delta} - B^{\beta}_{\alpha} B^{\alpha}_{\delta} = K \varepsilon^{\alpha\beta} \varepsilon_{\alpha\delta}. \tag{12.83}$$

Since $\varepsilon^{\alpha\beta} \varepsilon_{\alpha\delta} = \delta^{\beta}_{\delta}$, we find

$$B^{\alpha}_{\alpha} B^{\beta}_{\delta} - B^{\beta}_{\alpha} B^{\alpha}_{\delta} = K \delta^{\beta}_{\delta}. \tag{12.84}$$

This new form of Theorema Egregium relates the first, second, and third groundforms of the surface.

$$C_{\alpha\beta} = H B_{\alpha\beta} - K S_{\alpha\beta}. \tag{12.85}$$

Exercise 270. Show that the Gaussian curvature of the cylinder (10.101)–(10.103) is zero:

$$K = 0. \tag{12.86}$$

Exercise 271. Show that the Gaussian curvature of a cone is zero:

$$K = 0. \tag{12.87}$$

Exercise 272. Show that the Gaussian curvature of the sphere (10.92)–(10.94) is

$$K = \frac{1}{R^2}. \tag{12.88}$$

Exercise 273. Show that the Gaussian curvature of the torus (10.107)–(10.109) is

$$K = \frac{\cos\phi}{r\left(R + r\cos\phi\right)}. \tag{12.89}$$

Exercise 274. Show that the Gaussian curvature of the surface of revolution (10.115)–(10.117) is

$$K = -\frac{r''(z)}{r(z)\left(1 + r'(z)^2\right)^2}. \tag{12.90}$$

12.10 The Gauss–Bonnet Theorem

The Gauss–Bonnet theorem is a fundamental result in differential geometry and topology. It states that for a closed surface, the *total curvature*, defined as the integral of the Gaussian curvature K, depends on the genus of the surface and not its shape. The *genus* g *of a surface* is the number of topological holes. For example, the genus of a sphere is zero and the genus of a torus is one. According to the Gauss–Bonnet theorem, the total curvature is $4\pi\,(1-g)$:

$$\int_S K dS = 4\pi\,(1-g)\,. \tag{12.91}$$

In particular, the total curvature is 4π for any surface of genus zero and 0 for any surface of genus one. In Chap. 17, we discuss the Gauss–Bonnet theorem in greater detail and give a novel proof based on the calculus of moving surfaces.

Exercise 275. Verify the Gauss–Bonnet theorem for a sphere. That is

$$\int_S K dS = 4\pi. \tag{12.92}$$

Exercise 276. Verify the Gauss–Bonnet theorem for a torus. That is

$$\int_S K dS = 0. \tag{12.93}$$

12.11 Summary

This chapter, devoted to curvature, was one of the most exciting to write. Curvature is not only an intriguing geometric concept, but also the source of most of the analytical challenges that make tensor calculus an indispensable tool of differential geometry. Armed with the tensor calculus framework, we had hardly any difficulties with any of the analytical calculations. This is truly a tribute to the power of the tensor technique. Interestingly, all of the results discussed in this chapter were known to Gauss and Riemann. And these great mathematicians did not have contemporary techniques at their disposal, which makes their achievements all the more remarkable.

Chapter 13
Embedded Curves

13.1 Preview

In this chapter, we apply the methods of tensor calculus to embedded curves. In some ways curves are similar to surfaces and in some ways they are different. Of course, our focus is on the differences. When embedded in Euclidean spaces of dimension greater than two, curves are not hypersurfaces and therefore do not have a well-defined normal $\mathbf{N}$ and curvature tensor $B_{\alpha\beta}$. Furthermore, a number of interesting features of curves can be attributed to their one-dimensional nature. For example, curves are intrinsically Euclidean: they admit Cartesian coordinates (arc length) and their Riemann–Christoffel tensor vanishes. Other properties that stem from the curves' one-dimensional nature are captured by the Frenet formulas.

We also consider curves embedded in two-dimensional surfaces which are, in turn, embedded in the ambient Euclidean space. Curves embedded in two-dimensional surfaces are hypersurfaces with respect to that embedding. Therefore, much of our discussion of general hypersurfaces carries over to curves embedded in surfaces. However, the ambient surfaces are not Euclidean! This gives us an opportunity to discuss elements of differential geometry from a Riemannian point of view.

13.2 The Intrinsic Geometry of a Curve

Curves are intrinsically Euclidean. In analytical terms, the Riemann–Christoffel tensor vanishes for all curves. While we demonstrate this below, it is also clear from geometric intuition. A curve can be parameterized by its arc length, which is equivalent to introducing a Cartesian coordinate system. To convince yourself of this, imagine the curve as an inextensible string with etched–in arc length coordinates. Suppose now that the string is pulled into a straight line. Since the string is inextensible, no part of it is stretched or compressed. Thus the transformation was

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_13,

isometric—in other words, all distances along the curve are preserved. Now that the string is straight, it is clearly a Euclidean space referred to Cartesian coordinates. However, to an intrinsic observer only capable of measuring distances along the curve, nothing has changed. Therefore, the original curve is also Euclidean.

13.3 Different Parametrizations of a Curve

Curves are parametrized by a single variable often taken to be the (signed) arc length s. Most, if not all, classical texts approach curves in this way. The arc length approach has a number of advantages, not the least of which is the clear geometric interpretation. For example, the quantity

$$\mathbf{T} = \frac{d\mathbf{R}(s)}{ds}, \tag{13.1}$$

is the unit tangent. Could there be a more natural object associated with a curve? However, the arc length approach goes very much against the spirit of our subject. We have passionately advocated against choosing special coordinate systems in deriving general results. The reasons against doing so were outlined for general Euclidean spaces in Chap. 3. Those reasons remain valid for curves. More practically, what does one do in situations where arc length is an inconvenient parameter? In fact, for most curves, it may be quite difficult to come up with the arc length parametrization. For example, consider the ellipsoidal spiral

$$x(t) = A\cos t \tag{13.2}$$

$$y(t) = B\sin t \tag{13.3}$$

$$z(t) = Ct. \tag{13.4}$$

To come up with an arc length parameterization, one would need to calculate the integral

$$s(t) = \int_0^t \sqrt{A^2\sin^2 t' + B^2\cos^2 t' + C^2}dt', \tag{13.5}$$

and then invert the resulting function of t to obtain $t(s)$. Clearly, this is not a desirable approach in most practical applications.

We therefore refer the curve to an arbitrary coordinate system. The coordinate variable is denoted by U^1. Despite the fact that there is a single dimension, we add an index to the independent variable in order to preserve the tensor notation. Furthermore, we use capital Greek letters to represent the value 1. Thus, U^1 may be referred to as U^Φ or U^Ψ. This allows us to draw a close parallel between curves and surfaces and to reap full benefits of the tensor calculus framework.

13.4 The Fundamental Elements of Curves

Refer the ambient space to coordinates Z^i and that the embedded curve to the coordinate U^Φ:

$$Z^i = Z^i(U). \tag{13.6}$$

The shift tensor Z^i_Φ is obtained by partial differentiation

$$Z^i_\Phi = \frac{\partial Z^i}{\partial U^\Phi}. \tag{13.7}$$

As with embedded surfaces, a variant is considered a tensor if it changes according to the tensor rule with respect to changes of both coordinate systems.

The covariant basis $\mathbf{U}_\Phi$, consisting of a single vector, is defined by

$$\mathbf{U}_\Phi = \frac{\partial \mathbf{R}(U)}{\partial U^\Phi} \tag{13.8}$$

and is related to the ambient basis $\mathbf{Z}_i$ by contraction with the shift tensor

$$\mathbf{U}_\Phi = \mathbf{Z}_i Z^i_\Phi. \tag{13.9}$$

The covariant metric tensor $U_{\Phi\Psi}$, a single number which could be thought of as a 1×1 matrix, is defined by the dot product

$$U_{\Phi\Psi} = \mathbf{U}_\Phi \cdot \mathbf{U}_\Psi. \tag{13.10}$$

It is related to the ambient metric tensor Z_{ij} by a double contraction with the shift tensor:

$$U_{\Phi\Psi} = Z_{ij} Z^i_\Phi Z^j_\Psi. \tag{13.11}$$

The contravariant metric tensor $U^{\Phi\Psi}$ is the "matrix" inverse of $U_{\Phi\Psi}$:

$$U_{\Theta\Phi} U^{\Phi\Psi} = \delta^\Psi_\Theta. \tag{13.12}$$

Finally, the contravariant basis $\mathbf{U}^\Phi$ is given by

$$\mathbf{U}^\Phi = U^{\Phi\Psi} \mathbf{U}_\Psi. \tag{13.13}$$

The object U, defined as the determinant of the covariant metric tensor

$$U = |U_{\cdot\cdot}|$$

is a relative invariant of weight 2. It is given explicitly by the formula

$$U = Z_{ij}\frac{dZ^i}{dU^1}\frac{dZ^j}{dU^1}. \tag{13.14}$$

The *length element* or the *line element* $\sqrt{U}$ is the square root of U, and it is a relative invariant of weight 1. The length of curve s corresponding to the coordinate ranging from a and b is given by the integral

$$s = \int_a^b \sqrt{U}dU. \tag{13.15}$$

This equation was derived in Sect. 5.8.

The Levi-Civita symbols ε_Φ and ε^Φ are defined by

$$\varepsilon_\Phi = \sqrt{U} \tag{13.16}$$

$$\varepsilon^\Phi = \frac{1}{\sqrt{U}}. \tag{13.17}$$

The line element $\sqrt{U}$ is a relative tensor of weight 1 when treated as a variant of order zero. But it is an absolute tensor when treated as a variant of order one.

In Chap. 9, we discussed the Levi-Civita in two or more dimensions. Equations (13.16) and (13.17) are a generalization to one dimension. For curves, the Levi-Civita symbol loses its skew-symmetric characteristics. Nevertheless, it remains a very useful object. For example, the combination $\varepsilon^\Phi \mathbf{U}_\Phi$ is an invariant. It must therefore be the unit tangent.

Exercise 277. Explain why $\varepsilon^\Phi \mathbf{U}_\Phi$ is the unit tangent. Hint: Consider $U^1 = s$.

The Christoffel symbol $\Gamma^\Theta_{\Phi\Psi}$ is defined by

$$\Gamma^\Theta_{\Phi\Psi} = \mathbf{U}^\Theta \cdot \frac{\partial \mathbf{U}_\Phi}{\partial U^\Psi}. \tag{13.18}$$

The Christoffel symbol is expressed alternatively by

$$\Gamma^\Theta_{\Phi\Psi} = -\frac{\partial \mathbf{U}^\Theta}{\partial U^\Psi} \cdot \mathbf{U}_\Phi. \tag{13.19}$$

In terms of the metric tensor, the Christoffel symbol is given intrinsically by the usual identity

$$\Gamma^\Theta_{\Phi\Psi} = \frac{1}{2}U^{\Theta\Omega}\left(\frac{\partial U_{\Omega\Phi}}{\partial U^\Psi} + \frac{\partial U_{\Omega\Psi}}{\partial U^\Phi} - \frac{\partial U_{\Phi\Psi}}{\partial U^\Omega}\right). \tag{13.20}$$

When a curve is parameterized by the arc length s, the values of the fundamental objects are consistent with the Cartesian nature of such parameterization. Namely, with respect to the arc length s,

$$U_{\Phi\Psi} = [1]; \quad U^{\Phi\Psi} = [1]; \quad \Gamma^{\Theta}_{\Phi\Psi} = 0. \tag{13.21}$$

13.5 The Covariant Derivative

The covariant derivative ∇_Θ applied to a tensor $W^{i\Phi}_{j\Psi}$ with a representative collection of indices is entirely analogous to the surface covariant derivative:

$$\nabla_\Theta W^{i\Phi}_{j\Psi} = \frac{\partial W^{i\Phi}_{j\Psi}}{\partial U^\Theta} + Z^k_\Theta \Gamma^i_{km} W^{m\Phi}_{j\Psi} - Z^k_\Theta \Gamma^m_{kj} W^{i\Phi}_{m\Psi} + \Gamma^\Phi_{\Theta\Omega} W^{i\Omega}_{j\Psi} - \Gamma^\Omega_{\Theta\Psi} W^{i\Phi}_{j\Omega}. \tag{13.22}$$

This covariant derivative ∇_Θ has all of the same properties as the surface covariant derivative described in Chap. 11. In particular, it produces tensor outputs for tensor inputs, satisfies the sum and the product rules, commutes with contraction and obeys the chain rule for curve restrictions W^i_j of ambient variants

$$\nabla_\Theta W^i_j = Z^k_\Theta \nabla_k W^i_j. \tag{13.23}$$

There is nothing that can be said about the covariant derivative ∇_Θ that was not said in Chap. 11 on the covariant surface derivative ∇_α. The definition (13.22) does not rely on the one-dimensional nature of curves and is therefore valid for subspaces of arbitrary dimension. We highlight an important difference: since a curve is not a hypersurface with respect to the ambient space and there is no a priori normal direction, the quantity $\nabla_\Phi \mathbf{U}_\Psi$ cannot be used to define a curvature tensor $B_{\Phi\Psi}$. Thus, our approach to analyzing curvature must be modified. This is accomplished in the next section.

We can take advantage of the one-dimensional nature of curves and define a new differential operator that produces tensors of the same order as the argument. The new operator ∇_s is defined by the combination

$$\nabla_s = \varepsilon^\Theta \nabla_\Theta, \tag{13.24}$$

that is

$$\nabla_s W^{i\Phi}_{j\Psi} = \varepsilon^\Theta \nabla_\Theta W^{i\Phi}_{j\Psi}. \tag{13.25}$$

The following exercises describe the analytical properties of this operator.

Exercise 278. Show that ∇_s coincides with d/ds for invariants

$$\nabla_s W = \frac{dW}{ds}. \tag{13.26}$$

Exercise 279. Show that ∇_s satisfies the sum and product rules.

Exercise 280. Show that ∇_s commutes with contraction.

Exercise 281. Show that ∇_s satisfies the following chain rule for curve restrictions of ambient variants

$$\nabla_s W^i_j = \varepsilon^{\Theta} Z^k_{\Theta} \nabla_k W^i_j. \tag{13.27}$$

Therefore,

$$\nabla_s W^i_j = T^k \nabla_k W^i_j, \tag{13.28}$$

where T^k is the contravariant component of the unit tangent.

Exercise 282. Show that

$$\nabla_s^2 = \nabla_{\Phi} \nabla^{\Phi}. \tag{13.29}$$

13.6 The Curvature and the Principal Normal

As was the case for hypersurfaces, the object $\nabla_{\Phi} \mathbf{U}_{\Psi}$ does not vanish. Instead, it is orthogonal to the tangent space:

$$\mathbf{U}^{\Theta} \cdot \nabla_{\Phi} \mathbf{U}_{\Psi} = 0. \tag{13.30}$$

Derivation of this equation is analogous to that of equation (11.15). In the case of hypersurfaces, we argued that since $\nabla_{\Phi} \mathbf{U}_{\Psi}$ is orthogonal to the tangent plane, it must point along the normal N. For curves, there is no unique normal direction. For curves embedded in the three-dimensional Euclidean space, the normal space at each point on a curve is a plane. In general, the dimension of the normal space equals the co-dimension of the curve. Therefore, we are not able to represent $\nabla_{\Phi} \mathbf{U}_{\Psi}$ as a product of an invariant vector and a tensor with scalar elements. We refer to this object as the *vector curvature tensor* and denote it by $\mathbf{B}_{\Phi\Psi}$:

$$\mathbf{B}_{\Phi\Psi} = \nabla_{\Phi} \mathbf{U}_{\Psi}. \tag{13.31}$$

For curves, the vector curvature tensor consists of a single vector. The indices continue to play their important role: they remind us that $\mathbf{B}_{\Phi\Psi}$ is not an invariant vector and indicate how $\mathbf{B}_{\Phi\Psi}$ changes under a reparametrization of the curve. In order to obtain an invariant, one of the indices must be raised and contracted with the other to yield $\mathbf{B}_{\Phi}^{\Phi}$. The invariant $\mathbf{B}_{\Phi}^{\Phi}$ is called the *curvature normal*.

The curvature normal $\mathbf{B}_{\Phi}^{\Phi}$ can be viewed from two complementary points of view. If the curve is interpreted as a general embedded surface, the curvature normal is seen as the surface Laplacian of the position vector $\mathbf{R}$:

$$\mathbf{B}_{\Phi}^{\Phi} = \nabla_{\Phi}\nabla^{\Phi}\mathbf{R}. \tag{13.32}$$

This interpretation is an immediate consequence of the definition in equation (13.31). This is the interpretation of Chap. 12. Alternatively, if the emphasis is on the one-dimensional nature of the curve, $\mathbf{B}_{\Phi}^{\Phi}$ can be represented as

$$\mathbf{B}_{\Phi}^{\Phi} = \nabla_s^2\mathbf{R} = \frac{d^2\mathbf{R}(s)}{ds^2}, \tag{13.33}$$

where the invariant derivative ∇_s is defined in equation (13.24) and the equivalence of (13.32) and (13.33) follows from equation (13.29). The latter interpretation is particularly important in the following section on Frenet formulas.

At points where the vector $\mathbf{B}_{\Phi}^{\Phi}$ does not vanish it can be represented as a product its length κ, called the *curvature*, and a unit vector $\mathbf{P}$ called the *principal normal* $\mathbf{P}$:

$$\mathbf{B}_{\Phi}^{\Phi} = \kappa\mathbf{P}. \tag{13.34}$$

This concept of curvature is analogous to the mean curvature of a hyper surface. However, the curvature κ is always nonnegative and the principal normal $\mathbf{P}$ is only defined at points with positive curvature κ. By contrast, the mean curvature B_{α}^{α} of a hypersurface can be positive or negative and the normal N is defined everywhere. These differences are illustrated in Fig. 13.1, which shows a curve that can be considered either as a hypersurface in the plane

$$x(t) = t \tag{13.35}$$

$$y(t) = \sin t \tag{13.36}$$

and as a curve in the three-dimensional space

$$x(t) = t \tag{13.37}$$

$$y(t) = \sin t \tag{13.38}$$

$$z(t) = 0. \tag{13.39}$$

In the former case, the normal N is a continuous vector field and the mean curvature is given by

$$B_{\alpha}^{\alpha} = \frac{\sin\alpha}{\left(1+\cos^2\alpha\right)^{3/2}}. \tag{13.40}$$

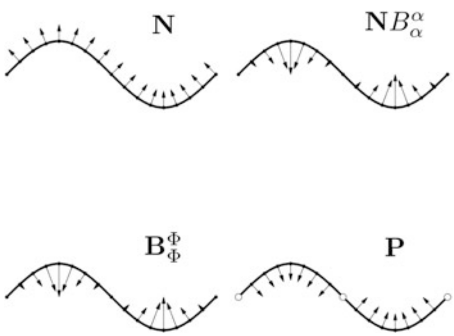


Fig. 13.1 The first two plots show a sinusoid viewed as embedded in the two-dimensional plane. With respect to this embedding, the sinusoid is a hypersurface and has a well-defined continuous unit normal field **N**. The other two plots show the sinusoid embedded in the three-dimensional space. The curvature normal $\mathbf{B}_\alpha^\alpha$ is a smooth vector field, while the principal normal **P** is only defined at points of nonzero curvature

In the latter case, the principal normal **P** is piecewise continuous and is undefined at points of zero curvature. Also, when the curve is interpreted as a hypersurface, the direction of the normal can be arbitrarily chosen, while the direction of the principal normal **P** is unique.

13.7 The Binormal and the Frenet Formulas

The focus of the preceding sections was on the elements that curves share with surfaces of higher dimensions. The focus of this section is on the one-dimensional nature of curves. We start with the position vector **R** and repeatedly apply the invariant derivative ∇_s. Since, according to equation (13.26), ∇_s is equivalent to d/ds when applied to invariants, we may as well refer the curve to the signed arc length s and use the ordinary derivative d/ds. Keep in mind, however, that, for invariants, d/ds is equivalent to $\varepsilon^\Phi \nabla_\Phi$, which means that the entire discussion can be easily translated to arbitrary coordinates.

The unit tangent vector **T** is obtained by applying d/ds to the position vector **T**:

$$\frac{d\mathbf{R}}{ds} = \mathbf{T}. \tag{13.41}$$

An application of d/ds to **T** yields the curvature and the principal normal:

$$\frac{d\mathbf{T}}{ds} = \kappa \mathbf{P}, \tag{13.42}$$

where the principal normal $\mathbf{P}$ is unit length and orthogonal to the tangent $\mathbf{T}$:

$$\mathbf{P}\cdot\mathbf{P}=1 \tag{13.43}$$

$$\mathbf{P}\cdot\mathbf{T}=0. \tag{13.44}$$

Exercise 283. Differentiate the identity $\mathbf{T}(s)\cdot\mathbf{T}(s)=1$ to conclude that $\mathbf{P}$ is orthogonal to $\mathbf{T}$.

Applying d/ds to equation (13.43):

$$\frac{d\mathbf{P}}{ds}\cdot\mathbf{P}=0, \tag{13.45}$$

and to equation (13.44):

$$\frac{d\mathbf{P}}{ds}\cdot\mathbf{T}+\kappa=0. \tag{13.46}$$

The last equation can be written as

$$\left(\frac{d\mathbf{P}}{ds}+\kappa\mathbf{T}\right)\cdot\mathbf{T}=0 \tag{13.47}$$

which shows that the vector

$$\frac{d\mathbf{P}}{ds}+\kappa\mathbf{T} \tag{13.48}$$

is orthogonal to $\mathbf{T}$. Furthermore, from equation (13.45) it follows that this vector is also orthogonal to $\mathbf{P}$. Denote this vector by $\tau\mathbf{Q}$:

$$\frac{d\mathbf{P}}{ds}+\kappa\mathbf{T}=\tau\mathbf{Q}, \tag{13.49}$$

where $\mathbf{Q}$ is a unit vector and the sign of τ is chosen such that the triplet $\mathbf{T},\mathbf{P},\mathbf{Q}$ forms a right-handed set in the following sense

$$\varepsilon_{ijk}T^iP^jQ^k=1. \tag{13.50}$$

The quantity τ is called the *torsion* and the vector $\mathbf{Q}$ is called the *binormal vector* or simply the *binormal*.

The binormal $\mathbf{Q}$ can be expressed as the vector product of the tangent and the principal normal

$$\mathbf{Q}=\mathbf{T}\times\mathbf{P}. \tag{13.51}$$

Applying d/ds to both sides of the equation, we find

$$\frac{d\mathbf{Q}}{ds} = \frac{d\mathbf{T}}{ds} \times \mathbf{P} + \mathbf{T} \times \frac{d\mathbf{P}}{ds}, \tag{13.52}$$

or

$$\frac{d\mathbf{Q}}{ds} = \mathbf{P} \times \mathbf{P} + \mathbf{T} \times (-\kappa\mathbf{T} + \tau\mathbf{Q}). \tag{13.53}$$

Since $\mathbf{T} \times \mathbf{Q} = -\mathbf{P}$, we have

$$\frac{d\mathbf{Q}}{ds} = -\tau\mathbf{P}. \tag{13.54}$$

The combination of equations (13.42), (13.49), and (13.54), summarized here:

$$\frac{d\mathbf{T}}{ds} = \kappa\mathbf{P} \tag{13.55}$$

$$\frac{d\mathbf{P}}{ds} = -\kappa\mathbf{T} + \tau\mathbf{Q} \tag{13.56}$$

$$\frac{d\mathbf{Q}}{ds} = -\tau\mathbf{P} \tag{13.57}$$

are called the *Frenet formulas*.

Exercise 284. Show that the vector $d\mathbf{P}/ds + \kappa\mathbf{T}$ is obtained from $d\mathbf{P}/ds$ by the Gram–Schmidt orthogonalization procedure.

Exercise 285. Calculate the principal normal, binormal, curvature, torsion for the spiral

$$x(t) = 4\cos t \tag{13.58}$$

$$y(t) = 4\sin t \tag{13.59}$$

$$z(t) = 3t. \tag{13.60}$$

Exercise 286. Repeat the calculation for the spiral

$$x(t) = 4\cos t \tag{13.61}$$

$$y(t) = 4\sin t \tag{13.62}$$

$$z(t) = -3t. \tag{13.63}$$

Note which answers change for this spiral.

13.8 The Frenet Formulas in Higher Dimensions

When the Frenet formulas (13.55)–(13.57) are written in matrix form

$$\begin{bmatrix} \frac{d\mathbf{T}}{ds} \\ \frac{d\mathbf{P}}{ds} \\ \frac{d\mathbf{Q}}{ds} \end{bmatrix} = \begin{bmatrix} & \kappa & \\ -\kappa & & \tau \\ & -\tau & \end{bmatrix} \begin{bmatrix} \mathbf{T} \\ \mathbf{P} \\ \mathbf{Q} \end{bmatrix}, \tag{13.64}$$

and we are struck by the skew symmetric property. This is part of a larger pattern.

Suppose that a curve is embedded in an N-dimensional Euclidean space. At a given point on the curve, construct an orthonormal basis according to the following procedure. The initial vector $\mathbf{T}_0$ is taken to be the unit tangent $\mathbf{T}$. Subsequently, $\mathbf{T}_m$ is obtained is obtained by applying a single step of Gram–Schmidt algorithm to $d\mathbf{T}_{m-1}/ds$ to make it orthogonal to each of the preceding vectors and factoring out a scalar multiple to make it unit length.

In the preceding section, we examined the first two steps of the procedure. The vector $d\mathbf{T}_0/ds$ is automatically orthogonal to $\mathbf{T}_0$, since $\mathbf{T}_0\,(s)$ is unit length for all s. Thus, in the context of the Gram–Schmidt procedure, it only needs to be normalized to produce $\mathbf{T}_1$:

$$\frac{d\mathbf{T}_0}{ds} = \kappa_1 \mathbf{T}_1. \tag{13.65}$$

Using the terminology of the preceding section, $\mathbf{T}_1$ is the principal normal $\mathbf{P}$ and κ_1 is the curvature κ. Thus, $\mathbf{T}_1$ satisfies the following identities

$$\mathbf{T}_1 \cdot \mathbf{T}_0 = 0 \tag{13.66}$$

$$\mathbf{T}_1 \cdot \mathbf{T}_1 = 1. \tag{13.67}$$

Differentiating the first identity we find

$$\frac{d\mathbf{T}_1}{ds} \cdot \mathbf{T}_0 + \mathbf{T}_1 \cdot \frac{d\mathbf{T}_0}{ds} = 0, \tag{13.68}$$

which, by equation (13.65), gives

$$\frac{d\mathbf{T}_1}{ds} \cdot \mathbf{T}_0 = -\kappa_1. \tag{13.69}$$

Differentiating the second identity yields the familiar orthogonality condition

$$\frac{d\mathbf{T}_1}{ds} \cdot \mathbf{T}_1 = 0. \tag{13.70}$$

Thus, $d\mathbf{T}_1/ds$ is orthogonal to $\mathbf{T}_1$. Therefore, the Gram–Schmidt formula includes the term that corresponds to $\mathbf{T}_0$ and a coefficient κ_2 chosen so that $\mathbf{T}_2$ is unit length:

$$\kappa_2\mathbf{T}_2 = \frac{d\mathbf{T}_1}{ds} - \left(\frac{d\mathbf{T}_1}{ds}\cdot\mathbf{T}_0\right)\mathbf{T}_0. \tag{13.71}$$

Therefore, by equation (13.69), we have

$$\kappa_2\mathbf{T}_2 = \frac{d\mathbf{T}_1}{ds} + \kappa_1\mathbf{T}_0, \tag{13.72}$$

or

$$\frac{d\mathbf{T}_1}{ds} = \kappa_2\mathbf{T}_2 - \kappa_1\mathbf{T}_0. \tag{13.73}$$

We recognize this relationship as equation (13.49), where $\mathbf{T}_2$ is the binormal $\mathbf{Q}$ and κ_2 is the torsion τ.

We now embark on the inductive step of the procedure. The unit vector $\mathbf{T}_m$ is orthogonal to all preceding vectors $\mathbf{T}_{k<m}$:

$$\mathbf{T}_m\cdot\mathbf{T}_0 = 0 \tag{13.74}$$
$$\mathbf{T}_m\cdot\mathbf{T}_1 = 0 \tag{13.75}$$
$$\cdots = \cdots$$
$$\mathbf{T}_m\cdot\mathbf{T}_k = 0 \tag{13.76}$$
$$\cdots = \cdots$$
$$\mathbf{T}_m\cdot\mathbf{T}_m = 1. \tag{13.77}$$

Furthermore, presume that up to $k = m - 1$, we have

$$\frac{d\mathbf{T}_k}{ds} = \kappa_{k+1}\mathbf{T}_{k+1} - \kappa_k\mathbf{T}_{k-1}. \tag{13.78}$$

which we demonstrated for $k = 0$ and 1.

Differentiate equation (13.76), valid for $k < m$, with respect to s:

$$\frac{d\mathbf{T}_m}{ds}\cdot\mathbf{T}_k + \mathbf{T}_m\cdot\frac{d\mathbf{T}_k}{ds} = 0. \tag{13.79}$$

Substitute equation (13.78) into equation (13.79):

$$\frac{d\mathbf{T}_m}{ds}\cdot\mathbf{T}_k + \kappa_{k+1}\mathbf{T}_m\cdot\mathbf{T}_{k+1} - \kappa_k\mathbf{T}_m\cdot\mathbf{T}_{k-1} = 0. \tag{13.80}$$

The dot product $\mathbf{T}_m \cdot \mathbf{T}_{k-1}$ vanishes for all $k < m$. Meanwhile, the dot product $\mathbf{T}_m \cdot \mathbf{T}_{k+1}$ vanishes for all $k < m - 1$ and equals 1 for $k = m - 1$. Therefore, equation (13.80) tells us the following

$$\frac{d\mathbf{T}_m}{ds} \cdot \mathbf{T}_k = 0, \text{ for } k < m - 1 \tag{13.81}$$

$$\frac{d\mathbf{T}_m}{ds} \cdot \mathbf{T}_{m-1} = -\kappa_m \tag{13.82}$$

In other words, $d\mathbf{T}_m/ds$ is orthogonal to $\mathbf{T}_0, \ldots, \mathbf{T}_{m-2}$, but not $\mathbf{T}_{m-1}$. The fact that $d\mathbf{T}_m/ds$ is orthogonal to $\mathbf{T}_m$ follows from equation (13.77). Therefore, the Gram–Schmidt formula always contains a single term

$$\kappa_{m+1}\mathbf{T}_{m+1} = \frac{d\mathbf{T}_m}{ds} - \left(\frac{d\mathbf{T}_m}{ds} \cdot \mathbf{T}_{m-1}\right)\mathbf{T}_{m-1}, \tag{13.83}$$

which, with the help of equation (13.82) reads

$$\kappa_{m+1}\mathbf{T}_{m+1} = \frac{d\mathbf{T}_m}{ds} + \kappa_m \mathbf{T}_{m-1}. \tag{13.84}$$

This equation can be written as

$$\frac{d\mathbf{T}_m}{ds} = \kappa_{m+1}\mathbf{T}_{m+1} - \kappa_m \mathbf{T}_{m-1}. \tag{13.85}$$

and the proof by induction is complete.

Exercise 287. Explain why the last equation in the Frenet procedure reads

$$\frac{d\mathbf{T}_{N-1}}{ds} = -\kappa_{N-1}\mathbf{T}_{N-2}, \tag{13.86}$$

consistent with equation (13.57).

13.9 Curves Embedded in Surfaces

Leonhard Euler's original investigation of surface curvature was based on analyzing curves embedded in surfaces. Euler's analysis focused on intersections of surfaces with planes since curves embedded in planes was well understood.

We have so far analyzed two different types of embedded manifolds: hypersurfaces and curves embedded in a Euclidean space. We are about to analyze a new combination—curves embedded in surfaces. This configuration has a fundamentally new characteristic: the ambient space is not Euclidean. This is a typical situation in Riemannian geometry, where the ambient space provides a metric from which the differential elements of the embedded manifold are derived. Therefore, the discussion contained in this section is essentially an example of Riemannian analysis.

Suppose the surface is given by the equation

$$Z^i = Z^i(S), \tag{10.3}$$

and the embedded surface is given by

$$S^\alpha = S^\alpha(U), \tag{13.87}$$

as before. We first discuss the embedding of the curve in the surface. This is the Riemannian portion of our analysis. While many of the equations are identical to previously encountered relationships, their interpretation is different.

The shift tensor S^α_Φ is given by

$$S^\alpha_\Phi = \frac{\partial S^\alpha(U)}{\partial U^\Phi}. \tag{13.88}$$

The covariant metric tensor $U_{\Phi\Psi}$ is defined by

$$U_{\Phi\Psi} = S_{\alpha\beta} S^\alpha_\Phi S^\beta_\Psi. \tag{13.89}$$

Despite the outward similarity of equations (13.89) and (13.11), their interpretations are completely different! Equation (13.11) is not a definition of the metric tensor $U_{\Phi\Psi}$, but rather a consequence of the definition (13.10). A definition based on the covariant basis is not possible here, since we are considering embedding the curve with respect to the surface and ignoring the larger ambient Euclidean space.

The next few steps follow a familiar pattern. The contravariant metric tensor $U^{\Phi\Psi}$ is the matrix inverse of $U_{\Phi\Psi}$:

$$U^{\Phi\Psi} U_{\Psi\Omega} = \delta^\Phi_\Omega. \tag{13.90}$$

From equation (13.89), we find

$$U_{\Phi\Psi} = S_{\Phi\beta} S^\beta_\Psi, \tag{13.91}$$

and

$$S^\Phi_\beta S^\beta_\Psi = \delta^\Phi_\Psi. \tag{13.92}$$

The Christoffel symbol $\Gamma^\Theta_{\Phi\Psi}$ also cannot be defined in the extrinsic fashion, as was done in equation (13.18). We must instead accept the intrinsic definition

$$\Gamma^{\Theta}_{\Phi\Psi} = \frac{1}{2}U^{\Theta\Omega}\left(\frac{\partial U_{\Omega\Phi}}{\partial U^{\Psi}} + \frac{\partial U_{\Omega\Psi}}{\partial U^{\Phi}} - \frac{\partial U_{\Phi\Psi}}{\partial U^{\Omega}}\right). \tag{13.18}$$

The curve Christoffel symbol $\Gamma^{\Theta}_{\Phi\Psi}$ is related to the surface Christoffel symbol $\Gamma^{\alpha}_{\beta\Gamma}$ by the relationship

$$\Gamma^{\Omega}_{\Phi\Psi} = \Gamma^{\alpha}_{\beta\Gamma}S^{\Omega}_{\alpha}S^{\beta}_{\Phi}S^{\Gamma}_{\Psi} + \frac{\partial S^{\alpha}_{\Phi}}{\partial S^{\Psi}}S^{\Omega}_{\alpha}. \tag{13.93}$$

The definition of the covariant derivative reads

$$\nabla_{\Theta}T^{\alpha\Phi}_{\beta\Psi} = \frac{\partial T^{\alpha\Phi}_{\beta\Psi}}{\partial U^{\Theta}} + S^{\gamma}_{\Theta}\Gamma^{\alpha}_{\gamma\omega}T^{\omega\Phi}_{\beta\Psi} - S^{\gamma}_{\Theta}\Gamma^{\omega}_{\gamma\beta}T^{\alpha\Phi}_{\omega\Psi} + \Gamma^{\Phi}_{\Theta\Omega}T^{\alpha\Omega}_{\beta\Psi} - \Gamma^{\Omega}_{\Theta\Psi}T^{\alpha\Phi}_{\beta\Omega}. \tag{13.94}$$

This derivative satisfies all the familiar properties.

The *length element* or *line element* $\sqrt{U}$ is defined as the square root of the determinant of the 1×1 covariant metric tensor $U_{\Phi\Psi}$:

$$\sqrt{U} = \sqrt{|U_{\cdot\cdot}|}. \tag{13.95}$$

The Levi-Civita symbols ε_{Φ} and ε^{Φ} are defined by

$$\varepsilon_{\Phi} = \sqrt{U} \tag{13.96}$$

$$\varepsilon^{\Phi} = \frac{1}{\sqrt{U}}. \tag{13.97}$$

The curve, being a hypersurface with respect to an ambient surface, has a well-defined normal n^{α} given by

$$n^{\alpha} = \varepsilon^{\alpha\beta}\varepsilon_{\Phi}S^{\Phi}_{\beta}. \tag{13.98}$$

We have already seen a definition analogous to this one in Sect. 10.7, where the normal was defined for curves embedded in two-dimensional planes. Since in this definition, nothing relies on the flatness of the ambient space, it remains valid for hypersurfaces embedded in Riemann spaces. The normal n^{α} is unit length

$$n_{\alpha}n^{\alpha} = 1 \tag{13.99}$$

and is orthogonal to the tangent space

$$n_{\alpha}S^{\alpha}_{\Phi} = 0. \tag{13.100}$$

We next the curvature tensor $b_{\Phi\Psi}$ according to the equation

$$\nabla_{\Psi}S^{\alpha}_{\Phi} = b_{\Phi\Psi}n^{\alpha}. \tag{13.101}$$

This equation is identical to (11.42) which, in turn, is an immediate consequence of the definition (11.16). However, the basis for equation (11.16) was the fact that $\nabla_\alpha \mathbf{S}_\beta$ is orthogonal to the tangent space spanned by $\mathbf{S}_\alpha$ and thus colinear with $\mathbf{N}$. If you recall, the demonstration of that fact relied on the extrinsic definition (10.80) of the Christoffel symbol. Therefore, in order to use equation (13.101) as the definition of curvature, we must show that $\nabla_\Psi S^\alpha_\Phi$ is orthogonal to the tangent space by using the intrinsic definition (13.18) of the Christoffel tensor or the equivalent equation (13.93).

The analytical expression for $\nabla_\Psi S^\alpha_\Phi$ being orthogonal to the tangent space reads:

$$S^\Theta_\alpha \nabla_\Psi S^\alpha_\Phi = 0. \tag{13.102}$$

Let us expand the covariant derivative $\nabla_\Psi S^\alpha_\Phi$

$$\nabla_\Psi S^\alpha_\Phi = \frac{\partial S^\alpha_\Phi}{\partial U^\Psi} + S^\beta_\Psi \Gamma^\alpha_{\beta\gamma} S^\gamma_\Phi - \Gamma^\Omega_{\Phi\Psi} S^\alpha_\Omega. \tag{13.103}$$

and contract both sides with S^Θ_α

$$S^\Theta_\alpha \nabla_\Psi S^\alpha_\Phi = S^\Theta_\alpha \frac{\partial S^\alpha_\Phi}{\partial U^\Psi} + \Gamma^\alpha_{\beta\gamma} S^\Theta_\alpha S^\beta_\Psi S^\gamma_\Phi - \Gamma^\Omega_{\Phi\Psi} S^\alpha_\Omega S^\Theta_\alpha. \tag{13.104}$$

According to equation (13.93), the first two terms produce $\Gamma^\Theta_{\Phi\Psi}$. The last term yields $-\Gamma^\Theta_{\Phi\Psi}$ since $S^\alpha_\Omega S^\Theta_\alpha = \delta^\Theta_\Omega$ and equation (13.102) is confirmed.

We are therefore able to accept equation (13.101) as the definition of the curvature tensor $b_{\Phi\Psi}$. The invariant

$$b^\Phi_\Phi, \tag{13.105}$$

is called the *geodesic curvature*. Of course, geodesic curvature is the same as the mean curvature. However, the term *geodesic curvature* is applied to one-dimensional curves while *mean curvature* is applied to hypersurfaces in higher dimensions.

At a particular point on a curve we have three curvature tensors in play: the curvature tensor B^α_α of the host surface, the curvature tensor $b_{\Phi\Psi}$ characteristic of the curve's embedding in the surface, and the vector curvature tensor $\mathbf{B}_{\Phi\Psi}$ characteristic of the curve's embedding in the ambient Euclidean space. These three tensors and linked by an identity that we are about to derive.

Express the equation of the curve (13.6) as a composition of the equation of the surface (10.3) and the equation of the curve (13.87) within the surface

$$Z^i(U) = Z^i(S(U)). \tag{13.106}$$

Differentiating this identity with respect to U^Φ yields a relationship among the three shift tensors Z^i_Φ, Z^i_α, and S^α_Φ:

$$Z^i_\Phi = Z^i_\alpha S^\alpha_\Phi. \tag{13.107}$$

The vector form of this equation is obtained by contracting both sides of this equation with the ambient covariant basis $\mathbf{Z}_i$:

$$\mathbf{U}_\Phi = \mathbf{S}_\alpha S^\alpha_\Phi. \tag{13.108}$$

Apply the covariant derivative ∇_Ψ to this identity:

$$\nabla_\Psi \mathbf{U}_\Phi = \nabla_\Psi \mathbf{S}_\alpha \ S^\alpha_\Phi + \mathbf{S}_\alpha \nabla_\Psi S^\alpha_\Phi \tag{13.109}$$

The combination on the left-hand side is $\mathbf{B}_{\Phi\Psi}$. The tensor $\nabla_\Psi \mathbf{S}_\alpha$ is expanded by the chain rule

$$\nabla_\Psi \mathbf{S}_\alpha = S^\beta_\Psi \nabla_\beta \mathbf{S}_\alpha = \mathbf{N} B_{\alpha\beta} S^\beta_\Psi. \tag{13.110}$$

Finally, $\nabla_\Psi S^\alpha_\Phi$ equals $b_{\Phi\Psi} n^\alpha$. Putting these elements together yields the identity

$$\mathbf{B}_{\Phi\Psi} = \mathbf{N} B_{\alpha\beta} S^\alpha_\Phi S^\beta_\Psi + \mathbf{S}_\alpha n^\alpha b_{\Phi\Psi}. \tag{13.111}$$

We are interested in one particular corollary of this rich and beautiful relationship. It states that $B_{\alpha\beta} S^\alpha_\Phi S^\beta_\Psi$ is the normal component and $n^\alpha b_{\Phi\Psi}$ are the tangential components of the vector curvature tensor $\mathbf{B}_{\Phi\Psi}$. Let us focus on the normal component, which can be isolated by dotting both sides of this identity with the normal $\mathbf{N}$:

$$\mathbf{N} \cdot \mathbf{B}_{\Phi\Psi} = B_{\alpha\beta} S^\alpha_\Phi S^\beta_\Psi \tag{13.112}$$

The tensor $\mathbf{N} \cdot \mathbf{B}_{\Phi\Psi}$ is called the *principal curvature tensor*. While vector curvature tensor $\mathbf{B}_{\Phi\Psi}$ depends on the curve's embedding in the ambient space, equation (13.112) tells us that its normal component $\mathbf{N} \cdot \mathbf{B}_{\Phi\Psi}$ depends only on the surface curvature tensor and the manner of the embedding.

As the next step, raise the index Ψ and contract with Φ:

$$\mathbf{N} \cdot \mathbf{B}^\Phi_\Phi = B_{\alpha\beta} S^\alpha_\Phi S^{\beta\Phi} \tag{13.113}$$

By equation (10.31) interpreted in the current context, the right-hand side becomes

$$\mathbf{N} \cdot \mathbf{B}^\Phi_\Phi = B^\alpha_\alpha - B_{\alpha\beta} n^\alpha n^\beta. \tag{13.114}$$

This equation leads to numerous geometric insights. The left-hand side can be rewritten as $\kappa \mathbf{N} \cdot \mathbf{P}$ where κ is the curvature and $\mathbf{P}$ is the principal normal. The right-hand side can be rewritten as $\left(B^\gamma_\gamma S_{\alpha\beta} - B_{\alpha\beta}\right) n^\alpha n^\beta$. We have

$$\kappa \mathbf{N} \cdot \mathbf{P} = \left(B_\gamma^\gamma S_{\alpha\beta} - B_{\alpha\beta} \right) n^\alpha n^\beta. \tag{13.115}$$

Let us limit our attention to curves where the principal normal $\mathbf{P}$ is aligned with the surface normal $\mathbf{N}$, which would, in fact, make them equal. For such curves, equation (13.115) becomes

$$\kappa = \left(B_\gamma^\gamma S_{\alpha\beta} - B_{\alpha\beta} \right) n^\alpha n^\beta, \tag{13.116}$$

which is a very interesting formula that states the curvature κ, characteristic of the curve's embedding in the larger Euclidean space, depends on the metric and curvature tensors of the surface and the orientation—given by n^α—of the curve.

Let us ask the following question which, in 1760, was asked by Euler: Among all possible embedded curves obtained by cutting the surface by an orthogonal plane, which curves have the highest and the lowest curvatures? Equation (13.116) holds the answer to this question.

By choosing different planes, we are essentially choosing n^α. Therefore, algebraically, the question becomes: what are the extreme values of

$$\left(B_\gamma^\gamma S_{\alpha\beta} - B_{\alpha\beta} \right) n^\alpha n^\beta \tag{13.117}$$

subject to the condition that n^α is unit length

$$S_{\alpha\beta} n^\alpha n^\beta = 1. \tag{13.118}$$

This is a fundamental question in linear algebra. It is well-known that the extremal values and directions are given by the eigenvalue problem

$$\left(B_\gamma^\gamma S_{\alpha\beta} - B_{\alpha\beta} \right) n^\beta = \lambda S_{\alpha\beta} n^\beta. \tag{13.119}$$

Juggling the indices yields

$$\left(B_\gamma^\gamma \delta_\beta^\alpha - B_\beta^\alpha \right) n^\beta = \lambda n^\alpha. \tag{13.120}$$

Therefore, the problem is reduced to finding the eigenvalues and eigenvectors of the matrix

$$B_\gamma^\gamma \delta_\beta^\alpha - B_\beta^\alpha. \tag{13.121}$$

The simple argument we are about to make is best seen in plane matrix notation. Denote the tensor B_β^α by the matrix B and let its eigenvalues be $\kappa_1$and κ_2. Then the tensor $B_\gamma^\gamma \delta_\beta^\alpha - B_\beta^\alpha$ is represented by the matrix

$$\left(\kappa_1 + \kappa_2 \right) I - B, \tag{13.122}$$

which has the same eigenvectors as B and the corresponding eigenvalues are

$$(\kappa_1 + \kappa_2) - \kappa_1 = \kappa_2 \tag{13.123}$$

$$(\kappa_1 + \kappa_2) - \kappa_2 = \kappa_1 \tag{13.124}$$

Therefore, the extremal values of the quadratic form (13.117) are the invariant eigenvalues κ_1 and κ_2 of the curvature tensor B^α_β. These eigenvalues are known as the *principal curvatures* of the surface. Their sum equals the mean curvature:

$$B^\alpha_\alpha = \kappa_1 + \kappa_2 \tag{13.125}$$

and their product equals Gaussian curvature:

$$K = \kappa_1 \kappa_2. \tag{13.126}$$

Furthermore, due to the symmetry of $B_{\alpha\beta}$ the corresponding directions are orthogonal.

13.10 Geodesics

The subject of embedded curves usually includes a discussion of geodesics. Given two points A and B on the surface, the *geodesic* is the curve of least length connecting A and B. Geodesics satisfy a set of second-order equations known as the geodesic equations. These equations are usually derived [31] by following the standard Euler–Lagrange procedure. We pursue an entirely different approach based on the calculus of moving surface. We therefore postpone the subject of geodesics until Chap. 17.

13.11 Summary

In this chapter, we studied curves from two perspectives: as general submanifolds of codimension greater than one, and as submanifold of dimension one. The former perspective led to the concepts of the curvature normal tensor $\mathbf{B}^\Phi_\Psi$ which reduces to $\mathbf{N}B^\Phi_\Psi$ for hypersurfaces. The latter perspective led to an equivalent concept of the principal normal $\mathbf{P}$, the binormal $\mathbf{Q}$ as well as higher-order derivatives with respect to arc length s.

Finally, we analyzed curves embedded in surfaces and recovered Euler's definitions of principal curvature.

Chapter 14
Integration and Gauss's Theorem

14.1 Preview

In this chapter, we pursue two goals. First, we discuss integration from a geometric point of view and establish the tensor calculus way of representing invariant integrals in arbitrary coordinates. Second, we prove a rather general form of Gauss's theorem. The starting point for the derivation is Gauss's theorem over flat domains referred to as Cartesian coordinates. Our task is to extend that result to arbitrary curved patches.

This chapter first defines a single kind of invariant integral in which a physical quantity is integrated over a geometric domain. The domain may be a section of a curve, a surface patch, or a three-dimensional domain. In all cases, the integral is defined by a procedure in which the domain is broken up into ever smaller parts and the limit of a finite sum is considered. All types of integrals (work, flux, circulation, etc.) encountered in applications can be reduced to an invariant integral.

Gauss's theorem and Stokes' theorem transform integrals over closed domains to integrals over the boundaries of those domains. Both theorems are multidimensional generalizations of the fundamental theorem of calculus. We demonstrate that the two theorems are closely related. The formulation of these theorems took several decades and efforts from some of the nineteenth-century's brightest scientific minds, including Gauss, Ostrogradsky, Cauchy, Poisson, Maxwell, and Riemann [24].

14.2 Integrals in Applications

In applications, integrals represent physical and geometric quantities. For example, the total mass M of a body Ω with density distribution ρ is given by the volume integral

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_14,

$$M = \int_{\Omega} \rho d\Omega. \tag{14.1}$$

The kinetic energy T of a fluid with density distribution ρ and velocity field $\mathbf{V}$ is given by the integral

$$T = \frac{1}{2} \int_{\Omega} \rho \mathbf{V} \cdot \mathbf{V} d\Omega. \tag{14.2}$$

The flux F of the fluid across a surface S is given by the surface integral

$$F = \int_{S} \mathbf{V} \cdot \mathbf{N} dS. \tag{14.3}$$

The total charge Q distributed with density σ over the surface S of a conductor is given by the surface integral

$$Q = \int_{S} \sigma dS. \tag{14.4}$$

The total force $\mathbf{F}$ of pressure exerted on a body Ω with boundary S immersed in fluid with pressure distribution P is given by the surface integral

$$\mathbf{F} = -\int_{S} P \mathbf{N}\, dS. \tag{14.5}$$

The total curvature T of a surface S is given by

$$T = \int_{S} K dS, \tag{14.6}$$

where K is the Gaussian curvature.

The total gravitational energy V of a string U with linear mass density ρ suspended in a gravitational field with potential P is given by the contour integral

$$V = \int_{U} \rho P \, dU. \tag{14.7}$$

Each of the integrals in these examples has the form

$$\int_{\text{Geometric Domain}} \text{Invariant.} \tag{14.8}$$

The definition of the integral (14.8) involves a limiting process, in which the domain (whether it is Ω, S, or U) is divided into small pieces and a finite sum approaches a limit. This procedure is nontrivial, but its independence from coordinates is evident.

14.3 The Arithmetic Space

The definition of the physical integral

$$\int_{\Omega} F d\Omega, \quad \int_{S} F dS, \quad \int_{U} F dU \tag{14.9}$$

is entirely coordinate free. However, in order to evaluate a physical integral, one almost always needs to introduce a coordinate system and transform the invariant integral into a repeated *arithmetic integral*. The transformation to a repeated integral consists of three steps. First, the physical field F is expressed as a function $F(Z)$, $F(S)$ or $F(U)$ with respect to the coordinates. Second, the domain of integration is translated into the corresponding arithmetic domain. Third, an additional factor is introduced that adjusts for the length, area, or volume mismatch between the physical domain and its arithmetic representation.

Consider an example. Let F be the scalar field defined as the distance squared from a point P. What is the integral of F over a unit disk Ω with a center that is a distance $\sqrt{2}$ away from P? This problem is illustrated in Fig. 14.1.

Can you solve this problem without introducing a coordinate system? Perhaps you could, but given the available power of coordinates, it is hardly worth the effort. Keep in mind that coordinate systems are a recent luxury. Ancient Greeks, in particular Archimedes of Syracuse whose portrait appears in Fig. 14.2, did not have analytical methods at their disposal, yet were remarkably good at evaluating areas and volumes. Famously, Archimedes discovered that the volume of a sphere is $4\pi/3$. The answer to the problem in Fig. 14.1 is $5\pi/2$, so perhaps Archimedes could solve it geometrically, as well. However, with the benefit of coordinates, one need not be Archimedes to solve this problem.

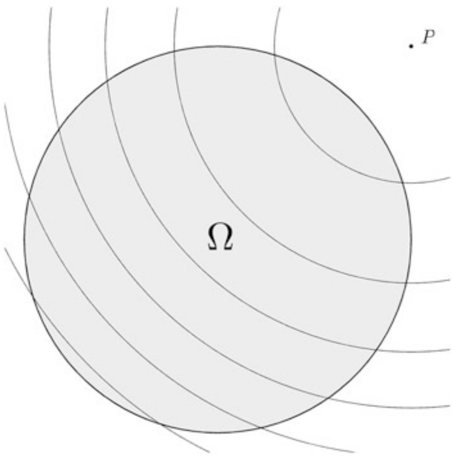


Fig. 14.1 A unit disk Ω and a point P a distance $\sqrt{2}$ from the center of the disk. The scalar field F is defined as the distance squared from the point P. The firgure includes six level lines of the function F for values betwen $1/2$ and $11/2$

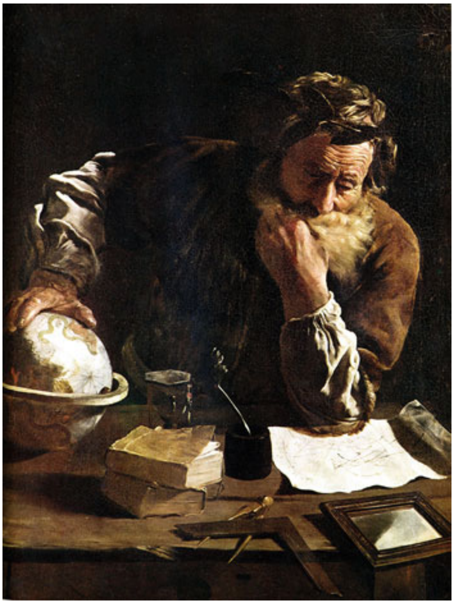


Fig. 14.2 This 1620 work by the Italian Baroque painter Domenico Fetti (c. 1589–1623) depicts Archimedes toiling over a problem in natural science without the benefit of a coordinate system

We ought to select a coordinate system that is well-suited for the particular problem at hand. Let us first consider Cartesian coordinates x, y with the origin at the center and coordinates $(1, 1)$ at the point P. Then the domain Ω is expressed by the relatively simple integration and the integral is given by

$$\int_\Omega F d\Omega = \int_{-1}^{1} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \left((x-1)^2 + (y-1)^2\right) dxdy \tag{14.10}$$

If we shift the coordinate system to the point P, then the integrand simplifies, while the limits become more complicated

$$\int_\Omega F d\Omega = \int_{-2}^{0} \int_{-1-\sqrt{1-(y+1)^2}}^{-1+\sqrt{1-(y+1)^2}} \left(x^2 + y^2\right) dxdy. \tag{14.11}$$

These integrals are not particularly easy to evaluate. That is because Cartesian coordinates are usually not the best choice when describing circular geometries.

A significant simplification comes from using polar coordinates. If the origin of the polar coordinates r, θ is at the center of the disk and the point P is at $\theta = 0$, then the domain Ω corresponds to a simple rectangle $(r, \theta) \in [0, 1] \times [0, 2\pi]$ in the arithmetic space. The distance squared is given by the law of cosines

$$F(r, \theta) = 2 + r^2 - 2\sqrt{2} r \cos\theta \tag{14.12}$$

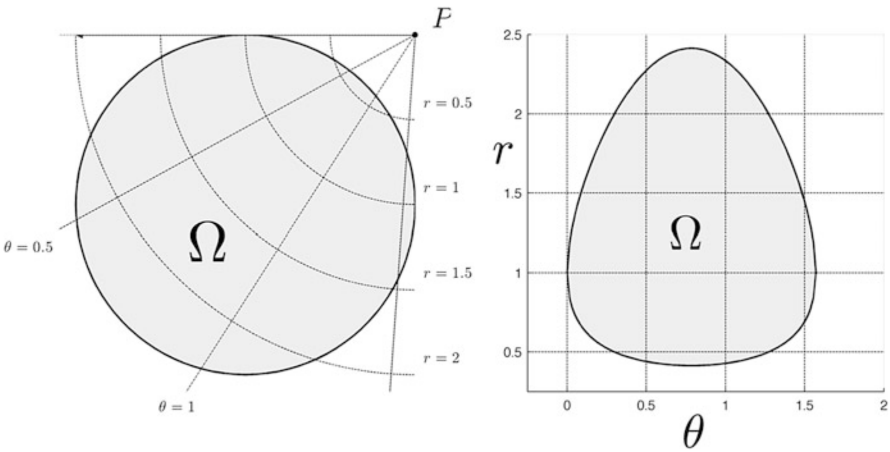


Fig. 14.3 With respect to the polar coordinate system, illustrated on the left, the domain Ω corresponds to the *arithmetic domain* shown on the right

and the invariant integral is given by repeated arithmetic integral

$$\int_{\Omega} F d\Omega = \int_0^{2\pi} \int_0^1 \left(2 + r^2 - 2\sqrt{2}r\cos\theta\right) r dr d\theta, \tag{14.13}$$

which is easy to evaluate. From multivariable calculus, the reader is familiar with the additional factor of r in the integrand. It is the adjustment for the distortion in area between the arithmetic and the geometric domains. Its interpretation from the tensor calculus point of view is discussed below.

The polar coordinate system that we just considered is perhaps the best choice for the posed problem because it leads to the simplest repeated integral. Nevertheless, let us consider one more polar coordinate system. It is not as convenient for the problem we have just solved, but it is better suited for the illustration of the concept of the arithmetic space. Place the origin of the coordinate system at P and direct the polar axis in the direction tangent to the disk as illustrated on the left of Fig. 14.3.

In this polar coordinate system, the field F has the simplest form: $F(r,\theta) = r^2$. The main challenge is to identify the limits of integration. It is evident that θ changes from 0 to $\pi/2$. It is not hard to demonstrate that, for each θ, r varies from $\sqrt{2}\cos\left(\frac{\pi}{4} - \theta\right) - \sqrt{\sin 2\theta}$ to $\sqrt{2}\cos\left(\frac{\pi}{4} - \theta\right) + \sqrt{\sin 2\theta}$. Therefore, the conversion to arithmetic integral is

$$\int_{\Omega} F d\Omega = \int_0^{\pi/2} \int_{\sqrt{2}\cos\left(\frac{\pi}{4}-\theta\right)-\sqrt{\sin 2\theta}}^{\sqrt{2}\cos\left(\frac{\pi}{4}-\theta\right)+\sqrt{\sin 2\theta}} r^2 \, r dr d\theta. \tag{14.14}$$

The repeated integral on the right hand side of equation (14.14) is called arithmetic because it involves algebraic expressions rather than geometric quantities. It can be evaluated by the analytical techniques that the reader learned while studying multivariable calculus. The immediate graphical interpretation of equation (14.14), illustrated on the right of Fig. 14.3, is Cartesian—thus the term *arithmetic* which is often used synonymously with $\mathbb{R}^n$.

This Cartesian feature of the plot on the right may not seem particularly significant at this point. After all, we could have chosen a Cartesian coordinate system to begin with. Instead, we chose a curvilinear coordinate system with curved coordinate lines. Subsequently, the geometric interpretation of the arithmetic integral (14.14) "straightened out" the coordinate lines. Both the original domain Ω and the arithmetic domain are Euclidean, and the curvature of the coordinate lines is only due to our choice of coordinates. Looking ahead to studying curved surfaces, **the Euclidean nature of the arithmetic domain and the associated Cartesian interpretation of the coordinates will take on much greater significance**.

14.4 The Invariant Arithmetic Form

We now discuss the crucial factor of r in equations (14.13) and (14.14). Most calculus textbooks include an illustration of the fact that $rdrd\theta$ represents the area of a small coordinate patch. Therefore, the factor of r can be thought of as the metric adjustment factor. The corresponding tensor object is the volume element $\sqrt{Z}$ introduced in Sect. 9.7, which may be referred to as the area element and length element depending on the context. Tensor calculus offers a simple and general recipe for expressing invariant integrals in arithmetic form. It is captured by the equation

$$\int_{\Omega} F d\Omega = \int_{A_3}^{B_3} \int_{A_2}^{B_2} \int_{A_1}^{B_1} F(Z) \sqrt{Z} dZ^1 dZ^2 dZ^3, \tag{14.15}$$

where the integration limits satisfy $A_1 < B_1$, $A_2 < B_2$ and $A_3 < B_3$ and are chosen so that the integration domain is properly captured.

We must demonstrate that the repeated arithmetic integral in equation (14.15) is invariant under a change of variables. We limit our demonstration to orientation-preserving coordinate changes. This orientation-preserving property of a coordinate change is given in Sect. 4.10. The explanation is based on the corresponding theorem from multivariate calculus. According to the change-of-variables theorem, when going from one coordinate system (unprimed) to another (primed), an integral transforms according to the following rule

$$\int_{A_3}^{B_3} \int_{A_2}^{B_2} \int_{A_1}^{B_1} F(Z) dZ^1 dZ^2 dZ^3 = \int_{A_{3'}}^{B_{3'}} \int_{A_{2'}}^{B_{2'}} \int_{A_{1'}}^{B_{1'}} F(Z') |J|^{-1} dZ^{1'} dZ^{2'} dZ^{3'}. \tag{14.16}$$

In words, the integrand needs to be re-expressed in terms of the new (primed) coordinate systems and multiplied by the absolute value of the determinant of the coordinate transformation and the limits of integration must be appropriately adjusted. Let us apply this rule to the integral in equation (14.15). Going from unprimed coordinates to the primed coordinates, we have

$$\int_{A_3}^{B_3}\int_{A_2}^{B_2}\int_{A_1}^{B_1} F(Z)\sqrt{Z}dZ^1dZ^2dZ^3$$
$$= \int_{A_{3'}}^{B_{3'}}\int_{A_{2'}}^{B_{2'}}\int_{A_{1'}}^{B_{1'}} F\left(Z\left(Z'\right)\right)\sqrt{Z}J^{-1}dZ^{1'}dZ^{2'}dZ^{3'} \qquad (14.17)$$

Recall from Sect. 9.9 that the volume element $\sqrt{Z}$ is a relative invariant of weight -1, that is

$$\sqrt{Z'} = J^{-1}\sqrt{Z}. \qquad (14.18)$$

Since $F(Z') = F(Z(Z'))$, we have

$$\int_{A_3}^{B_3}\int_{A_2}^{B_2}\int_{A_1}^{B_1} F(Z)\sqrt{Z}dZ^1dZ^2dZ^3 = \int_{A_{3'}}^{B_{3'}}\int_{A_{2'}}^{B_{2'}}\int_{A_{1'}}^{B_{1'}} F\left(Z'\right)\sqrt{Z'}dZ^{1'}dZ^{2'}dZ^{3'} \qquad (14.19)$$

which confirms the invariant nature of the integral in equation (14.15).

14.5 Gauss's Theorem

Gauss's theorem is of fundamental importance in applied mathematics. After all, it is a multidimensional generalization of the fundamental theorem of calculus. For a domain Ω with boundary S, Gauss's theorem reads

$$\int_\Omega \nabla_i T^i d\Omega = \int_S N_i T^i dS. \qquad (14.20)$$

It holds for a sufficiently smooth tensor field T^i. Importantly, Gauss's theorem holds in Riemannian as well as Euclidean spaces. For instance, on a curved patch S bounded by a contour U, Gauss's theorem for a tensor T^α reads

$$\int_S \nabla_\alpha T^\alpha dS = \int_U n_\alpha T^\alpha dU. \qquad (14.21)$$

The applications of this theorem are too numerous to even begin to mention. In this book, its most frequent application will be in the context of calculus of variations.

Let us assume that Gauss's theorem holds in the Euclidean space and prove its more general Riemannian form (14.21). Convert the volume integral to the arithmetic form

$$\int_\Omega \nabla_i T^i d\Omega = \iiint \nabla_i T^i \sqrt{Z} dZ^1 dZ^2 dZ^3 \tag{14.22}$$

and expand the covariant derivative according to its definition

$$\int_\Omega \nabla_i T^i d\Omega = \iiint \left(\frac{\partial T^i}{\partial Z^i} \sqrt{Z} + \Gamma^i_{ik} T^k \sqrt{Z} \right) dZ^1 dZ^2 dZ^3 \tag{14.23}$$

Transform the first term in the integrand by the product rule

$$\int_\Omega \nabla_i T^i d\Omega = \iiint \left(\frac{\partial \left(T^i \sqrt{Z} \right)}{\partial Z^i} - T^i \frac{\partial \sqrt{Z}}{\partial Z^i} + \Gamma^i_{ik} T^k \sqrt{Z} \right) dZ^1 dZ^2 dZ^3 \tag{14.24}$$

Sincev

$$\frac{\partial \sqrt{Z}}{\partial Z^i} = \sqrt{Z} \Gamma^k_{ki}, \tag{14.25}$$

the last two terms cancel, leaving the single term

$$\int_\Omega \nabla_i T^i d\Omega = \iiint \frac{\partial \left(T^i \sqrt{Z} \right)}{\partial Z^i} dZ^1 dZ^2 dZ^3. \tag{14.26}$$

This term is subject to Gauss's theorem in the arithmetic space. After all, the arithmetic space is a Euclidean space referred to as Cartesian coordinates. Thus,

$$\int_\Omega \nabla_i T^i d\Omega = \int_{\bar{S}} \bar{N}_i T^i \sqrt{Z} d\bar{S}, \tag{14.27}$$

where the bar above the letter indicates that the symbol refers to an object in the arithmetic space. Express the surface integral in this equation in arithmetic form

$$\int_\Omega \nabla_i T^i d\Omega = \int_{\bar{S}} \bar{N}_i T^i \sqrt{Z} \sqrt{\bar{S}} dS^1 dS^2. \tag{14.28}$$

Let us determine the relationship between the arithmetic and the actual normals $\bar{N}_i$ and N_i. These normals are given by the expressions

$$\bar{N}_i = \frac{1}{2} \bar{\varepsilon}_{ijk} \bar{\varepsilon}^{\alpha\beta} Z^j_\alpha Z^k_\beta, \text{ and} \tag{14.29}$$

$$N_i = \frac{1}{2} \varepsilon_{ijk} \varepsilon^{\alpha\beta} Z^j_\alpha Z^k_\beta. \tag{14.30}$$

Note that there are no bars over the shift tensors. The entries of the shift tensor are derived from the parametric equations that describe the boundary—and those are the same in the arithmetic and the actual spaces. Thus the two normals are related by the identity

$$\bar{N}_i \frac{\sqrt{\bar{S}}}{\sqrt{\bar{Z}}} = N_i \frac{\sqrt{S}}{\sqrt{Z}}, \tag{14.31}$$

where $\bar{Z} = 1$ because the coordinates are Cartesian in the arithmetic space. Therefore, equation (14.27) can be rewritten as

$$\int_\Omega \nabla_i T^i d\Omega = \iint N_i T^i \sqrt{S} dS^1 dS^2, \tag{14.32}$$

which is, in arithmetic form, precisely the statement of Gauss's theorem:

$$\int_\Omega \nabla_i T^i d\Omega = \int_S N_i T^i dS. \tag{14.33}$$

14.6 Several Applications of Gauss's Theorem

As the first striking application let us show that the integral of the normal over closed surface vanishes:

$$\int_S \mathbf{N} dS = \mathbf{0}. \tag{14.34}$$

The integral of the normal may be interpreted as the total force on a body immersed in a uniform pressure field. The demonstration of equation (14.34) is entirely straightforward. Since the integral in question can be written as

$$\int_S \mathbf{N} dS = \int_S N^i \mathbf{Z}_i dS, \tag{14.35}$$

an application of Gauss's theorem yields

$$\int_S \mathbf{N} dS = \int_\Omega \nabla^i \mathbf{Z}_i \, d\Omega \tag{14.36}$$

and the integrand vanishes due to the metrinilic property of the covariant derivative.

As another even more striking and, perhaps, even more straightforward application, let us show that the integral of curvature normal over vanishes over a closed surface

$$\int_S \mathbf{N} B^\alpha_\alpha dS = 0. \tag{14.37}$$

This identity is a near triviality since the curvature normal equals the surface Laplacian of the position vector

$$\mathbf{N}B_\alpha^\alpha = \nabla_\alpha \nabla^\alpha \mathbf{R}. \tag{14.38}$$

Thus, by Gauss's theorem,

$$\int_S \nabla_\alpha \nabla^\alpha \mathbf{R}\, dS = \int_U n_\alpha \nabla^\alpha \mathbf{R}\, dU = \int_U \mathbf{n}\, dU, \tag{14.39}$$

where U is the nonexistent boundary of the closed surface S. Thus, equation (14.37) holds.

If the same integral is considered over a patch rather than a closed surface, then the result

$$\int_S \mathbf{N}B_\alpha^\alpha dS = \int_U \mathbf{n}\, dU \tag{14.40}$$

gives a vivid geometric interpretation of mean curvature: Mean curvature measures the degree to which the contour boundary normal field $\mathbf{n}$ points consistently out of the plane. Furthermore, the curvature normal is given by the limit

$$\mathbf{N}B_\alpha^\alpha = \lim_{S\to 0} \frac{\int_U \mathbf{n}\, dU}{\int_S dS}, \tag{14.41}$$

where the integrals are calculated for a patch S that is appropriately shrinking to a point at which the normal and the mean curvature are evaluated. Dotting both sides with $\mathbf{N}$ gives an explicit limit for the mean curvature

$$B_\alpha^\alpha = \mathbf{N} \cdot \lim_{S\to 0} \frac{\int_U \mathbf{n}\, dU}{\int_S dS}. \tag{14.42}$$

14.7 Stokes' Theorem

In order to demonstrate the Stokes' theorem, introduce the unit tangent vector $\mathbf{T}$. Its ambient components T_i are given by expression

$$T_i = \varepsilon_{ijk} N^j n^\alpha z_\alpha^k \tag{14.43}$$

Exercise 288. Show that the vector $\mathbf{T}$ with components given by (14.43) is indeed the unit tangent to the contour boundary. That is, $\mathbf{T}$ is unit length

$$\mathbf{T} \cdot \mathbf{T} = 1, \tag{14.44}$$

orthogonal to the surface normal $\mathbf{N}$

$$\mathbf{T}\cdot\mathbf{N}=0, \tag{14.45}$$

and the contour normal $\mathbf{n}$

$$\mathbf{T}\cdot\mathbf{n}=0. \tag{14.46}$$

Now, onto the proof of the Stokes' theorem. The contour integral is given by

$$\int_U F^i T_i \, dU = \int_U F^i \varepsilon_{ijk} N^j n^\alpha Z^k_\alpha dU, \tag{14.47}$$

where the integral on the right is immediately subject to Gauss's theorem:

$$\int_U F^i T_i \, dU = \int_S \nabla^\alpha \left(F^i \varepsilon_{ijk} N^j Z^k_\alpha\right) dS \tag{14.48}$$

Expand the integrand by the product rule

$$\int_U F^i T_i \, dU = \int_S \left(\nabla^\alpha F^i \varepsilon_{ijk} N^j Z^k_\alpha + F^i \varepsilon_{ijk} \nabla^\alpha N^j Z^k_\alpha + F^i \varepsilon_{ijk} N^j \nabla^\alpha Z^k_\alpha\right) dS, \tag{14.49}$$

and apply the appropriate differential identity to each term

$$\int_U F^i T_i \, dU = \int_S \left(Z^\alpha_m \nabla^m F^i \varepsilon_{ijk} N^j Z^k_\alpha - F^i \varepsilon_{ijk} Z^j_\beta B^{\alpha\beta} Z^k_\alpha + F^i \varepsilon_{ijk} N^j N^k B^\alpha_\alpha\right) dS. \tag{14.50}$$

The second and the third terms vanish due to the skew-symmetry of the Levi-Civita symbol ε_{ijk}. Thus, we are left with a single term

$$\int_U F^i T_i \, dU = \int_S \nabla^m F^i \varepsilon_{ijk} N^j Z^k_\alpha Z^\alpha_m dS. \tag{14.51}$$

By the projection formula (10.55), we have

$$\int_U F^i T_i \, dU = \int_S \nabla^m F^i \varepsilon_{ijk} N^j \delta^k_m dS - \nabla^m F^i \varepsilon_{ijk} N^j N^k N_m dS. \tag{14.52}$$

The term with the three normals vanishes due to the skew-symmetry of ε_{ijk} and the remaining term (with the indices rotated $i \to j \to k \to i$) gives us precisely the statement of Stokes' theorem

$$\int_U F^i T_i \, dU = \int_S \varepsilon_{ijk} \nabla^i F^j N^k dS, \text{ Q.E.D.} \tag{14.53}$$

Exercise 289. Explain why the term $\varepsilon_{ijk} Z^j_\beta Z^k_\alpha B^{\alpha\beta}$ vanishes.

14.8 Summary

This chapter was devoted to integration. Integration proved to be another topic that is well organized by the tensor framework. We have highlighted that physical integrals are expressed in coordinates by equation (14.15). Furthermore, integrals with sufficiently smooth integrands are governed by Gauss's theorem (14.21). Gauss's theorem has numerous applications. In this book, Gauss's theorem will find important applications in shape optimization and other problems with moving surfaces.

Part III
The Calculus of Moving Surfaces

Chapter 15
The Foundations of the Calculus of Moving Surfaces

15.1 Preview

Moving surfaces are ubiquitous in life and mathematics. Moving surfaces, like stationary surfaces, need their own language. We have shown that, in many ways, tensor calculus is an ideal language for describing stationary surfaces. For moving surfaces, the language of tensors must be extended to capture the particularities of moving surfaces. This extension comes in the form of the *calculus of moving surfaces* (CMS).

The term *moving surface* likely invokes an image of dynamically deforming physical surfaces, such as waves in water, soap films, biological membranes, or a fluttering flag. However, in applications, moving surfaces arise in numerous other contexts. For example, in *shape optimization*—a branch of the calculus of variations where the unknown quantity is the shape of a domain—moving surfaces arise as a parametrized family of allowable variations. In *shape perturbation theory*—I think this term is self-descriptive enough—moving surfaces arise as evolutions from the unperturbed to the perturbed configurations. Finally, moving surfaces can be effectively introduced in problems where one may not think that moving surfaces can play any role at all. For example, in Chap. 17 we use the calculus of moving surfaces to prove a special case of the celebrated Gauss–Bonnet theorem which states that the integral of Gaussian curvature over a closed surface is independent of the shape of the surface and only depends on its genus (the number of topological holes).

The fundamental ideas of the calculus of moving surfaces were introduced by the French mathematician Jacques Hadamard (Fig. 15.1) who studied the propagation of discontinuities in continuous media. The basic concepts were introduced in his *Mémoire sur le problème d'analyse relatif à l'équilibre des plaques elastiques encastrées* [21]. Since Hadamard's pioneering work, numerous contributions, large and small, were made by several distinguished applied mathematicians [16, 45].

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_15,

Fig. 15.1 The great French mathematician Jacques Hadamard (1865–1963) is the originator of the calculus of moving surfaces

15.2 The Kinematics of a Moving Surface

A moving surface $S(t)$ can be thought of as a family of surfaces parameterized by a time-like variable t. For simplicity, we refer to t simply as time, although in nonphysical scenarios t can be any parameter. We almost always drop the argument t and use the symbol S to denote a moving surface.

A moving surface can be described parametrically by a set of equations

$$Z^i = Z^i(t, S), \tag{15.1}$$

where S^α are the coordinates to which the surface is referred at time t. As always, we suppress the tensor index of function arguments. The ambient coordinates Z^i and the surface coordinates S^α are arbitrary, subject to sufficient smoothness conditions discussed in the next paragraph. Equation (15.1) is analogous to equation (10.3) for a stationary surface. At any particular moment of time t, the surface is subject to classical differential geometry analysis described in the preceding chapters.

We assume that surface coordinates evolve *smoothly*. Analytically, this means that $Z^i(t, S)$ is a smooth function of t. Intuitively, this implies that nearby points on nearby surfaces (i.e., surfaces corresponding to nearby times t and $t+h$) have nearby surface coordinates. This describes *continuity* while smoothness is stronger: our coordinates are not only evolving continuously, but differentiably to a sufficient number of orders. In particular, the trajectory $Z^i(t, S_0)$ of a point with fixed surface coordinates S_0^α is a smooth curve.

Equation (15.1) describe a family of invariant surfaces for a particular combination of ambient and surface coordinates. In the spirit of tensor calculus, we must enable ourselves to parameterize this family of surfaces in a completely arbitrary way, only subject to the smoothness requirements just discussed. Consider an alternative family of parametrizations

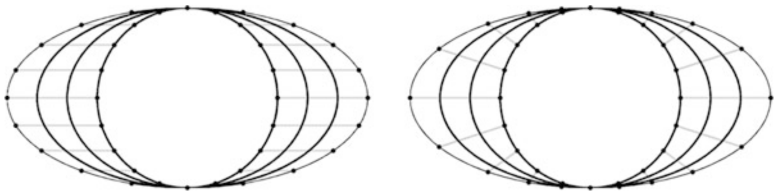


Fig. 15.2 The same evolution of an ellipse, in which the horizontal semiaxis equals At, is parameterized in two alternative ways. The figure on the left shows parametrization (15.5), while the figure on the right shows parametrization (15.6). The *dotted lines* show the trajectories of *coordinate particles*, that is, points of constant surface coordinates

$$Z^i = Z^i\left(t, S'\right), \tag{15.2}$$

where we have kept the ambient coordinates unchanged, but introduced new surface coordinates $S^{\alpha'}$. What is the relationships between the old coordinates S^α and the new coordinates $S^{\alpha'}$? The key point to realize is that the relationship between the two coordinate systems must depend on time:

$$S^{\alpha'} = S^{\alpha'}(t, S) \tag{15.3}$$

$$S^\alpha = S^\alpha\left(t, S'\right). \tag{15.4}$$

In other words, equations (10.27) and (10.28) would no longer be sufficient. The calculus of moving surfaces, following in the footsteps of tensor calculus, is adamant about constructing an analytical framework that is completely independent of parameterization, leaving the freedom to choose a parametrization to the analyst.

Consider the evolution of an ellipse in which one of the semiaxes is growing at a constant rate A. Figure 15.2 shows this evolution from $t = 1$ to $t = 2$. Perhaps the simplest way to parametrize this evolution is

$$\begin{cases} x(t,\alpha) = At\cos\alpha \\ y(t,\alpha) = \sin\alpha \end{cases}. \tag{15.5}$$

In this parameterization, the value of α does not correspond to the polar angle θ. The same evolution parameterized by θ appears as

$$\begin{cases} x(t,\theta) = \frac{At\cos\theta}{\sqrt{\cos^2\theta + A^2t^2\sin^2\theta}} \\ y(t,\theta) = \frac{At\sin\theta}{\sqrt{\cos^2\theta + A^2t^2\sin^2\theta}} \end{cases}. \tag{15.6}$$

The former parametrization is seen on the left of Fig. 15.2 and the latter on the right.

The relationship between $S^1 = \alpha$ and $S^{1'} = \theta$ is, of course, time-dependent

$$\theta(t, \alpha) = \operatorname{arccot}(At \cot \alpha) \tag{15.7}$$

and that is the key point. To reiterate: in order to assure the freedom to choose any parameterization for a moving surface, one must consider a time-dependent change of variables (15.3)–(15.4). If we were to restrict ourselves to time-independent changes of parametrization, then there would be no need for the new derivative operator $\dot{\nabla}$, which is at the heart of the calculus of moving surfaces.

It is the inescapable time dependence in the relationships (15.3) and (15.4) that poses the challenge that the invariant time derivative $\dot{\nabla}$ seeks to overcome. The crux of the problem is that the partial derivative $\partial/\partial t$ of a invariant $T(t, S)$ is not itself an invariant. We demonstrate this in the following paragraph. So our starting point for moving surfaces is more complicated than it was for stationary surfaces where the partial derivative $\partial/\partial S^\alpha$ produced a covariant tensor out of an invariant. Of course, matters become even more complicated when the partial time derivative $\partial/\partial t$ is applied to a tensor $T^{i\alpha}_{j\beta}$. As you would expect, the result $\partial T^{i\alpha}_{j\beta}(t, S)/\partial t$ is not a tensor.

We now show that the partial derivative $\partial/\partial t$ does not preserve the invariant property. Suppose that T is an invariant defined on a moving surface, such as the mean curvature B^α_α or the Gaussian curvature K. Let U be the partial derivative of T with respect to time

$$U = \frac{\partial T(t, S)}{\partial t}, \tag{15.8}$$

and evaluate U according to the same rule in the alternative coordinate system $S^{\alpha'}$

$$U' = \frac{\partial T(t, S')}{\partial t}. \tag{15.9}$$

In order to establish the relationship between U and U', differentiate the identity

$$T\left(t, S'\right) = T\left(t, S\left(t, S'\right)\right)$$

with respect to t. By the chain rule, we find

$$U' = \frac{\partial T(t, S)}{\partial t} + \frac{\partial T}{\partial S^\alpha}\frac{\partial S^\alpha(t, S')}{\partial t}. \tag{15.10}$$

Introduce the Jacobian-like object J^α_t defined as

$$J^\alpha_t\left(t, S'\right) = \frac{\partial S^\alpha(t, S')}{\partial t}. \tag{15.11}$$

The definition of the Jacobians $J^{\alpha}_{\alpha'}$ and $J^{\alpha'}_{\alpha}$ is the same as before except now these objects are functions of t:

$$J^{\alpha}_{\alpha'}\left(t,S'\right) = \frac{\partial S^{\alpha}\left(t,S'\right)}{\partial S^{\alpha'}} \tag{15.12}$$

$$J^{\alpha'}_{\alpha}\left(t,S\right) = \frac{\partial S^{\alpha'}\left(t,S\right)}{\partial S^{\alpha}}. \tag{15.13}$$

Using the newly defined object J^{α}_{t}, equation (15.10) can be rewritten as

$$U' = U + J^{\alpha}_{t}\nabla_{\alpha}T. \tag{15.14}$$

This identity indicates that U is not an invariant; after all, $U' \neq U$ because of the term $J^{\alpha}_{t}\nabla_{\alpha}T$. Thus, the tensor property is lost even for invariants. This is not the case for the partial derivatives $\partial/\partial Z^i$ and $\partial/\partial S^{\alpha}$, which produced covariant tensors out of invariants. In this chapter, we construct a new differential operator $\dot{\nabla}$ that preserves the invariant property. In the next chapter, the new operator is extended to arbitrary tensors.

Exercise 291. Calculate the Jacobians J^{α}_{t} and $J^{\alpha'}_{t}$ for the change of variables in equation (15.7).

15.3 The Coordinate Velocity V^i

In this section we introduce the coordinate velocity V^i. It is an essential ingredient in defining the invariant velocity C, which is the fundamental quantity in the calculus of moving surfaces. The *coordinate velocity* V^i is defined as

$$V^i = \frac{\partial Z^i\left(t,S\right)}{\partial t}. \tag{15.15}$$

It is the ambient component of the velocity of the coordinate particle with fixed coordinates S^{α}. To show this, note that the position vector $\mathbf{R}$ tracking the coordinate particle S^{α} is given by

$$\mathbf{R}\left(t,S_0\right) = \mathbf{R}\left(Z\left(t,S_0\right)\right). \tag{15.16}$$

Differentiating, we find

$$\frac{\partial \mathbf{R}\left(t,S_0\right)}{\partial t} = \frac{\partial \mathbf{R}}{\partial Z^i}\frac{\partial Z^i\left(t,S_0\right)}{\partial t}, \tag{15.17}$$

or

$$\mathbf{V} = V^i \mathbf{Z}_i, \tag{15.18}$$

confirming that V^i is the ambient coordinate of velocity.

The projection of the coordinate velocity onto the surface is called the *tangential coordinate velocity* V^α

$$V^\alpha = V^i Z_i^\alpha. \tag{15.19}$$

Exercise 292. Compute V^i for the parametrization (15.5) of the ellipse evolution.

Exercise 293. Compute V^i for the parametrization (15.6) of the ellipse evolution.

Exercise 294. Note that the answers obtained in the two preceding exercises are different, indicating that V^i is not a tensor with respect to changes in surface coordinates.

Exercise 295. Compute V^α for the parametrization (15.5) of the ellipse evolution.

Exercise 296. Compute $V^{\alpha'}$ for the parametrization (15.6) of the ellipse evolution.

Exercise 297. Are the two answers the same? What does that tell you about the tensor nature of V^α with respect to surface coordinate changes?

Exercise 298. Show that V^i is tensor with respect to changes in ambient coordinates.

Exercise 299. From the preceding exercise and equation (15.19), conclude that V^α is also a tensor with respect to changes in ambient coordinates.

Exercise 300. Show that the acceleration of a coordinate particle is given by

$$A^i = \frac{\partial V^i}{\partial t} + \Gamma^i_{jk} V^j V^k. \tag{15.20}$$

The preceding exercises dealt with one particular example of a change of surface coordinates. Let us now derive the general rules by which V^i and V^α transform under a simultaneous change of ambient and surface coordinates. In the primed coordinates, $V^{i'}$ is given by

$$V^{i'} = \frac{\partial Z^{i'}(t, S')}{\partial t}. \tag{15.21}$$

To relate V^i and $V^{i'}$, substitute the relationships between the coordinate system into equation (15.21)

$$V^{i'} = \frac{\partial Z^{i'}(Z(t, S(t, S')))}{\partial t}. \tag{15.22}$$

Expanding the right-hand side, we find by a repeated application of the chain rule:

$$V^{i'} = \frac{\partial Z^{i'}}{\partial Z^i}\left(\frac{\partial Z^i(t,S)}{\partial t} + \frac{\partial Z^i(t,S')}{\partial S^\alpha}\frac{\partial S^\alpha(t,S')}{\partial t}\right). \tag{15.23}$$

In other words,

$$V^{i'} = V^i J_i^{i'} + Z_\alpha^i J_i^{i'} J_t^\alpha. \tag{15.24}$$

Thus, we officially conclude that V^i is not a tensor with respect to changes in surface coordinates. Note two interesting aspects of equation (15.24). First, if we were to limit ourselves to time-independent changes of surface coordinates, then—since J_t^α would vanish—V^i would be tensor. Second, the nontensor part of the relationship in equation (15.24) is proportional to the shift tensor Z_α^i. Thus, the projection $V^i N_i$ has a legitimate chance at being an invariant. It is, in fact, an invariant and is known as the surface velocity C and is introduced in the next section.

Exercise 301. Show that coordinate velocities V^i and $V^{i'}$ associated with parametrizations (15.5) and (15.6) satisfy equation (15.24).

Exercise 302. Show that V^α transforms according to

$$V^{\alpha'} = V^\alpha J_\alpha^{\alpha'} + J_\alpha^{\alpha'} J_t^\alpha. \tag{15.25}$$

Exercise 303. Show that coordinate velocities V^α and $V^{\alpha'}$ associated with parametrizations (15.5) and (15.6) satisfy equation (15.25).

15.4 The Velocity C of an Interface

In the preceding section, we established that the coordinate velocity V^i in the a tensor on the account of the term $Z_\alpha^i J_i^{i'} J_t^\alpha$ in equation (15.24). This term contains Z_α^i and is therefore orthogonal to the normal N_i (more accurately: $N_{i'}$). To take advantage of this, multiply both sides by $N_{i'}$

$$V^{i'} N_{i'} = V^i J_i^{i'} N_{i'} + Z_\alpha^i J_i^{i'} N_{i'} J_t^\alpha. \tag{15.26}$$

Since $N_{i'} J_i^{i'} = N_i$, the first term on the right-hand side equals $N_i V^i$, while the second term vanishes. Thus,

$$V^{i'} N_{i'} = V^i N_i. \tag{15.27}$$

In other words, the normal component $V^i N_i$ of the coordinate velocity is invariant. This combination is called the *velocity of the surface* and is denoted by the letter C

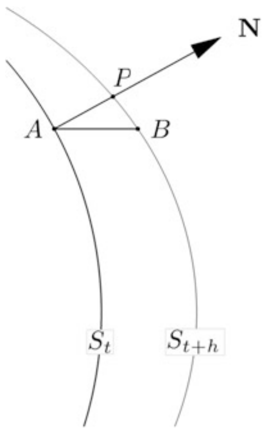


Fig. 15.3 Geometric construction of the surface velocity C

$$C = V^i N_i. \tag{15.28}$$

Note that the sign of C depends on the choice of the normal.

The velocity C has a clear geometric interpretation illustrated in Fig. 15.3. Denote the surfaces at two nearby moments of time t and $t + h$ by S_t and S_{t+h}. Suppose that A is a point on S_t and B is a point on S_t that has the same surface coordinates as A. Thus, the vector from A to B is approximately $\mathbf{V}h$. Let P be the point where the unit normal $\mathbf{N}$ of S_t intersects S_{t+h}. Since h is small, the angle APB is nearly $\pi/2$. The length of AP is approximately

$$|AP| = \mathbf{V} \cdot \mathbf{N}h = Ch. \tag{15.29}$$

Therefore, C can be defined as the limit

$$C = \lim_{h \to 0} \frac{|AP|}{h} \tag{15.30}$$

and is interpreted as the instantaneous velocity of the interface in the normal direction.

We have noted previously that the sign of C depends on the choice of the normal. This property of C is also evident in its geometric interpretation: the length $|AP|$ is considered positive if the vector $\overrightarrow{AP}$ points in the same direction as $\mathbf{N}$ and negative otherwise. This convention can be explicitly incorporated into the geometric definition as follows

$$C = \lim_{h \to 0} \frac{\overrightarrow{AP} \cdot \mathbf{N}}{h}. \tag{15.31}$$

The geometric approach makes no reference to coordinate systems, and it is therefore clear that C is an invariant quantity. Furthermore, the geometric construction of C is consistent with the analytical definition, since for "small enough" h, AP is "nearly" an orthogonal projection of AB. There is no doubt that this argument can be turned into a rigorous proof of the equivalence. We have not presented a rigorous demonstration of the equivalence and therefore invite the reader to view equation (15.28) as the definition of C and equation (15.31) as an intuitive geometric insight.

Although the quantity C is a scalar, it is called a *velocity* because the normal direction is implied. One may define the *vector normal velocity* **C**:

$$\mathbf{C} = C\mathbf{N}. \tag{15.32}$$

The vector velocity **C** is convenient in several contexts, particularly equation (15.46). However, all the information regarding the motion of the surface is contained in the scalar quantity C.

Exercise 304. Compute C for $A = 1$ at $t = 1$ for the evolution (15.5).

Exercise 305. Compute C for $A = 1$ at $t = 1$ for the evolution (15.6).

15.5 The Invariant Time Derivative $\dot{\nabla}$

The covariant derivative ∇_i was born out of the need to preserve the tensor property of objects, since the partial derivative $\partial/\partial Z^i$ did not have this property. The covariant surface derivative ∇_α had similar motivations. In the study of moving surfaces, the partial derivative $\partial/\partial t$ does not preserve the tensor property, and therefore a new operator is needed.

When we defined the invariant surface velocity C, our initial approach was analytical and we later backed it up with geometric intuition. In the case of the invariant derivative operator $\dot{\nabla}$, we start with the geometric construction, which is analogous to that of C. It is illustrated in Fig. 15.4.

Suppose that an invariant field T is defined on the surface at all times t. The field T could be the normal **N**, the mean curvature B^α_α, a surface restriction of an invariant field defined in the ambient space, or anything else. The idea of the derivative operator $\dot{\nabla}$ is to capture the rate of change in T in the normal direction. This is similar to how C measures the rate of deformation in the normal direction. Given a point A on S_t, find the points B and P as we did previously: B lies on S_{t+h} and has the same coordinates as A, and P is the intersection of S_{t+h} and the straight line orthogonal to S_h. The geometric definition of $\dot{\nabla}T$ involves only points A and P and corresponds to our intuition of instantaneous rate of change in the normal direction:

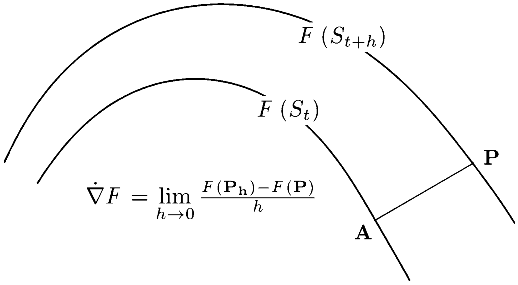


Fig. 15.4 Geometric construction of the invariant time derivative $\dot{\nabla}$ as applied to invariant quantities

$$\dot{\nabla} T = \lim_{h\to 0} \frac{T(P) - T(A)}{h}. \tag{15.33}$$

The definition (15.33) is entirely geometric and therefore $\dot{\nabla} T$ is an invariant. Our next goal is to find an analytical expression for $\dot{\nabla} T$ in a particular coordinate system. By virtue of the geometric construction, the resulting expression is invariant. The desired expression can be determined by analyzing the value of T at point B. Since B has the same surface coordinates as A, we have

$$T(B) \approx T(A) + h\frac{\partial T(t,S)}{\partial t}. \tag{15.34}$$

On the other hand, $T(B)$ can be related to $T(P)$ because B and P are nearby points on the surface S_{t+h}. The relation reads

$$T(B) \approx T(P) + hV^\alpha \nabla_\alpha T, \tag{15.35}$$

since $\nabla_\alpha T$ captures the rates of change in T along the surface and hV^α captures the directed distance from B to P. Eliminating $T(B)$ from equations (15.34) and (15.35), we find

$$T(P) - T(A) \approx h\left(\frac{\partial T}{\partial t} - V^\alpha \nabla_\alpha T\right). \tag{15.36}$$

Motivated by equation (15.36) we give the following definition of the invariant time derivative $\dot{\nabla}$ when applied to invariants:

$$\dot{\nabla} T = \frac{\partial T(t,S)}{\partial t} - V^\alpha \nabla_\alpha T. \tag{15.37}$$

As before, we have not drawn a rigorous connection between the definitions (15.33) and (15.37). Therefore, we consider equation (15.37) as the primary definition and equation (15.33) as the geometric intuition behind the analytical definition.

Exercise 306. With the help of equations (15.25) and (15.14), show analytically that $\dot{\nabla} T$ is an invariant.

Exercise 307. Show that $\dot{\nabla}$ satisfies the sum rule

$$\dot{\nabla}(F+G) = \dot{\nabla}F + \dot{\nabla}G. \tag{15.38}$$

Exercise 308. Show that $\dot{\nabla}$ satisfies the product rule

$$\dot{\nabla}(FG) = \dot{\nabla}F\, G + F\, \dot{\nabla}G. \tag{15.39}$$

Exercise 309. Show that $\dot{\nabla}$ satisfies the product rule with respect to the dot

$$\dot{\nabla}(\mathbf{F}\cdot\mathbf{G}) = \dot{\nabla}\mathbf{F}\cdot\mathbf{G} + \mathbf{F}\cdot\dot{\nabla}\mathbf{G}. \tag{15.40}$$

Exercise 310. Show that $\dot{\nabla}$ applied to a constant field $T(t,S) \equiv A$ vanishes.

Definition (15.37) applies only variants of order zero. In Chap. 16, we expand $\dot{\nabla}$ to variants of arbitrary order. So far, we are able to apply it to the position vector $\mathbf{R}$, the normal $\mathbf{N}$ and the mean curvature B^α_α. We would now like to present the results of applying $\dot{\nabla}$ to these objects. However, in the case of the normal $\mathbf{N}$ and the mean curvature B^α_α the derivation is postponed until Chap. 16.

By definition, $\dot{\nabla}\mathbf{R}$ is given by

$$\dot{\nabla}\mathbf{R} = \frac{\partial \mathbf{R}(t,S)}{\partial t} - V^\alpha \nabla_\alpha \mathbf{R}. \tag{15.41}$$

The first term on the right-hand side can be expanded by the chain rule

$$\frac{\partial \mathbf{R}(t,S)}{\partial t} = \frac{\partial \mathbf{R}(Z(t,S))}{\partial t} = \frac{\partial \mathbf{R}}{\partial Z^i}\frac{\partial Z^i(t,S)}{\partial t} = V^i \mathbf{Z}_i, \tag{15.42}$$

as we saw previously. Thus,

$$\dot{\nabla}\mathbf{R} = V^i \mathbf{Z}_i - V^\alpha \mathbf{S}_\alpha, \tag{15.43}$$

or

$$\dot{\nabla}\mathbf{R} = \left(V^i - V^\alpha Z^i_\alpha\right)\mathbf{Z}_i. \tag{15.44}$$

Finally, combining equations (15.19) and (10.55), we find

$$\dot{\nabla}\mathbf{R} = V^j N_j N^i \mathbf{Z}_i, \tag{15.45}$$

or

$$\dot{\nabla}\mathbf{R} = C\mathbf{N} = \mathbf{C}. \tag{15.46}$$

This is the first equation from the fundamental differentiation table of the calculus of moving surfaces.

Exercise 311. Derive equation (15.46) by the geometric definition that (15.33).

Exercise 312. Fill in the details of the transition from equation (15.43) to (15.44).

Exercise 313. Fill in the details of the transition from equation (15.44) to (15.45).

As was mentioned earlier, we postpone the derivation of the formulas for $\dot{\nabla}\mathbf{N}$ and $\dot{\nabla} B^{\alpha}_{\alpha}$ until the following chapter. However, one relationship involving $\dot{\nabla}\mathbf{N}$ is easy to determine. Differentiating the identity $\mathbf{N} \cdot \mathbf{N} = 1$, we find by the product rule that

$$\mathbf{N} \cdot \dot{\nabla}\mathbf{N} = 0. \tag{15.47}$$

This relationship tells us that $\dot{\nabla}\mathbf{N}$ is orthogonal to the normal space. It must therefore lie in the tangent space and be represented by a linear combination of the covariant basis vectors $\mathbf{S}_{\alpha}$. The coefficients of the linear combination turn out to be $-\nabla^{\alpha} C$. The result is known as the *Thomas formula*

$$\dot{\nabla}\mathbf{N} = -\mathbf{S}_{\alpha}\nabla^{\alpha} C. \tag{15.48}$$

This formula first appeared in Tracy Thomas's book [46]. Finally, we mention that the derivative $\dot{\nabla} B^{\alpha}_{\alpha}$ is given by the following formula, perhaps the most beautiful among all calculus of moving surfaces relationships:

$$\dot{\nabla} B^{\alpha}_{\alpha} = \nabla_{\alpha}\nabla^{\alpha} C + C B^{\alpha}_{\beta} B^{\beta}_{\alpha}. \tag{15.49}$$

This formula can be found in [16].

15.6 The Chain Rule

The rule discussed in this section is analogous to the chain rule for the covariant surface derivative ∇_{α} discussed in Sect. 11.8. Suppose that a field T is defined in the ambient space. Its surface restriction to the surface S is subject to the invariant derivative $\dot{\nabla}$. By definition,

$$\dot{\nabla} T = \frac{\partial T\,(t, S)}{\partial t} - V^{\alpha}\nabla_{\alpha} T. \tag{15.50}$$

If T is a surface restriction of an ambient field $T(t,Z)$, then $T(t,S)$ may be expressed as $T(t,Z(t,S))$. By the chain rule of multivariable calculus and tensor calculus on stationary manifolds, we find:

$$\dot{\nabla} T = \frac{\partial T(t,Z)}{\partial t} + \frac{\partial T}{\partial Z^i}\frac{\partial Z^i(t,S)}{\partial t} - V^\alpha Z^i_\alpha \nabla_i T. \tag{15.51}$$

Since $\partial T/\partial Z^i = \nabla_i T$ and $\partial Z^i(t,S)/\partial t = V^i$,

$$\dot{\nabla} T = \left(V^i - V^\alpha Z^i_\alpha\right)\nabla_i T. \tag{15.52}$$

As we showed above while deriving $\dot{\nabla}\mathbf{R}$, the quantity in parentheses equals CN^i. We have thus arrived at the *chain rule* of the calculus of moving surfaces:

$$\dot{\nabla} T = \frac{\partial T(t,Z)}{\partial t} + CN^i \nabla_i T. \tag{15.53}$$

This chain rule has a clear geometric interpretation illustrated in Fig. 15.3. It states that the difference in the quantity T between the points A and P is due to two influences: the difference of $h\ \partial T/\partial t$ in the value of T due to the passage of time and the difference of $hCN^i\nabla_i T$ between the values of T at the points A and P at time t.

15.7 Time Evolution of Integrals

The remarkable usefulness of the calculus of moving surfaces becomes evident from two upcoming formulas (15.54) and (15.56) that govern the rates of change of volume and surface integrals due to the deformation of the domain. Many critical quantities in physics and engineering are represented by integrals. These quantities include mass, energy, entropy, electrical current, and so forth. In applications one often needs to find the rate of change in these quantities induced by changes in the domain. In fact, most problems in the calculus of moving surfaces—whether in shape optimization, boundary variations or physical applications—begin with an integral expression. Thus, equations (15.54) and (15.56) are indispensable tools in the analysis of moving surfaces.

Consider a Euclidean domain Ω with boundary S evolving with surface velocity C. Figure 15.5 illustrates this discussion. Suppose that the invariant field, vector, or scalar, $F(t,Z)$ depends on time and is sufficiently smooth within Ω. Then the rate of change of the volume integral $\int_\Omega F d\Omega$ is given by the formula

$$\frac{d}{dt}\int_\Omega F d\Omega = \int_\Omega \frac{\partial F}{\partial t} d\Omega + \int_S CF dS. \tag{15.54}$$

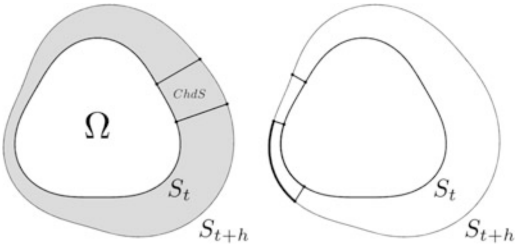


Fig. 15.5 Illustrations of the integral laws (15.54) and (15.56). The figure on the left shows the annexed region that has a thickness proportional to C. The thicker portion of the boundary of the figure on the right illustrates that the change in area is proportional to the mean curvature B_α^α

The volume term on the right captures the rate of change of F and the surface term captures the rate at which the moving boundary annexes or gives up the neighboring regions. The annexing term is illustrated on the left of Fig. 15.5.

We give this law without a general proof, but I believe that this law is entirely intuitive. Notice the similarities between this equation and the fundamental theorem of calculus in the case when the integrand and integration domain depend on the parameter t:

$$\frac{d}{dt}\int_a^{b(t)} F(t,x)\,dx = \int_a^{b(t)} \frac{\partial F(t,x)}{\partial t}dx + b'(t)\,F(t,b(t)). \tag{15.55}$$

In fact, as an exercise below shows, a special case of equation (15.54) follows easily from the fundamental theorem of calculus.

The other fundamental formula addresses the surface integral $\int_S FdS$ for an invariant, vector or scalar, defined on the *closed* surface. The rate of change of this integral is governed by the formula

$$\frac{d}{dt}\int_S FdS = \int_S \dot{\nabla} F\,dS - \int_S CB_\alpha^\alpha F\,dS. \tag{15.56}$$

This formula was first given in [16]. The first term once again corresponds to the rate of change in F. Notably, the operator that figures in this term is the invariant time derivative $\dot{\nabla}$! It could hardly have been any other operator: the left-hand side of equation (15.56) is coordinate-free, and therefore the result ought to consist only of invariant elements.

In Chap. 14, we proved that the surface integral of the unit normal over a closed surface vanishes

$$\int_S \mathbf{N}dS = \mathbf{0}. \tag{15.57}$$

As an example of using equation (15.56), we now show a special case of this equation by means of the calculus of moving surfaces. Namely, we show that the rate of change of this integral is zero

$$\frac{d}{dt}\int_S \mathbf{N}dS = \mathbf{0}. \tag{15.58}$$

By equation (15.56), we have

$$\frac{d}{dt}\int_S \mathbf{N}dS = \int_S \dot{\nabla}\mathbf{N}dS - \int_S CB_\alpha^\alpha \mathbf{N}dS. \tag{15.59}$$

According to the Thomas formula (15.48)

$$\frac{d}{dt}\int_S \mathbf{N}dS = -\int_S \mathbf{S}_\alpha \nabla^\alpha C dS - \int_S CB_\alpha^\alpha \mathbf{N}dS. \tag{15.60}$$

By Gauss's theorem,

$$\frac{d}{dt}\int_S \mathbf{N}dS = \int_S \nabla^\alpha \mathbf{S}_\alpha \, C \, dS - \int_S CB_\alpha^\alpha \mathbf{N}dS. \tag{15.61}$$

Because $\nabla^\alpha \mathbf{S}_\alpha = \mathbf{N}B_\alpha^\alpha$, the right-hand side vanishes and equation (15.58) is confirmed. Since the integral of the unit normal over a sphere vanishes, we have actually shown that it vanishes over all smooth surfaces of genus zero—that is, all closed surfaces that can be obtained by a smooth deformation of a sphere.

Exercise 314. Show that

$$\frac{d}{dt}\int_S \mathbf{N}B_\alpha^\alpha dS = \mathbf{0}.$$

You will enjoy this exercise as it draws upon virtually all of the surface identities, including the Codazzi equation.

Exercise 315. Conclude that the integral

$$\int_S \mathbf{N}B_\alpha^\alpha dS \tag{15.62}$$

vanishes over all closed smooth surfaces of genus zero.

Exercise 316. This exercise gives the proof of a special case of the volume law (15.54). Suppose that, with a Cartesian coordinate system imposed, the planar domain Ω is the region between two fixed values of x in the horizontal direction, and between the x-axis and a positive time-dependent function $b(t,x)$ in the vertical direction. Consider a time-independent function $F(x,y)$ and its area integral

$$I = \int_{\Omega} F d\Omega = \int_{A_1}^{A_2} \int_{0}^{b(t,x)} F(x, y)\, dy dx. \tag{15.63}$$

Use the fundamental theorem of calculus to show dI/dt is given by formula (15.54).

15.8 A Need for Further Development

The definition (15.37) gives us an operator that produces invariants when applied to invariants. However, for tensors of order greater than zero, the result of applying $\dot{\nabla}$ is not a tensor. The following exercises show this. Therefore, more work remains to be done. The important task of constructing such an operator is undertaken in the next chapter.

Exercise 317. For a tensor T^i, derive the following transformation rule for the variant $U^i = \partial T^i (t, S) / \partial t$:

$$U^{i'} = U^i J_i^{i'} + \frac{\partial T^i}{\partial S^\alpha} J_t^\alpha J_i^{i'} + T^i J_{ij}^{i'} V^j + T^i J_{ij}^{i'} Z_\alpha^j J_t^\alpha. \tag{15.64}$$

Exercise 318. For a tensor T_i, derive the following transformation rule for the variant $U_i = \partial T_i (t, S) / \partial t$:

$$U_{i'} = U_i J_{i'}^i + \frac{\partial T_i}{\partial S^\alpha} J_t^\alpha J_{i'}^i + T_i J_{i'j'}^i V^{j'}. \tag{15.65}$$

Exercise 319. For a tensor T^α, derive the following transformation rule for the variant $U^\alpha = \partial T^\alpha (t, S) / \partial t$:

$$U^{\alpha'} = U^\alpha J_\alpha^{\alpha'} + \frac{\partial T^\alpha}{\partial S^\beta} J_t^\beta J_\alpha^{\alpha'} + T^\alpha J_{\alpha t}^{\alpha'} + T^\alpha J_{\alpha\beta}^{\alpha'} J_t^\beta. \tag{15.66}$$

Exercise 320. For a tensor $T_{\alpha'}$, derive the following transformation rule for the variant $U_{\alpha'} = \partial T_{\alpha'} (t, S) / \partial t$:

$$U_{\alpha'} = U_\alpha J_{\alpha'}^\alpha + \frac{\partial T_\alpha}{\partial S^\beta} J_t^\beta J_{\alpha'}^\alpha + T_\alpha J_{\alpha' t}^\alpha. \tag{15.67}$$

Exercise 321. For a tensor T^i, derive the following transformation rule for the variant $W^i = \partial T^i (t, S) / \partial t - V^\alpha \nabla_\alpha T^i$:

$$W^{i'} = W^i J_i^{i'} - \Gamma_{ij}^k T^i Z_\alpha^j J_t^\alpha J_k^{i'} + T^i J_{ij}^{i'} V^j + T^i J_{ij}^{i'} Z_\alpha^j J_t^\alpha. \tag{15.68}$$

Exercise 322. For a tensor T_i, derive the following transformation rule for the variant $W_i = \partial T_i (t, S) / \partial t - V^{\alpha} \nabla_{\alpha} T_i$:

$$W_{i'} = W_i J^i_{i'} + \Gamma^k_{ji} Z^j_{\alpha} T_k J^{\alpha}_t J^i_{i'} + T_i J^i_{i'j'} V^{j'} - J^{\alpha'}_{\alpha} J^{\alpha}_t \nabla_{\alpha'} T_{i'}. \tag{15.69}$$

Exercise 323. For a tensor T^{α}, derive the following transformation rule for the variant $W^{\alpha} = \partial T^{\alpha} (t, S) / \partial t - V^{\beta} \nabla_{\beta} T^{\alpha}$:

$$W^{\alpha'} = W^{\alpha} J^{\alpha'}_{\alpha} - \Gamma^{\alpha}_{\beta\gamma} T^{\gamma} J^{\beta}_t J^{\alpha'}_{\alpha} + T^{\alpha} J^{\alpha'}_{\alpha t} + T^{\alpha} J^{\alpha'}_{\alpha\beta} J^{\beta}_t. \tag{15.70}$$

Exercise 324. For a tensor $T_{\alpha'}$, derive the following transformation rule for the variant $W_{\alpha'} = \partial T_{\alpha'} (t, S) / \partial t - V^{\beta'} \nabla_{\beta'} T_{\alpha'}$:

$$W_{\alpha'} = W_{\alpha} J^{\alpha}_{\alpha'} + \Gamma^{\gamma}_{\alpha\beta} T_{\gamma} J^{\beta}_t J^{\alpha}_{\alpha'} + T_{\alpha} J^{\alpha}_{\alpha' t}. \tag{15.71}$$

15.9 Summary

In this chapter, we introduced the fundamental elements of the calculus of moving surfaces. We defined the velocity of the surface C and a new time derivative operator $\dot{\nabla}$ that preserves the invariant property. The two fundamental formulas (15.54) and (15.56) for differentiating integrals over deforming domains showed the importance of developing the techniques of the calculus of moving surfaces.

The elements of the calculus of moving surfaces presented in this chapter are only the beginning. The technique is not fully developed until the time derivative $\dot{\nabla}$ is extended to variants of arbitrary indicial signature. This is the central task of the next chapter.

Chapter 16
Extension to Arbitrary Tensors

16.1 Preview

In Chap. 15, the invariant derivative operator $\dot{\nabla}$ was defined for variants of order zero. The new operator proved to have a number of essential features including the tensor property—that is, producing tensor outputs for tensor inputs. Even with that narrow definition, the new operator demonstrated its impressive utility in equations (15.54) and (15.56) for evaluating the rates of change of volume and surface integrals. However, analysis of all but a few problems is impossible, unless the new derivative is extended to all objects encountered on surfaces which includes variants with arbitrary indicial signatures. The development of this extension is the subject of this chapter.

16.2 The Extension to Ambient Indices

The extension of $\dot{\nabla}$ to variants with ambient indices can be induced if $\dot{\nabla}$ is to satisfy the following properties, the sum and product rules, commutativity with contraction, and the metrinilic property with respect to the ambient basis $\mathbf{Z}_i$.

Consider the variant $\mathbf{T}$ of order zero

$$\mathbf{T} = T^i \mathbf{Z}_i . \tag{16.1}$$

If the desired properties hold, then

$$\dot{\nabla}\mathbf{T} = \dot{\nabla}T^i \ \mathbf{Z}_i . \tag{16.2}$$

Exercise 325. Explain how each of the three desired properties contribute to the validity of equation (16.2).

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6_16,

On the left-hand side, the invariant derivative $\dot{\nabla}$ is applied to a variant of order zero and is therefore given by

$$\dot{\nabla}\mathbf{T} = \frac{\partial \mathbf{T}}{\partial t} - V^{\alpha}\nabla_{\alpha}\mathbf{T}. \tag{16.3}$$

Substitute equation (16.1) and apply the product rule:

$$\dot{\nabla}\mathbf{T} = \frac{\partial T^i}{\partial t}\mathbf{Z}_i + T^i\frac{\partial \mathbf{Z}_i}{\partial t} - V^{\gamma}\nabla_{\gamma}T^i\,\mathbf{Z}_i. \tag{16.4}$$

Exercise 326. Show that

$$\frac{\partial \mathbf{Z}_i}{\partial t} = \Gamma^k_{ij}V^j\mathbf{Z}_k. \tag{16.5}$$

Therefore,

$$\dot{\nabla}\mathbf{T} = \frac{\partial T^i}{\partial t}\mathbf{Z}_i + T^k\Gamma^i_{kj}V^j\mathbf{Z}_i - V^{\alpha}\nabla_{\alpha}T^i\mathbf{Z}_i. \tag{16.6}$$

Equating this expression to the right-hand side of equation (16.2), we arrive at a reasonable definition of $\dot{\nabla}$ for a variant T^i with a contravariant ambient index:

$$\dot{\nabla}T^i = \frac{\partial T^i}{\partial t} - V^{\gamma}\nabla_{\gamma}T^i + V^j\Gamma^i_{jk}T^k. \tag{16.7}$$

Exercise 327. Show that similar reasoning leads to the following definition for a tensor with a covariant ambient index

$$\dot{\nabla}T_i = \frac{\partial T_i}{\partial t} - V^{\gamma}\nabla_{\gamma}T_i - V^j\Gamma^k_{ji}T_k. \tag{16.8}$$

Motivated by equations (16.7) and (16.8) we give the following definition for a tensor T^i_j with a representative collection of ambient indices:

$$\dot{\nabla}T^i_j = \frac{\partial T^i_j}{\partial t} - V^{\gamma}\nabla_{\gamma}T^i_j + V^k\Gamma^i_{km}T^m_j - V^k\Gamma^m_{kj}T^i_m. \tag{16.9}$$

Algorithmically, equation (16.9) is interpreted similarly to other definition of tensor differential operators. The term $-V^{\gamma}\nabla_{\gamma}$ appears only once, and to each index there corresponds a term containing the Christoffel symbols. The new extended operator retains all of its properties, and it is left as exercises to demonstrate that it is so.

Exercise 328. Show that $\dot{\nabla}T^i$ is a tensor for a tensor argument T^i. You should show this in two ways: by deriving the rule by which $\dot{\nabla}T^i$ transforms under a change of variables and by analyzing equation (16.3).

Exercise 329. Similarly, show that $\dot{\nabla} T_i$ is a tensor for a tensor argument T_i.

Exercise 330. Show that $\dot{\nabla} T^i_j$ is a tensor for a tensor operand T^i_j. Once again, there are two ways of showing this: by deriving the rule by which $\dot{\nabla} T^i_j$ transforms under a change of variables, and by leveraging the corresponding property for $\dot{\nabla} T^i$ or $\dot{\nabla} T_i$.

Exercise 331. Conclude that $\dot{\nabla}$ possesses the tensor property for arbitrary tensors with ambient indices.

Exercise 332. Show that $\dot{\nabla}$ commutes with contraction.

Exercise 333. Show that $\dot{\nabla}$ satisfies the sum and product rules.

Exercise 334. Show that if $T^i_j(t, Z)$ is defined in the entire ambient space, then the invariant derivative $\dot{\nabla} T^i_j$ applied to its surface restriction $T^i_j(t, S)$ satisfies the chain rule

$$\dot{\nabla} T^i_j = \frac{\partial T(t, Z)}{\partial t} + C N^k \nabla_k T^i_j. \tag{16.10}$$

Exercise 335. Conclude that the invariant time derivative is metrinilic with respect to the ambient metrics:

$$\dot{\nabla} \mathbf{Z}_i, \ \ \dot{\nabla} \mathbf{Z}^i = \mathbf{0} \tag{16.11}$$

$$\dot{\nabla} Z_{ij}, \ \ \dot{\nabla} \delta^i_j, \ \ \dot{\nabla} Z^{ij}, \ \ \dot{\nabla} \varepsilon_{ijk}, \ \ \dot{\nabla} \varepsilon^{ijk} = 0. \tag{16.12}$$

16.3 The Extension to Surface Indices

Historically, the extension of the invariant time derivative to ambient indices followed a long and winding path. Several incorrect attempts were made until a proper definition was finally given in [16], where it referred to as the $\delta/\delta t$-derivative. The operator $\dot{\nabla}$, as applied to variants with surface indices, was first defined in [20] where its advantages over the $\delta/\delta t$-derivative was described.

The construction of the operator $\dot{\nabla}$ for the ambient indices can proceed in a way similar to the ambient indices. We stipulate that $\dot{\nabla}$ satisfies the following properties: the sum and product rules, commutativity with contraction and, when applied to $\mathbf{S}_\alpha$, the result proportional to the normal $\mathbf{N}$:

$$\dot{\nabla} \mathbf{S}_\alpha = A_\alpha \mathbf{N}. \tag{16.13}$$

Consider the following variant $\mathbf{T}$ of order zero:

$$\mathbf{T} = T^\alpha \mathbf{S}_\alpha. \tag{16.14}$$

If the desired properties hold, then

$$\dot{\nabla}\mathbf{T} = \dot{\nabla}T^{\alpha}\,\mathbf{S}_{\alpha} + T^{\alpha}A_{\alpha}\mathbf{N}. \tag{16.15}$$

The left-hand side can be analyzed by the definition of $\dot{\nabla}$ as applied to invariants:

$$\dot{\nabla}\mathbf{T} = \frac{\partial\mathbf{T}}{\partial t} - V^{\beta}\nabla_{\beta}\mathbf{T}. \tag{16.16}$$

Substituting equation (16.14), we find

$$\dot{\nabla}\mathbf{T} = \frac{\partial\left(T^{\alpha}\mathbf{S}_{\alpha}\right)}{\partial t} - V^{\beta}\nabla_{\beta}\left(T^{\alpha}\mathbf{S}_{\alpha}\right), \tag{16.17}$$

which, by the product rule, equation (11.16), and renaming $\alpha \to \beta$ in the second term, yields

$$\dot{\nabla}\mathbf{T} = \frac{\partial T^{\alpha}}{\partial t}\mathbf{S}_{\alpha} + T^{\beta}\frac{\partial\mathbf{S}_{\beta}}{\partial t} - V^{\beta}\nabla_{\beta}T^{\alpha}\,\mathbf{S}_{\alpha} - V^{\beta}T^{\alpha}\mathbf{N}B_{\beta\alpha}. \tag{16.18}$$

Exercise 336. Show that

$$\frac{\partial\mathbf{S}_{\beta}}{\partial t} = \nabla_{\beta}\mathbf{V}. \tag{16.19}$$

Exercise 337. Thus,

$$\frac{\partial\mathbf{S}_{\beta}}{\partial t} = \nabla_{\beta}\left(V^{\alpha}\mathbf{S}_{\alpha} + C\mathbf{N}\right). \tag{16.20}$$

Exercise 338. Expand the right-hand side and show that

$$\frac{\partial\mathbf{S}_{\beta}}{\partial t} = \nabla_{\beta}V^{\alpha}\mathbf{S}_{\alpha} + V^{\alpha}\mathbf{N}B_{\beta\alpha} + \mathbf{N}\nabla_{\beta}C - CB^{\alpha}_{\beta}\mathbf{S}_{\alpha}. \tag{16.21}$$

Exercise 339. Show that substituting this result into equation (16.18), yields

$$\dot{\nabla}\mathbf{T} = \frac{\partial T^{\alpha}}{\partial t}\mathbf{S}_{\alpha} + T^{\beta}\nabla_{\beta}V^{\alpha}\mathbf{S}_{\alpha} + T^{\alpha}\nabla_{\alpha}C\mathbf{N} - T^{\beta}CB^{\alpha}_{\beta}\mathbf{S}_{\alpha} - V^{\beta}\nabla_{\beta}T^{\alpha}\,\mathbf{S}_{\alpha}. \tag{16.22}$$

Comparing this result to equation (16.17) yields the definition of $\dot{\nabla}$ for T^{α}

$$\dot{\nabla}T^{\alpha} = \frac{\partial T^{\alpha}}{\partial t} - V^{\beta}\nabla_{\beta}T^{\alpha} + \left(\nabla_{\beta}V^{\alpha} - CB^{\alpha}_{\beta}\right)T^{\beta}, \tag{16.23}$$

as well as the expected formula for $\dot{\nabla}\mathbf{S}_\alpha$:

$$\dot{\nabla}\mathbf{S}_\alpha = \mathbf{N}\nabla_\alpha C. \tag{16.24}$$

Exercise 340. Show that the same approach suggests the following definition for covariant arguments:

$$\dot{\nabla}T_\alpha = \frac{\partial T_\alpha}{\partial t} - V^\beta\nabla_\beta T_\alpha - \left(\nabla_\alpha V^\beta - CB_\alpha^\beta\right)T_\beta. \tag{16.25}$$

16.4 The General Invariant Derivative $\dot{\nabla}$

We are now ready to give the full definition of $\dot{\nabla}$ for arbitrary tensors. We combine equations (16.7), (16.8), (16.23), and (16.25) for each type of index and give the following definition for the operator $\dot{\nabla}$ for a variant $T_{j\beta}^{i\alpha}$ with a representative collection of indices. Define the Christoffel symbol $\dot{\Gamma}_\beta^\alpha$ for moving surfaces as

$$\dot{\Gamma}_\beta^\alpha = \nabla_\beta V^\alpha - CB_\beta^\alpha. \tag{16.26}$$

Then the definition for the operator $\dot{\nabla}$ as applied to $T_{j\beta}^{i\alpha}$ is:

$$\dot{\nabla}T_{j\beta}^{i\alpha} = \frac{\partial T_{j\beta}^{i\alpha}}{\partial t} - V^\gamma\nabla_\gamma T_{j\beta}^{i\alpha} + V^k\Gamma_{km}^i T_{j\beta}^{m\alpha} - V^k\Gamma_{kj}^m T_{m\beta}^{i\alpha} + \dot{\Gamma}_\omega^\alpha T_{j\beta}^{i\omega} - \dot{\Gamma}_\beta^\omega T_{j\omega}^{i\alpha}. \tag{16.27}$$

The proofs of these properties of $\dot{\nabla}$ are left as exercises.

1. The operator $\dot{\nabla}$ produces tensor outputs for tensor inputs
2. The operator $\dot{\nabla}$ commutes with contraction
3. The operator $\dot{\nabla}$ satisfies the sum and product rules
4. The operator $\dot{\nabla}$ satisfies the chain rule. That is, if $T_j^i(t,Z)$ is defined in the entire ambient space, then $\dot{\nabla}$ applied to its surface restriction $T_j^i(t,S)$ satisfies the chain rule

$$\dot{\nabla}T_j^i = \frac{\partial T(t,Z)}{\partial t} + CN^k\nabla_k T_j^i. \tag{16.28}$$

5. The operator $\dot{\nabla}$ is metrinilic with respect to the ambient metrics

$$\dot{\nabla}\mathbf{Z}_i,\quad \dot{\nabla}\mathbf{Z}^i = \mathbf{0} \tag{16.29}$$

$$\dot{\nabla}Z_{ij},\quad \dot{\nabla}Z^{ij},\quad \dot{\nabla}\delta_j^i,\quad \dot{\nabla}\varepsilon_{ijk},\quad \dot{\nabla}\varepsilon^{ijk} = 0. \tag{16.30}$$

6. The invariant derivative $\dot{\nabla}$ does not commute with the surface derivative ∇_α. The general rule for the commutator $\dot{\nabla}\nabla_\alpha - \nabla_\alpha\dot{\nabla}$ depends on the indicial signature of the variant to which it is applied. For variants T of order zero, the rule reads

$$\left(\dot{\nabla}\nabla_\alpha - \nabla_\alpha\dot{\nabla}\right) T = -CB_\alpha^\omega \nabla_\omega T. \tag{16.31}$$

Next, we evaluate the invariant time derivative for the fundamental differential objects on moving surfaces.

16.5 The Formula for $\dot{\nabla}\mathbf{S}_\alpha$

By the definition (16.27), we have

$$\dot{\nabla}\mathbf{S}_\alpha = \frac{\partial \mathbf{S}_\alpha}{\partial t} - V^\omega \nabla_\omega \mathbf{S}_\alpha - \left(\nabla_\alpha V^\omega - CB_\alpha^\omega\right) \mathbf{S}_\omega. \tag{16.32}$$

By equation (11.16), this becomes

$$\dot{\nabla}\mathbf{S}_\alpha = \nabla_\alpha \mathbf{V} - V^\omega \mathbf{N} B_{\omega\alpha} - \left(\nabla_\alpha V^\omega - CB_\alpha^\omega\right) \mathbf{S}_\omega. \tag{16.33}$$

Decompose $\mathbf{V}$ in tangential and normal components:

$$\dot{\nabla}\mathbf{S}_\alpha = \nabla_\alpha \left(V^\omega \mathbf{S}_\omega + C\mathbf{N}\right) - V^\gamma \mathbf{N} B_{\gamma\alpha} - \left(\nabla_\alpha V^\omega - CB_\alpha^\omega\right) \mathbf{S}_\omega. \tag{16.34}$$

Apply the product rule to the first term. Several terms cancel and we find the final formula

$$\dot{\nabla}\mathbf{S}_\alpha = \mathbf{N}\nabla_\alpha C. \tag{16.35}$$

The metrinilic property of $\dot{\nabla}$ with respect to $\mathbf{Z}_i$ yields the following component form of equation (16.35)

$$\nabla Z_\alpha^i = N^i \nabla_\alpha C. \tag{16.36}$$

This completes perhaps the most challenging derivation of this chapter. The rest of the fundamental differential relationships follow relatively easily from equations (16.35) and (16.36).

16.6 The Metrinilic Property of $\dot{\nabla}$

The metrinilic property of the invariant derivative $\dot{\nabla}$ is similar to that of the covariant surface derivative ∇_α. That is, the metrinilic property holds when applied to ambient and surface metrics with the exception of $\mathbf{S}_\alpha$ and $\mathbf{S}^\alpha$. The metrinilic property with respect to the ambient metrics follows from the chain rule and has been established above. In this section, we focus on the metrinilic property with respect to the surface metrics.

For the covariant basis $S_{\alpha\beta}$, we have

$$\dot{\nabla} S_{\alpha\beta} = \dot{\nabla}\left(\mathbf{S}_\alpha \cdot \mathbf{S}_\beta\right) = \dot{\nabla}\mathbf{S}_\alpha \cdot \mathbf{S}_\beta + \mathbf{S}_\alpha \cdot \dot{\nabla}\mathbf{S}_\beta. \tag{16.37}$$

By equation (16.35), we find

$$\dot{\nabla} S_{\alpha\beta} = \mathbf{N}\nabla_\alpha C \cdot \mathbf{S}_\beta + \mathbf{S}_\alpha \cdot \mathbf{N}\nabla_\beta C. \tag{16.38}$$

Therefore, the result is zero by orthogonality of $\mathbf{N}$ and $\mathbf{S}_\alpha$:

$$\dot{\nabla} S_{\alpha\beta} = 0. \tag{16.39}$$

To prove that $\dot{\nabla}\delta^\alpha_\beta$ vanishes, we appeal to the definition (16.27):

$$\dot{\nabla}\delta^\alpha_\beta = \frac{\partial \delta^\alpha_\beta}{\partial t} - V^\omega \nabla_\omega \delta^\alpha_\beta + \Gamma^\alpha_\omega \delta^\omega_\beta - \Gamma^\omega_\beta \delta^\alpha_\omega. \tag{16.40}$$

Each of the first two terms vanish, while the third and the fourth terms cancel each other. Therefore

$$\dot{\nabla}\delta^\alpha_\beta = 0. \tag{16.41}$$

To show that $\dot{\nabla} S^{\alpha\beta}$ vanishes, apply $\dot{\nabla}$ to the identity $S^{\alpha\beta} S_{\beta\gamma} = \delta^\alpha_\gamma$. By equation (16.41)

$$\dot{\nabla}\left(S^{\alpha\beta} S_{\beta\gamma}\right) = 0. \tag{16.42}$$

By the product rule,

$$\dot{\nabla} S^{\alpha\beta}\ S_{\beta\gamma} + S^{\alpha\beta}\ \dot{\nabla} S_{\beta\gamma} = 0, \tag{16.43}$$

and by equation (16.39),

$$\dot{\nabla} S^{\alpha\beta}\ S_{\beta\gamma} = 0 \tag{16.44}$$

which immediately yields

$$\dot{\nabla} S^{\alpha\beta} = 0. \tag{16.45}$$

The metrinilic property of the $\dot{\nabla}$ with respect to the surface bases $S_{\alpha\beta}$ and $S^{\alpha\beta}$ allows the juggling of indices under $\dot{\nabla}$. For example, by raising the covariant index in equation (16.35), we find

$$\dot{\nabla} \mathbf{S}^{\alpha} = \mathbf{N} \nabla^{\alpha} C. \tag{16.46}$$

Similarly, by raising the covariant index in equation (16.36), we find

$$\nabla Z^{i\alpha} = N^{i} \nabla^{\alpha} C. \tag{16.47}$$

Exercise 341. Show from equation (16.39) that

$$\frac{\partial S_{\alpha\beta}}{\partial t} = \nabla_{\alpha} V_{\beta} + \nabla_{\beta} V_{\alpha} - 2CB_{\alpha\beta} \tag{16.48}$$

and from equation (16.45) that

$$\frac{\partial S^{\alpha\beta}}{\partial t} = -\nabla^{\alpha} V^{\beta} - \nabla^{\beta} V^{\alpha} + 2CB^{\alpha\beta}. \tag{16.49}$$

Exercise 342. Show that

$$\frac{\partial S}{\partial t} = 2S \left(\nabla_{\alpha} V^{\alpha} - CB_{\alpha}^{\alpha} \right) \tag{16.50}$$

and therefore

$$\frac{\partial \sqrt{S}}{\partial t} = \sqrt{S} \left(\nabla_{\alpha} V^{\alpha} - CB_{\alpha}^{\alpha} \right). \tag{16.51}$$

and

$$\frac{\partial \left(1/\sqrt{S} \right)}{\partial t} = -\frac{1}{\sqrt{S}} \left(\nabla_{\alpha} V^{\alpha} - CB_{\alpha}^{\alpha} \right). \tag{16.52}$$

Exercise 343. Show that

$$\frac{\partial \varepsilon_{\alpha\beta}}{\partial t} = \varepsilon_{\alpha\beta} \left(\nabla_{\gamma} V^{\gamma} - CB_{\gamma}^{\gamma} \right). \tag{16.53}$$

Exercise 344. Show that

$$\frac{\partial \varepsilon^{\alpha\beta}}{\partial t} = -\varepsilon^{\alpha\beta}\left(\nabla_\gamma V^\gamma - CB_\gamma^\gamma\right). \tag{16.54}$$

The two preceding exercises enable us to consider the Levi-Civita symbols $\varepsilon_{\alpha\beta}$ and $\varepsilon^{\alpha\beta}$. By definition (16.27),

$$\dot{\nabla}\varepsilon_{\alpha\beta} = \frac{\partial \varepsilon_{\alpha\beta}}{\partial t} - V^\gamma \nabla_\gamma \varepsilon_{\alpha\beta} - \dot{\Gamma}_\alpha^\omega \varepsilon_{\omega\beta} - \dot{\Gamma}_\beta^\omega \varepsilon_{\alpha\omega}. \tag{16.55}$$

The first term is found in equation (16.53), the second term vanishes, and the remaining terms are expanded according to the definition (16.26) of $\dot{\Gamma}_\beta^\alpha$:

$$\dot{\nabla}\varepsilon_{\alpha\beta} = \left(\nabla_\gamma V^\gamma - CB_\gamma^\gamma\right)\varepsilon_{\alpha\beta} - \left(\nabla_\alpha V^\omega - CB_\alpha^\omega\right)\varepsilon_{\omega\beta} - \left(\nabla_\beta V^\omega - CB_\beta^\omega\right)\varepsilon_{\alpha\omega}. \tag{16.56}$$

To show that the right-hand side vanishes, consider $\dot{\nabla}\varepsilon_{12}$:

$$\dot{\nabla}\varepsilon_{12} = \left(\nabla_\gamma V^\gamma - CB_\gamma^\gamma\right)\varepsilon_{12} - \left(\nabla_1 V^\omega - CB_1^\omega\right)\varepsilon_{\omega 2} - \left(\nabla_2 V^\omega - CB_2^\omega\right)\varepsilon_{1\omega}. \tag{16.57}$$

Keep the nonzero terms on the right-hand side

$$\dot{\nabla}\varepsilon_{12} = \left(\nabla_\gamma V^\gamma - CB_\gamma^\gamma - \nabla_1 V^1 + CB_1^1 - \nabla_2 V^2 + CB_2^2\right)\varepsilon_{12}. \tag{16.58}$$

It is clear that this sum vanishes which shows that $\dot{\nabla}\varepsilon_{12} = 0$. The other four entries of the tensor $\dot{\nabla}\varepsilon_{\alpha\beta}$ can be analyzed in a similar fashion and we conclude that

$$\dot{\nabla}\varepsilon_{\alpha\beta} = 0. \tag{16.59}$$

The metrinilic property of $\dot{\nabla}$ with respect to the contravariant Levi-Civita symbol $\varepsilon^{\alpha\beta}$ is left as an exercise.

Exercise 345. Show that

$$\dot{\nabla}\varepsilon^{\alpha\beta} = 0. \tag{16.60}$$

16.7 The Formula for $\dot{\nabla}\mathbf{N}$

From equation (16.35) it is easy to obtain the formula for $\dot{\nabla}\mathbf{N}$. Dot both sides of (16.35) with $\mathbf{N}$:

$$\mathbf{N}\cdot\dot{\nabla}\mathbf{S}_\alpha = \nabla_\alpha C. \tag{16.61}$$

By the product rule,

$$-\dot{\nabla}\mathbf{N}\cdot\mathbf{S}_\alpha = \nabla_\alpha C. \tag{16.62}$$

Therefore

$$\dot{\nabla}\mathbf{N} = -\mathbf{S}^\alpha\nabla_\alpha C. \tag{16.63}$$

Equation (16.63) is called the *Thomas formula* for Tracy Thomas who was the first to derive it. In component form, the Thomas formula reads

$$\dot{\nabla}N^i = -Z^i_\alpha\nabla^\alpha C. \tag{16.64}$$

16.8 The Formula for $\dot{\nabla}B^\alpha_\beta$

Start with equation

$$B^\alpha_\beta = -\mathbf{S}^\alpha\cdot\nabla_\beta\mathbf{N}, \tag{16.65}$$

which is the vector form of equation (12.41). Apply $\dot{\nabla}$ to both sides:

$$\dot{\nabla}B^\alpha_\beta = -\dot{\nabla}\mathbf{S}^\alpha\cdot\nabla_\beta\mathbf{N} - \mathbf{S}^\alpha\cdot\dot{\nabla}\nabla_\beta\mathbf{N}. \tag{16.66}$$

By equation (16.31), we find

$$\dot{\nabla}B^\alpha_\beta = -\dot{\nabla}\mathbf{S}^\alpha\cdot\nabla_\beta\mathbf{N} - \mathbf{S}^\alpha\cdot\nabla_\beta\dot{\nabla}\mathbf{N} - \mathbf{S}^\alpha\cdot CB^\omega_\beta\nabla_\omega\mathbf{N}. \tag{16.67}$$

Since the first and the third terms cancel, the formula for $\dot{\nabla}B^\alpha_\beta$ reads

$$\dot{\nabla}B^\alpha_\beta = \nabla^\alpha\nabla_\beta C + CB^\alpha_\gamma B^\gamma_\beta. \tag{16.68}$$

This formula is found in [16]. Contracting the free indices yields the previously mentioned equation (15.49)

$$\dot{\nabla}B^\alpha_\alpha = \nabla_\alpha\nabla^\alpha C + CB^\alpha_\beta B^\beta_\alpha. \tag{15.49}$$

Exercise 346. Show that the formula for the derivative of the Gaussian curvature K reads

$$\dot{\nabla}K = B^\alpha_\alpha\nabla_\beta\nabla^\beta C - B^\alpha_\beta\nabla_\alpha\nabla^\beta C + CKB^\alpha_\alpha. \tag{16.69}$$

In the next chapter, this formula is used to prove a version of the Gauss–Bonnet theorem.

16.9 Summary

In this chapter, we extended the invariant time derivative $\dot{\nabla}$ to variants arbitrary indicial signature. The general definition of $\dot{\nabla}$ is given in equation (16.27). In summary, we give the differentiation table of the calculus of moving surfaces:

$$\dot{\nabla}\mathbf{Z}_i,\ \ \dot{\nabla}\mathbf{Z}^i = \mathbf{0} \tag{16.70}$$

$$\dot{\nabla}Z_{ij},\ \ \dot{\nabla}Z^{ij},\ \ \dot{\nabla}\delta^i_j = 0 \tag{16.71}$$

$$\dot{\nabla}\varepsilon_{ijk},\ \ \dot{\nabla}\varepsilon^{ijk} = 0 \tag{16.72}$$

$$\dot{\nabla}\mathbf{S}_\alpha = \mathbf{N}\nabla_\alpha C \tag{16.73}$$

$$\dot{\nabla}\mathbf{N} = -\mathbf{S}_\alpha\nabla^\alpha C \tag{16.74}$$

$$\dot{\nabla}Z^i_\alpha = N^i\nabla_\alpha C \tag{16.75}$$

$$\dot{\nabla}S_{\alpha\beta},\ \ \dot{\nabla}S^{\alpha\beta},\ \ \dot{\nabla}\delta^\alpha_\beta = 0 \tag{16.76}$$

$$\dot{\nabla}\varepsilon_{\alpha\beta},\ \ \dot{\nabla}\varepsilon^{\alpha\beta} = 0 \tag{16.77}$$

$$\dot{\nabla}B^\beta_\alpha = \nabla_\alpha\nabla^\beta C + CB^\gamma_\alpha B^\beta_\gamma \tag{16.78}$$

Chapter 17
Applications of the Calculus of Moving Surfaces

17.1 Preview

The applications of the calculus of moving surfaces are remarkably broad. Of course, many of the applications come from problems in physics and engineering in which physical surfaces move. On the other hand, numerous applications come from problems in which, at least in the statement of the problem, there are no moving surfaces. There are at least three categories of such problems: shape optimization, boundary perturbation, and a third category illustrated by the proof of a version of the Gauss–Bonnet theorem.

Shape optimization is a topic in the calculus of variations, in which a certain quantity is minimized with respect to a function. In shape optimization, the unknown function is the shape of a domain. The central idea in the calculus of variations is to reduce what is essentially an infinite-dimensional optimization problem to a one-dimensional problem by introducing a smooth family of functions, parameterized by a single variable t, and then searching for the best function in the given family. When this idea is carried over to shape optimization, a smooth family of functions becomes a smooth family of shapes. In other words, we are dealing with a moving surface!

A classical shape optimization problem in finding a surface of minimal surface area that has a given contour boundary. Such surfaces are called *minimal*. We show that minimal surfaces are characterized by zero mean curvature B_α^α. We also derive the geodesic equation.

Boundary perturbation problems often arise in the study of partial differential equations. Suppose that a boundary value problem has been solved on some domain Ω and let ψ denote the solution. It is often necessary to solve a similar problem on another domain Ω^*, the solution being ψ^*. An approximation to ψ^* can be found by constructing an evolution of domains Ω_t (now we have a moving surface!) that starts as Ω at time zero and arrives at Ω^* at a later time t^*. Let ψ_t be solution to the boundary value problem on Ω_t. Then at $t = 0$, ψ_t coincides with ψ and at $t = t^*$, ψ_t coincides with ψ^*. The calculus of moving surfaces provides tools

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*, DOI 10.1007/978-1-4614-7867-6_17,

for expressing the derivatives of ψ_t with respect to time. The derivatives of ψ_t can typically be evaluated at time $t = 0$, thereby enabling us to form a Taylor series approximation for ψ^*.

In the last category of problems, moving surfaces are introduced as an artificial device. Consider the statement of the celebrated Gauss–Bonnet theorem. In its simplest form, it states that for a closed smooth surface S of genus zero (that is, without bagel-like holes), the integral of Gaussian curvature K is independent of the shape and equals 4π:

$$\int_S KdS = 4\pi. \tag{17.1}$$

The integral of K is called the *total curvature* of the surface. It is easy to verify that the Gauss–Bonnet theorem holds for a unit sphere. For an arbitrary smooth surface S of genus zero, we can construct a smooth evolution from the unit sphere to S and show that the rate of change in the total curvature vanishes. Therefore, the total curvature of S is 4π as well. What a nice and unexpected application of the calculus of moving surfaces!

17.2 Shape Optimization

17.2.1 The Minimal Surface Equation

A classical problem in shape optimization is to find a surface S of least possible area that spans a prescribed contour boundary Γ in three dimensions. This problem has a long and rich history. A discussion of this problem is included in almost every text on the calculus of variations. The reader might recall a lengthy derivation of the governing equation based on the Euler–Lagrange equation.

In shape optimization problems such as this one, the calculus of moving surfaces shines at its brightest (perhaps by contrast with the alternative methods). With the help of the calculus of moving surfaces, the governing equation is derived in a single line. Consider a one-parameter family of surfaces $S(t)$, each spanned by a stationary contour Γ. In the following discussion, we omit the reference to the parameter t and denote the family of surfaces simply by S. The area A of S is given by the integral

$$A = \int_S dS. \tag{17.2}$$

Differentiating with respect to t, we find by equation (15.56)

$$\frac{dA}{dt} = -\int_S CB^\alpha_\alpha dS. \tag{17.3}$$

Since, for an optimal surface, dA/dt must vanish for all independent variations and since C can be chosen arbitrarily, the equilibrium equation reads

$$B_\alpha^\alpha = 0. \tag{17.4}$$

In words, *a minimum surface has zero mean curvature*.

Two aspects of this derivation are noteworthy. First is its remarkable compactness, which shows that the calculus of moving surfaces is ideally suited to this type of problems. Second is that the derivation and therefore equation (17.4) are valid for hypersurfaces in a Euclidean space of arbitrary dimension. For example, in a two-dimensional Euclidean space, surface area corresponds to contour length. Therefore, not surprisingly, equation (17.4) tells us that a minimal surface is a straight line.

17.2.2 The Isoperimetric Problem

The classical isoperimetric problem is to find a surface of least surface area that encloses a given volume V. The constraint is accounted for by introducing a Lagrange multiplier λ:

$$E = \int_S dS + \lambda \left(\int_\Omega d\Omega - V \right). \tag{17.5}$$

The derivative of E reads

$$\frac{dE}{dt} = \int_S C \left(-B_\alpha^\alpha + \lambda \right) dS. \tag{17.6}$$

Therefore, the equilibrium equation reads

$$B_\alpha^\alpha = \lambda. \tag{17.7}$$

In words, a minimal surface that incloses a given volume has *constant mean curvature*.

It is easy to see that a sphere of radius R, such that $4\pi R^3/3 = V$, satisfies this condition and the constraint, since for a sphere in three dimensions, $B_\alpha^\alpha = -2/R$, the equilibrium configuration has

$$\lambda = -\frac{2}{R}. \tag{17.8}$$

The equilibrium equations (17.4) and (17.7) constitute necessary, but not sufficient conditions for a minimum. A shape satisfying equation (17.4) or (17.7) may yield a maximum or a stationary point. The sign of the second derivative d^2E/dt^2 can help distinguish between local minima, maxima, and stationary points.

17.2.3 The Second Variation Analysis for the Isoperimetric Problem

Apply equation (15.54) to (17.6)

$$\frac{d^2E}{dt^2} = \int_S \dot{\nabla}\left(C\left(-B_\alpha^\alpha + \lambda\right)\right) dS - \int_S C^2 B_\beta^\beta \left(-B_\alpha^\alpha + \lambda\right) dS. \tag{17.9}$$

By the product rule, the first integral becomes

$$\frac{d^2E}{dt^2} = \int_S \dot{\nabla}C\left(-B_\alpha^\alpha + \lambda\right) dS + \int_S C\dot{\nabla}\left(-B_\alpha^\alpha + \lambda\right) dS - \int_S C^2 B_\beta^\beta \left(-B_\alpha^\alpha + \lambda\right) dS. \tag{17.10}$$

We are interested in evaluating d^2E/dt^2 for the equilibrium configuration that satisfies equation (17.7). Therefore, the first and the third integrals in (17.10) vanish:

$$\frac{d^2E}{dt^2} = \int_S C\dot{\nabla}\left(-B_\alpha^\alpha + \lambda\right) dS. \tag{17.11}$$

By equation (15.49),

$$\frac{d^2E}{dt^2} = \int_S \left(-C\nabla_\alpha\nabla^\alpha C - C^2 B_\beta^\alpha B_\alpha^\beta\right) dS. \tag{17.12}$$

By Gauss's theorem,

$$\frac{d^2E}{dt^2} = \int_S \left(\nabla_\alpha C\nabla^\alpha C - C^2 B_\beta^\alpha B_\alpha^\beta\right) dS. \tag{17.13}$$

The first term in equation (17.13) is positive (or zero) and the second is negative for any nontrivial C.

Equation (17.12) is best for determining the sign of d^2E/dt^2 for a general variation C. Decompose C as a series in spherical harmonics Y_{lm}

$$C = \sum_{l>0,\ |m|\le l} C_{lm} Y_{lm}, \tag{17.14}$$

where $l = 0$ is disallowed because it violates the volume constraint. Spherical harmonics Y_{lm} are eigenvalues of the surface Laplacian

$$\nabla_\alpha\nabla^\alpha Y_{lm} = -\frac{l\,(l+1)}{R^2} \tag{17.15}$$

and form an orthonormal set on the unit sphere S_0

$$\int_{S_0} Y_{l_1 m_1} Y_{l_2 m_2} dS = \begin{cases} 1, \text{if } l_1 = l_2 \text{ and } m_1 = m_2 \\ 0, \text{otherwise} \end{cases}. \tag{17.16}$$

For a sphere of radius R, we have

$$B^\alpha_\beta B^\beta_\alpha = \frac{2}{R^2}. \tag{17.17}$$

Substituting these relationships into equation (17.12), we find

$$\frac{d^2E}{dt^2} = 4\pi \sum_{l>0,m} C^2_{lm} (l-1)(l+2). \tag{17.18}$$

The spherical harmonic $Y_{00}(\theta,\phi) = (4\pi)^{-1/2}$, which corresponds to $l = 0$, is excluded because $\int Y_{00} dS > 0$ and it violates the constant area constraint. From this expression, we conclude that the sphere is neutrally stable ($d^2E/dt^2 = 0$) with respect to translation as a rigid body ($l = 1$) and is stable ($d^2E/dt^2 > 0$) with respect to all other smooth deformations.

17.2.4 The Geodesic Equation

Consider two points A and B on a two-dimensional surface S characterized by the metric tensor $S_{\alpha\beta}$. A *geodesic* is the shortest surface curve that connects A and B. Let the geodesic be given parametrically by

$$S^\alpha = S^\alpha (U). \tag{17.19}$$

A common derivation [31] of the equation characterizing geodesics is based on formulating the Euler–Lagrange equations for the arithmetic integral that yields the length of the curve

$$L = \int_\alpha^\beta \sqrt{S_{\alpha\beta} \frac{dS^\alpha}{dU} \frac{dS^\beta}{dU}} dU. \tag{17.20}$$

We outline a different approach based on minimal surfaces. After all, geodesic is a minimal hypersurface and must therefore be characterized by an equation analogous to (17.4). The difficulty lies in the fact that equation (17.4) is valid for minimal surfaces embedded in *Euclidean* spaces. Meanwhile, the surface S is not a Euclidean space. Nevertheless, we state (without proof) that the general definition (16.27) of the invariant time derivative $\dot\nabla$ remains valid in Riemann spaces. Furthermore, the law (15.56) governing the time evolution of surface integrals remains valid as well. Therefore, the derivation of Sect. 17.2.1 continues to apply to minimal surfaces in Riemann spaces. We reach the conclusion that *minimal surfaces in Riemann spaces are characterized by zero mean curvature*. When applied to one-dimensional curves embedded in two-dimensional surfaces, the term *mean curvature* becomes *geodesic curvature*. This was discussed in Sect. 13.9. Thus, geodesics are characterized by zero geodesic curvature:

$$b_\Phi^\Phi = 0. \tag{17.21}$$

Equation (17.21) is known as the *geodesic equation* or *equation of a geodesic*.

Let us translate equation (17.21) into an explicit equation for $S^\alpha(U)$. Multiply both sides of equation (17.21) by n^α

$$b_\Phi^\Phi n^\alpha = 0 \tag{17.22}$$

and recall that $b_\Phi^\Phi n^\alpha$ is given by

$$b_\Phi^\Phi n_\alpha = U^{\Phi\Psi} \nabla_\Phi S_\Psi^\alpha. \tag{17.23}$$

Therefore, equation (17.21) is equivalent to

$$U^{\Phi\Psi} \nabla_\Phi S_\Psi^\alpha = 0. \tag{17.24}$$

Expanding the covariant derivative ∇_Φ, we find

$$U^{\Phi\Psi}\left(\frac{\partial S_\Psi^\alpha}{\partial U^\Phi} + S_\Phi^\beta \Gamma_{\beta\gamma}^\alpha S_\Psi^\gamma - \Gamma_{\Phi\Psi}^\Omega S_\Omega^\alpha\right) = 0. \tag{17.25}$$

In terms of $S^\alpha(U)$, this equation reads

$$U^{\Phi\Psi}\left(\frac{\partial^2 S^\alpha}{\partial U^\Phi \partial U^\Psi} + \Gamma_{\beta\gamma}^\alpha \frac{\partial S^\beta}{\partial U^\Phi}\frac{\partial S^\gamma}{\partial U^\Psi} - \Gamma_{\Phi\Psi}^\Omega \frac{\partial S^\alpha}{\partial U^\Omega}\right) = 0. \tag{17.26}$$

Equation (17.26) characterizes minimal hypersurfaces of Riemann spaces of arbitrary dimension. Note that α is a free index and therefore there is a single equation for each $S^\alpha(U)$. In order to adapt this equation to the case of one-dimensional curves, denote by $\tilde{\Gamma}$ the single entry of the Christoffel tensor $\Gamma_{\Phi\Psi}^\Omega$ on the curve and cancel U^{11}:

$$\frac{d^2 S^\alpha}{dU^2} + \Gamma_{\beta\gamma}^\alpha \frac{dS^\beta}{dU}\frac{dS^\gamma}{dU} - \tilde{\Gamma}\frac{dS^\alpha}{dU} = 0. \tag{17.27}$$

Finally, if the curve is referred to arc length s, $\tilde{\Gamma}$ vanishes and equation (17.27) reads

$$\frac{d^2 S^\alpha}{ds^2} + \Gamma_{\beta\gamma}^\alpha \frac{dS^\beta}{ds}\frac{dS^\gamma}{ds} = 0. \tag{17.28}$$

Equation (17.28) gives the form in which the geodesic equation usually appears. Letting

$$T^\alpha = \frac{dS^\alpha}{ds}, \tag{17.29}$$

denote the unit tangent, equation (17.28) may be written in the form

$$\frac{\delta T^\alpha}{\delta s} = 0. \tag{17.30}$$

In words, the unit tangent forms a parallel vector field (in the sense of Sects. 8.9 and 12.7) along a geodesic referred to the arc length s.

17.3 Evolution of Boundary Conditions in Boundary Value Problems

A boundary value problem is solved on a domain Ω with boundary S. The unknown function u is determined by a partial differential equation inside Ω and a boundary condition prescribed on the boundary S. The classical boundary value problem is the Poisson equation

$$\nabla_i \nabla^i u = f, \tag{17.31}$$

where f is a known function on Ω.

There are three fundamental types of boundary conditions. *Dirichlet* boundary conditions prescribe the value of the function u along the boundary S. Our focus is the zero Dirichlet boundary conditions

$$u|_S = 0. \tag{17.32}$$

Neumann boundary conditions prescribe the value of the normal derivative along S. Our focus is once again the zero Neumann condition

$$N^i \nabla_i u = 0. \tag{17.33}$$

Finally, *mixed* boundary conditions prescribe the value of a linear combination $Au + BN^i \nabla_i u$ of the function u and its normal derivative $N^i \nabla_i u$. The quantities A and B may be variable functions on the boundary. We focus on the case where A is a constant, $B = 1$ and the prescribed value is zero

$$Au + N^i \nabla_i u = 0. \tag{17.34}$$

Suppose that the domain Ω evolves with time t and that $u(t, Z)$ is the solution to a particular boundary value problem subject to one of the three fundamental types of boundary conditions. Our goal is to derive the boundary value problem that governs the rate of change

$$u_1 = \frac{\partial u(t, Z)}{\partial t} \tag{17.35}$$

of $u(t, Z)$. Regardless of the differential relationship that $u(t, Z)$ satisfies in the interior of Ω, the boundary conditions for u_1 can be determined from the boundary conditions for u. In each case, the boundary condition is derived by applying the $\dot{\nabla}$ operator to the boundary condition for u.

For Dirichlet boundary conditions (17.32), an application of $\dot{\nabla}$ yields

$$\dot{\nabla} u = 0. \tag{17.36}$$

Expanding, we find

$$\frac{\partial u}{\partial t} + CN^i \nabla_i u = 0, \tag{17.37}$$

or

$$u_1 = -CN^i \nabla_i u. \tag{17.38}$$

Thus, we conclude that Dirichlet boundary conditions for u result in Dirichlet boundary conditions for u_1.

For Neumann boundary conditions (17.33), we find

$$\dot{\nabla} \left(N^i \nabla_i u \right) = 0. \tag{17.39}$$

By the product rule and the Thomas rule (16.64), we find

$$-Z^i_\alpha \nabla^\alpha C \nabla_i u + N^i \dot{\nabla} \nabla_i u = 0. \tag{17.40}$$

Apply the chain rule (15.53) to the expression $\dot{\nabla} \nabla_i u$:

$$\dot{\nabla} \nabla_i u = \frac{\partial}{\partial t} \nabla_i u + CN^j \nabla_j \nabla_i u. \tag{17.41}$$

Since the covariant derivative ∇_i commutes with the partial derivative $\partial / \partial t$, we have

$$\dot{\nabla} \nabla_i u = \nabla_i u_1 + CN^j \nabla_j \nabla_i u. \tag{17.42}$$

We thus arrive at the following condition for u_1

$$N^i \nabla_i u_1 = Z^i_\alpha \nabla^\alpha C \nabla_i u - CN^i N^j \nabla_i \nabla_j u. \tag{17.43}$$

or, equivalently,

$$N^i \nabla_i u_1 = \nabla^\alpha C \nabla_\alpha u - C N^i N^j \nabla_i \nabla_j u. \tag{17.44}$$

Thus, similarly to the Dirichlet case, Neumann boundary conditions for u lead to Neumann boundary conditions for u_1.

Finally, we turn to mixed boundary conditions (17.34) for constant A and B. Applying the $\dot{\nabla}$ operator to equation (17.34), we find

$$A\dot{\nabla} u + B\dot{\nabla}\left(N^i \nabla_i u\right) = 0. \tag{17.45}$$

Combining the analyses of the Dirichlet and Neumann cases, we find

$$Au_1 + BN^i \nabla_i u_1 = -ACN^i \nabla_i u + B\nabla^\alpha C \nabla_\alpha u - CBN^i N^j \nabla_j \nabla_i u. \tag{17.46}$$

Thus, mixed conditions for u lead to mixed conditions for u_1.

17.4 Eigenvalue Evolution and the Hadamard Formula

In this section, we study the dependence of Laplace eigenvalues on shape. More specifically, we analyze the rate of change in the eigenvalue induced by smooth deformations of the domain. The eigenvalues λ are determined by the Laplace eigenvalue equation

$$\nabla_i \nabla^i u = -\lambda u \tag{17.47}$$

in combination with Dirichlet (17.32), Neumann (17.33), or mixed (17.34) boundary condition on the boundary S of the domain Ω on which the interior condition (17.47) is to be satisfied. The interior equation and the boundary condition, define the eigenfunction u to within a multiplicative constant. The usual way to remove this arbitrariness is to normalize the eigenfunction to unity according to equation

$$\int_\Omega u^2 d\Omega = 1. \tag{17.48}$$

With the normalization condition in place, the eigenfunction u is determined to within a sign.

When the domain Ω, with its boundary S, evolve as a function of time t, so do the Laplace eigenvalues λ. If we denote the evolution of the domain by $\Omega(t)$ and $S(t)$ and the evolution of the eigenvalues by $\lambda(t)$, our goal is to determine the expression for the time derivative λ_1 of $\lambda(t)$:

$$\lambda_1 = \frac{d\lambda(t)}{dt}. \tag{17.49}$$

The eigenvalue λ corresponding to the eigenfunction u is given by the Rayleigh quotient

$$\lambda = \frac{\int_\Omega \nabla_i u \nabla^i u d\Omega}{\int_\Omega u^2 d\Omega}. \tag{17.50}$$

Since we have introduced the normalization condition (17.48), the Rayleigh quotient reduces to its numerator

$$\lambda = \int_\Omega \nabla_i u \nabla^i u d\Omega. \tag{17.51}$$

The Rayleigh quotient is the starting point of our analysis. However, let us save this analysis for last and first determine the relationships that the interior equation (17.47) and the normalization condition (17.48) imply about the time derivative u_1 of the eigenfunction u:

$$u_1(t, Z) = \frac{\partial u(t, Z)}{\partial t}. \tag{17.52}$$

The function u_1 is called the first variation of the eigenfunction u.

Differentiate the interior equation (17.47) with respect to time

$$\frac{\partial}{\partial t} \nabla_i \nabla^i u(t, Z) = -\frac{\partial}{\partial t} (\lambda(t) u(t, Z)). \tag{17.53}$$

It is left as an exercise to show that the covariant and contravariant derivatives ∇_i and ∇^i commute with partial time differentiation. The right-hand side is expanded by the product rule. The result is

$$\nabla_i \nabla^i u_1 = -\lambda_1 u - \lambda u_1. \tag{17.54}$$

Transferring the term $-\lambda u_1$ to the left-hand side, we obtain the final equation for u_1

$$\nabla_i \nabla^i u_1 + \lambda u_1 = -\lambda_1 u. \tag{17.55}$$

Therefore, the first variation of the eigenfunction satisfies an inhomogeneous Helmholtz equation.

The normalization condition (17.48) also gives us information regarding u_1. Differentiate equation (17.48) with respect to time t:

$$\frac{d}{dt} \int_\Omega u^2 d\Omega = 0. \tag{17.56}$$

By equation (15.54), we find

$$\int_\Omega 2uu_1 d\Omega + \int_S Cu^2 dS = 0. \tag{17.57}$$

Under Dirichlet conditions (17.32), the surface integral vanishes identically in time. Therefore, under Dirichlet conditions, the first eigenfunction variation u_1 is orthogonal to the eigenfunction u in the sense

$$\int_\Omega uu_1 d\Omega = 0. \tag{17.58}$$

We are now in a position to derive the Hadamard formula for λ_1. Let us begin with Dirichlet conditions. As we started before, our starting point is the Rayleigh quotient (17.51). Applying the ordinary time derivative, we have

$$\lambda_1 = \frac{d}{dt}\int_\Omega \nabla_i u \nabla^i u d\Omega. \tag{17.59}$$

By the volume integral law (15.54), we have

$$\lambda_1 = \int_\Omega \frac{\partial\left(\nabla_i u \nabla^i u\right)}{\partial t} d\Omega + \int_S C\nabla_i u \nabla^i u dS. \tag{17.60}$$

By a combination of the product rule and the fact that partial differentiation commutes with covariant differentiation, we find

$$\lambda_1 = \int_\Omega 2\nabla_i u \nabla^i u_1 d\Omega + \int_S C\nabla_i u \nabla^i u dS. \tag{17.61}$$

Use Gauss's theorem to transfer the covariant derivative ∇^i from u to $\nabla_i u_1$:

$$\lambda_1 = 2\int_S u\, N_i \nabla^i u_1 dS - 2\int_\Omega u \nabla_i \nabla^i u_1 d\Omega + \int_S C\nabla_i u \nabla^i u dS. \tag{17.62}$$

By the Dirichlet condition (17.32) the first surface integral vanishes. Furthermore, equation (17.54) gives us an expression for $\nabla_i \nabla^i u_1$. Thus

$$\lambda_1 = 2\int_\Omega \left(\lambda_1 u^2 + \lambda u u_1\right) d\Omega + \int_S C\nabla_i u \nabla^i u dS. \tag{17.63}$$

Split the first integral into two:

$$\lambda_1 = 2\lambda_1 \int_\Omega u^2 d\Omega + 2\lambda \int_\Omega uu_1 d\Omega + \int_S C\nabla_i u \nabla^i u dS. \tag{17.64}$$

The first term equals $2\lambda_1$ by the normalization condition (17.48). The second term vanishes by equation (17.58). We have therefore arrived at the following expression for λ_1

$$\lambda_1 = -\int_S C \nabla_i u \nabla^i u dS. \tag{17.65}$$

This result is known as the *Hadamard formula*.

Exercise 347. Show that, under sufficient conditions of smoothness, the partial time derivative $\partial/\partial t$ commutes with the covariant derivative $\nabla_i u$.

Exercise 348. The Laplace eigenfunctions corresponding to the simple eigenvalues on the unit circle under Dirichlet boundary conditions are given by

$$u(r,\theta) = \frac{J_0(\rho r)}{\sqrt{\pi} J_1(\rho)}, \tag{17.66}$$

where J_n are Bessel functions of the first kind and ρ is a root of J_0. Note that the Bessel functions J_0 and J_1 are connected by the relationship

$$J_0'(x) = -J_1(x). \tag{17.67}$$

Consider the following evolution from the unit circle to an ellipse with semi-axes $1+a$ and $1+b$:

$$\begin{cases} x(t,\alpha) = (1+at)\cos\alpha \\ y(t,\alpha) = (1+bt)\sin\alpha \end{cases} \tag{17.68}$$

Show that at $t = 0$, according to the Hadamard formula, we have

$$\lambda_1 = -(a+b)\lambda. \tag{17.69}$$

Exercise 349. Show that equation (17.69) can be derived without the calculus of moving surfaces by appealing to Euler's homogeneous function theorem. (Hint: Laplace eigenvalues are homogeneous functions of the ellipse semi-axis lengths).

We now derive the expression for u_1 under Neumann conditions. We once again start with the Rayleigh quotient and, by applying d/dt, arrive at the following equation

$$\lambda_1 = \int_\Omega 2\nabla_i u \nabla^i u_i d\Omega + \int_S C \nabla_i u \nabla^i u dS. \tag{17.70}$$

This time, we use Gauss's theorem to transfer the covariant derivative ∇_i from u_1 to $\nabla^i u$:

$$\lambda_1 = 2\int_S u_1\, N_i \nabla^i u dS - 2\int_\Omega u_1 \nabla_i \nabla^i u d\Omega + \int_S C \nabla_i u \nabla^i u dS. \tag{17.71}$$

The first integral vanishes by the Neumann condition (17.33). In the second integral, substitute $-\lambda u$ for $\nabla_i \nabla^i u$ by equation (17.47):

$$\lambda_1 = 2\lambda \int_\Omega u u_1 d\Omega + \int_S C \nabla_i u \nabla^i u dS. \tag{17.72}$$

Since the first integral is given by equation (17.57), we arrive at the final expression

$$\lambda_1 = \int_S C \left(\nabla_i u \nabla^i u - \lambda u^2\right) dS. \tag{17.73}$$

This equation is known as the *Hadamard formula for Neumann boundary conditions.*

Derivation of the Hadamard formula under mixed conditions is left as an exercise.

Exercise 350. Show that, under all boundary conditions among (17.32), (17.33), and (17.34),

$$\lambda_1 = \int_S \left(u_1 N_i \nabla^i u \; - u\, N_i \nabla^i u_1\right) dS. \tag{17.74}$$

Exercise 351. Use the divergence theorem to convert equation (17.74) to the following form

$$\lambda_1 = \int_\Omega \left(u \nabla_i \nabla^i u_1 - u_1 \nabla_i \nabla^i u\right) d\Omega. \tag{17.75}$$

Exercise 352. Derive the Hadamard formula for mixed boundary conditions

$$\lambda' = \int_S C \left(-\nabla_i u \nabla^i u + 2\nabla_\alpha u \nabla^\alpha u - \lambda u^2 + B^\alpha_\alpha u\, N_i \nabla^i u\right) dS \tag{17.76}$$

Exercise 353. Show that the Hadamard formula (17.76) reduces to equation (17.65) under Dirichlet conditions and equation (17.73) under Neumann conditions.

17.5 A Proof of the Gauss–Bonnet Theorem

Our proof of a special case of the Gauss–Bonnet theorem is a particularly neat application of the calculus of moving surfaces. After all, the Gauss–Bonnet theorem is a result in classical differential geometry on stationary manifolds. There are no moving surfaces in sight! Nevertheless, we reformulate the Gauss–Bonnet theorem as a problem in moving surfaces. The presented proof is excellent as a final demonstration in this book as it draws on many of the fundamental relationships in tensor calculus and the calculus of moving surfaces.

The Gauss–Bonnet theorem states that the integral of Gaussian curvature over a closed surface is independent of its shape and only depends on its *genus*. The genus (*genera* in plural) of a surface is a topological concept. In plain language, the genus of a closed surface is the number of holes that it has. Thus, a sphere is a surface of genus zero and a torus is a surface of genus one. If g is the genus of the closed surface S, the Gauss–Bonnet theorem states that

$$\int_S K dS = 4\pi (1-g) . \tag{17.77}$$

Thus, for surfaces of genus zero

$$\int_S K dS = 4\pi \tag{17.78}$$

and for surfaces of genus one

$$\int_S K dS = 0. \tag{17.79}$$

Equation (17.77) is a special case of the Gauss–Bonnet theorem. A more general Gauss–Bonnet theorem holds for nonclosed surfaces with contour boundaries. Furthermore, the Gauss–Bonnet holds for general two-dimensional Riemannian manifolds in the intrinsic sense—that is, without a reference to an embedding. Finally, in topology, the Gauss–Bonnet theorem is formulated in an even more general setting that makes no reference to the Riemannian metric.

We can prove our special case of the Gauss–Bonnet theorem by showing that the integral of Gaussian curvature

$$T = \int_S K dS,$$

known as the *total curvature*, is unchanged under sufficiently smooth evolutions. This would show that T is independent of shape, but does not quite prove that $T = 4\pi (1-g)$. However, for surfaces of genera zero and one, this can be verified by calculating T for a sphere and a torus. For surfaces of higher genera, the result can be argued (although this is not as simple as it may seem) by induction by annexing a torus to a surface of genus g to obtain a surface of genus $g+1$.

So, our goal is to show that, under a smooth evolution, the total curvature T is unchanged:

$$\frac{d}{dt}\int_S K dS = 0. \tag{17.80}$$

By the surface integral law (15.56), we have

$$\frac{d}{dt}\int_S K dS = \int_S \dot{\nabla} K dS - \int_S CKB_\alpha^\alpha dS. \tag{17.81}$$

Recall from Chap. 16 that $\dot{\nabla} K$ is given by the beautiful formula

$$\dot{\nabla} K = B_\alpha^\alpha \nabla_\beta \nabla^\beta C - B_\beta^\alpha \nabla_\alpha \nabla^\beta C + CKB_\alpha^\alpha. \tag{16.69}$$

Therefore, equation (17.81) becomes

$$\frac{d}{dt}\int_S K dS = \int_S \left(B_\alpha^\alpha \nabla_\beta \nabla^\beta C - B_\beta^\alpha \nabla_\alpha \nabla^\beta C\right) dS. \tag{17.82}$$

We next apply Gauss's theorem to each of the terms in the integral on the right-hand side. In the first term, we use Gauss's theorem to transfer ∇_β from ∇^β to B_α^α. In the second integral, Gauss's theorem helps transfer ∇_α from $\nabla^\beta C$ to B_β^α. Since the surface is closed, there arise no contour integrals. As a result, we arrive at the following expression

$$\frac{d}{dt}\int_S K dS = \int_S \left(-\nabla_\beta B_\alpha^\alpha + \nabla_\alpha B_\beta^\alpha\right) \nabla^\beta C dS. \tag{17.83}$$

We next use the Codazzi identity

$$\nabla_\alpha B_{\beta\gamma} = \nabla_\beta B_{\alpha\gamma}. \tag{12.73}$$

derived in Chap. 12. By raising the index γ and contracting it with α, we find

$$\nabla_\alpha B_\beta^\alpha = \nabla_\beta B_\alpha^\alpha. \tag{17.84}$$

Therefore, the integrand in (17.83) vanishes and equation (17.80) is confirmed!

17.6 The Dynamic Fluid Film Equations

Finally, we come to a dynamic application of the calculus of moving surfaces, and one that is perhaps closest to the author's heart. We discuss the exact equations for the dynamics of free fluid films under the influence of surface tension. We consider a fluid film that spans a stationary three-dimensional smooth curved contour. The free fluid film is modeled as a two-dimensional surface S with a thickness represented by variable surface mass density ρ. We decompose the fluid film velocity $\mathbf{V}$ in the normal component C and the tangential components V^α

$$\mathbf{V} = C\mathbf{N} + V^\alpha \mathbf{S}_\alpha. \tag{17.85}$$

The potential energy of the fluid film is captured by Laplace's model of surface tension

$$P = \sigma \int_S dS. \tag{17.86}$$

Therefore, the Lagrangian L is given by

$$L = \frac{1}{2} \int_S \rho \left(C^2 + V^\alpha V_\alpha \right) dS - \sigma \int_S dS. \tag{17.87}$$

The dynamic system [17–19] reads:

$$\dot{\nabla} \rho + \nabla_\alpha \left(\rho V^\alpha \right) = \rho C B_\alpha^\alpha \tag{17.88}$$

$$\rho \left(\dot{\nabla} C + 2 V^\alpha \nabla_\alpha C + B_{\alpha\beta} V^\alpha V^\beta \right) = \sigma B_\alpha^\alpha \tag{17.89}$$

$$\dot{\nabla} V^\alpha + V^\beta \nabla_\beta V^\alpha - C \nabla^\alpha C - C V^\beta B_\beta^\alpha = 0 \tag{17.90}$$

The first equation (17.88) represents conservation of mass. Equations (17.89) and (17.90) are the momentum equations for the normal and tangential components. The normal velocity C vanishes at the boundary Γ

$$C = 0 \tag{17.91}$$

while the in-surface components V^α are tangential to the boundary

$$n_\alpha V^\alpha = 0. \tag{17.92}$$

As an illustration, we demonstrate two key properties of these equations: conservation of mass and conservation of energy. The total mass M of the fluid film is given by the integral

$$M = \int_S \rho dS. \tag{17.93}$$

Thus, according to the surface law (15.56), M changes at the rate

$$\frac{dM}{dt} = \int_S \dot{\nabla} \rho dS - \int_S \rho C B_\alpha^\alpha dS. \tag{17.94}$$

An application of equation (17.88) yields

$$\frac{dM}{dt} = -\int_S \nabla_\alpha \left(\rho V^\alpha\right) dS. \tag{17.95}$$

Therefore, by the Gauss theorem,

$$\frac{dM}{dt} = \int_S n_\alpha \rho V^\alpha dS \tag{17.96}$$

This integral vanishes according to the boundary condition (17.92) and conservation of mass is therefore proven.

To prove conservation of energy, apply the surface law (15.56) to the total energy

$$E = \frac{1}{2}\int_S \rho \left(C^2 + V_\alpha V^\alpha\right) dS + \sigma \int_S dS. \tag{17.97}$$

The result is

$$E = \int_S \begin{pmatrix} \frac{1}{2}\dot{\nabla}\rho\left(C^2 + V^\alpha V_\alpha\right) + \rho\left(C\dot{\nabla}C + V_\alpha \dot{\nabla} V^\alpha\right) \\ -\frac{1}{2}\rho C B^\beta_\beta \left(C^2 + V^\alpha V_\alpha\right) - \sigma C B^\alpha_\alpha \end{pmatrix} dS. \tag{17.98}$$

Use equations (17.88)–(17.90) to eliminate $\dot{\nabla}\rho$, $\dot{\nabla}C$ and $\dot{\nabla}V^\alpha$ in this equation. Upon the cancellation of several terms we arrive at the following equation

$$\frac{dE}{dt} = \int_S \left(-\frac{1}{2}\nabla_\beta\left(\rho V^\beta\right)\left(C^2 + V^\alpha V_\alpha\right) - \rho C V^\alpha \nabla_\alpha C - \rho V_\alpha V^\beta \nabla_\beta V^\alpha\right) dS. \tag{17.99}$$

An application of Gauss's theorem cancels the rest of the terms and the resulting boundary term vanishes due to the boundary condition (17.92). This completes the proof on the conservation of energy.

Exercise 354. Combine equations (17.89)–(17.90) into a single second-order vector equation.

17.7 Summary

In this final chapter of the book, you experienced the versatility of the calculus of moving surfaces. Interestingly, we presented the calculus of moving surfaces in a relatively narrow setting: embedded surfaces in a Euclidean space. There exists an extension of the calculus of moving surfaces to surfaces of arbitrary dimension embedded in deforming Riemannian spaces. Nevertheless, even the relatively narrow version of the calculus of moving surfaces presented here finds a broad range of applications.

Bibliography

[1] L. Bewley. *Tensor Analysis of Electric Circuits and Machines*. Ronald Press, 1961.

[2] A. Borisenko and I.E.Tarapov. *Vector and Tensor Analysis with Applications*. Dover Publications, New York, 1979.

[3] S. Carroll. *Spacetime and Geometry: An Introduction to General Relativity*. Benjamin Cummings, 2003.

[4] E. Cartan. *Geometry of Riemannian Spaces*. Math Science Pr, 1983.

[5] E. Cartan. *Riemannian Geometry in an Orthogonal Frame: From Lectures Delivered by Elie Cartan at the Sorbonne in 1926–27*. World Scientific Pub Co Inc, 2002.

[6] I. Chavel. *Riemannian Geometry: A Modern Introduction (Cambridge Studies in Advanced Mathematics)*. Cambridge University Press, 2006.

[7] E. Christoffel. Sul problema delle temperature stazionarie e la rappresentazione di una data superficie. *Annali di Matematica Pura ed Applicata*, 1(1):89–103, 1867. On the problem of stationary temperature and the representation of a given area.

[8] D. Danielson. *Vectors And Tensors In Engineering And Physics: Second Edition*. Westview Press, 2003.

[9] R. Descartes. *The Geometry*. Dover Publications, New York, 1954.

[10] A. Einstein. Die grundlage der allgemeinen relativittstheorie. *Ann. der Physik*, 49:769–822, 1916.

[11] Euclid. *The Elements: Books I - XIII - Complete and Unabridged*. Barnes & Noble, 2006.

[12] L. Euler. Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici lattissimo sensu accepti. *Opera Omnia*, 24(1), 1897.

[13] I. Gelfand. *Lectures on Linear Algebra*. Dover Publications, New York, 1989.

[14] J. W. GIbbs. *Vector analysis: A text-book for the use of students of mathematics and physics*. Dover Publications, 1960.

[15] I. Grattan-Guiness. *From the calculus to set theory, 1630–1910: an introductory history*. Princeton University Press, Princeton, 2000.

[16] M. Grinfeld. *Thermodynamic Methods in the Theory of Heterogeneous Systems*. Longman, New York, NY, 1991.

[17] P. Grinfeld. Exact nonlinear equations for fluid films and proper adaptations of conservation theorems from classical hydrodynamics. *J. Geom. Symm. Phys.*, 16:1–21, 2009.

[18] P. Grinfeld. Hamiltonian dynamic equations for fluid films. *Stud. Appl. Math.*, 125:223–264, 2010.

[19] P. Grinfeld. A variable thickness model for fluid films under large displacements. *Phys. Rev. Lett.*, 105:137802, 2010.

[20] P. Grinfeld. A better calculus of moving surfaces. *J. Geom. Symm. Phys.*, 26:61–69, 2012.

[21] J. Hadamard. *Mmoire sur le problme d'analyse relatif l'quilibre des plaques elastiques encastres, Oeuvres, tome 2*. Hermann, 1968.

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*, DOI 10.1007/978-1-4614-7867-6,

[22] P. Halmos. *Finite-dimensional vector spaces*. Springer-Verlag, New York, 1974.
[23] F. Harley. *Differential forms, with applications to the physical sciences*. Academic Press, New York, 1963.
[24] V. Katz. The history of stokes' theorem. *Helv. Phys. Acta. Supp.*, 52(3):146–156, 1979.
[25] L. Kollros. Albert einstein en suisse souvenirs. *Helv. Phys. Acta. Supp.*, 4:271–281, 1956.
[26] J. Lagrange. Essai d'une nouvelle mthode pour dterminer les maxima et les minima des formules intgrales indfinies. *Miscellanea Taurinensia*, 1761.
[27] P. Lax. *Linear algebra and its applications*. Wiley-Interscience, Hoboken, N.J, 2007.
[28] T. Levi-Civita. *The Absolute Differential Calculus (Calculus of Tensors)*. Dover Publications, 1977.
[29] S. Lovett. *Differential Geometry of Manifolds*. A K Peters Ltd, 2010.
[30] J. Maxwell. *Treatise on Electricity and Magnetism*. Cambridge University Press, Cambridge, 2010.
[31] A. McConnell. *Applications of Tensor Analysis*. Dover Publications, New York, 1957.
[32] F. Morgan. *Riemannian Geometry: A Beginners Guide, Second Edition*. A K Peters/CRC Press, 1998.
[33] P. Nastasia and R. Tazzioli. Toward a scientific and personal biography of Tullio Levi-Civita. *Historia Mathematica*, 32(2):203–236, 2005.
[34] G. Ricci and T. Levi-Civita. Mthodes de calcul diffrentiel absolu et leurs applications. *Mathematische Annalen*, 54:125–201, 1900.
[35] B. Riemann. Ueber die hypothesen, welche der geometrie zu grunde liegen. *Abhandlungen der Kniglichen Gesellschaft der Wissenschaften zu Gttingen*, 13, 1867. On the hypotheses that lie at the foundation of geometry.
[36] B. Riemann. *Gesammelte Mathematische Werke*. Unknown, 1919. The Collected Mathematical Works with commentary by Hermann Weyl.
[37] B. Riemann. *Sochineniya*. GITTL, 1948. The Collected Mathematical Works with commentary by Hermann Weyl, in Russian.
[38] W. Rudin. *Principles of mathematical analysis*. McGraw-Hill, New York, 1976.
[39] H. Schwarz. Sur une definition erronee de aire d'une surface courbe. *Ges. Math. Abhandl*, 2:309–311, 369–370, 1882.
[40] J. Simmonds. *A brief on Tensor Analysis*. Springer, New York Berlin, 1994.
[41] I. Sokolnikoff. *Tensor Analysis: Theory and Applications to Geometry and Mechanics of Continua*. Krieger Pub Co, 1990.
[42] B. Spain. *Tensor Calculus: a concise course*. Dover Publications, Mineola, N.Y, 2003.
[43] M. Spivak. *Calculus on manifolds : a modern approach to classical theorems of advanced calculus*. W.A. Benjamin, New York, 1965.
[44] G. Strang. *Introduction to Linear Algebra, 4th edition*. Wellesley-Cambridge Press, 2009.
[45] T. Thomas. *Plastic Flow and Fracture in Solids*. Academic Press, New York, NY, 1961.
[46] T. Thomas. *Concepts from Tensor Analysis And Differential Geometry*. Academic Press, New York, 1965.
[47] H. Weyl. *Space, Time, Matter*. Dover Publications, 1952.
[48] F. Zames. Surface area and the cylinder area paradox. *The Two-Year College Mathematics Journal*, 8(4):207–211, 1977.

Index

P. Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*,
DOI 10.1007/978-1-4614-7867-6,