20/04/2024 04:15

(Article begins on next page)

Tensors: A guide for undergraduate students

Franco Battaglia
Dipartimento di Ingegneria Enzo Ferrari, Università di Modena e Reggio Emilia, Modena 41125, Italy

Thomas F. George
Departments of Chemistry/Biochemistry and Physics/Astronomy, University of Missouri–St. Louis, St. Louis, Missouri 63121

A guide on tensors is proposed for undergraduate students in physics or engineering that ties directly to vector calculus in orthonormal coordinate systems. We show that once orthonormality is relaxed, a dual basis, together with the contravariant and covariant components, naturally emerges. Manipulating these components requires some skill that can be acquired more easily and quickly once a new notation is adopted. This notation distinguishes multi-component quantities in different coordinate systems by a differentiating sign on the *index* labelling the component rather than on the label of the quantity itself. This tiny stratagem, together with simple rules openly stated at the beginning of this guide, allows an almost automatic, easy-to-pursue procedure for what is otherwise a cumbersome algebra. By the end of the paper, the reader will be skillful enough to tackle many applications involving tensors of any rank in any coordinate system, without index-manipulation obstacles standing in the way. © *2013 American Association of Physics Teachers.*
[http://dx.doi.org/10.1119/1.4802811]

I. INTRODUCTION

Undergraduate students, to whom this paper is addressed, are generally aware of the importance of tensors. However, as pointed out by others,[1] unless students have been engaged in a course dedicated to the study of tensors, the subject remains covered in a veil of mystery. Given that an understanding of tensors helps aid in understanding even simple experiments carried out in undergraduate laboratory courses,[2] this is unfortunate and is due, in our opinion, to two major causes. First, a lack in the literature of a presentation of tensors that is tied to what students already know from courses they have already taken, such as calculus, vector algebra, and vector calculus. And second, the use of a notation that turns the necessary use of multiple indices into an unnecessarily cumbersome process.

To be sure, there exist plenty of excellent books, such as *The Absolute Differential Calculus* by Levi-Civita,[3] to name a classic book written by one of the founders of the field. This outstanding treatise, however, starts with three long chapters on functional determinants and matrices, systems of total differential equations, and linear partial differential equations, before entering into the algebraic foundations of tensor calculus. And even so, little connection is made to the vector algebra and calculus that are a standard part of the undergraduate physics curriculum. Given the treatise nature of this book (which, by the way, is a must read by anybody who aims at being a professional on the subject), this circumstance is not terribly surprising, but it does make the book only somewhat appropriate to those readers (undergraduates) we have in mind here.

The excellent book by Schouten,[4] another founder of tensor analysis, in spite of its promising title, *Tensor Analysis for Physicists,* leaves us with some doubts on its appropriateness for undergraduate students. It is enough to browse through the book to realize that the target audience is quite advanced. The author himself, in the preface to the first edition, claims that some readers "can avoid actually working through the whole of the earlier chapters" (namely, the first five chapters, after which a "brief summary of the salient points of the theory" are provided). In the preface to the second edition, the author contradicts his previous statement with the warning, "Do not use this book in the wrong way… It is advisable to study the first chapters I-V thoroughly." We tend to agree with the latter advice and believe that, as with any subject, there are no short cuts. Yet, we also believe that the background necessary to get started can be presented more efficiently than is currently available in the literature.

Besides the classics by Levi-Civita and Schouten, several other books exist at both the introductory[5–10] and advanced[11–16] levels (the books referenced are simply a sampling of those we consider worthwhile to an undergraduate student). The second barrier mentioned is related to notation. To become acquainted with the field, it is necessary to gain some virtuosity with the so-called "index gymnastics." Without this skill, it is not possible for anybody to follow the calculations of others or to perform their own. The origin of much difficulty resides, we believe, in a notation that is widely used in the journal literature, in specific books on tensors (certainly in all the books referenced above, with only three exceptions), and in most popular books on mathematical methods for physicists.[17] Being a tensor refers to a property possessed by a multi-component quantity, defined within a coordinate system, whose components transform in a specific way upon changing the coordinate system. The notation we are referring to here is that if the components of the multi-component quantity $\mathbf{A}$ are denoted by A_a in a particular coordinate system, those in another coordinate system are often denoted by, for instance, A'_a or $\bar{A}_a$—of the sixteen books quoted in Refs. 3–17, seven use the former notation and six use the latter. As reasonable as it might seem, such a practice is the origin of an unnecessary confusion, perfectly avoidable as long as the prime (or the bar) is attached to the index, such as in $A_{a'}$. This tiny change in notation might seem extravagant and unnatural but, as we shall see, it greatly simplifies the manipulations involved. Schouten (b. 1883) is seemingly the one who first introduced this notation so we will name it after him. What we propose in this paper is that Schouten's notation be widely adopted.

Only three authors among the ones quoted in Refs. 3–17 follow Schouten's notation: Jeevanjee, Lichnerowicz, and, of course, Schouten himself. However, Lichnerowicz uses the notation inappropriately and adds confusion to the matter. In Schouten's notation, a sum over the components of $\mathbf{A}$, which in the unprimed coordinate system is written as $\sum_{a=1}^{N} A_a$, in the primed coordinate system is correctly written as $\sum_{a=1}^{N} A_{a'}$ and not as $\sum_{a'=1}^{N} A_{a'}$, which is what Lichnerowicz (incorrectly) writes.

A common notation that we adopt is the *Einstein summation rule*. As we will see, we need two types of indices that we distinguish from each other by locating them either in a lower position, as in A_a or in an upper position, as in A^a. The Einstein summation rule is:

Rule 0: Whenever an index is repeated twice in a product—once in an upper position and once in a lower position—the index is called a *dummy index* and summation over it is implied. For instance, $A_{a'b'} B_c^{b'}$ means $\sum_b A_{a'b'} B_c^{b'}$ $= A_{a'1'} B_c^{1'} + A_{a'2'} B_c^{2'} + \cdots$, and C_a^a means $\sum_a C_a^a = C_1^1 + C_2^2 + \cdots$.

An index that is not a dummy index is called a *free index*. Dummy and free indices follow three rules that, as trivial as they might sound, are so crucial and used continuously that we better state them explicitly:

Rule 1: In the case of a multiple sum, different letters must denote the dummy indices even though one letter is primed and the other is not. For instance, we will not write $A_{a'a} B^{a'a}$ but rather $A_{a'b} B^{a'b}$ to imply a double sum over the two indices.

Rule 2: A dummy index may be renamed, at will and as convenience requires, within any single factor (as long there is no conflict with Rule 1).

Rule 3: Any free index must appear with the same name and position on each side of an equation; thereafter, if one wishes to rename it, it must be done throughout the *entire* equation.

The remainder of this paper is organized as follows. In Sec. II, we discuss the background assumed, and presumably shared, by undergraduate students in physics or engineering—vector calculus in a Cartesian coordinate system.[18] As we will see, there is no need to have upper and lower indices in such a coordinate system; all indices appear as lower indices. Hence, *only* in Sec. II, do we allow the Einstein summation rule to apply to each repeated index, even though located twice in a lower position.

In subsequent sections, we tie the background of Sec. II to the algebra and calculus of tensors. In Sec. III, we adopt an arbitrary basis (not necessarily orthonormal) and see how the dual basis naturally emerges, so that for each vector one has to deal with two types of components—those with respect to the original basis and those with respect to its dual. In Sec. IV, we see how the dual basis and the two types of components transform under a coordinate transformation, and we introduce the tensor-character concept. In Sec. V, we consider scalar, vector, and tensor fields, introduce the covariant derivative, and provide coordinate-independent expressions for the gradient and Laplacian of a scalar field and for the divergence and curl of a vector field. Finally, in Sec. VI, we tailor these expressions to general orthogonal coordinates, thereby completing our journey back to what our readers are likely to be familiar with.

II. A REVIEW: ORTHONORMAL BASES

Let $\{\hat{x}_a\,; a = 1, 2, 3\}$ be an orthonormal basis spanning the vectors of the ordinary Euclidean three-dimensional (3D)

space. The vectors may be associatively and commutatively summed, there is a unique zero-vector, and each vector has its unique opposite. The vectors may be associatively multiplied by real numbers, an operation that is distributive with respect to both the sum of real numbers and the sum of vectors.[19] Between any two vectors is defined a *dot* (or *scalar*) product, a commutative rule that associates a real number to each pair of vectors. This dot-product rule is distributive with respect to a linear combination of vectors and provides a non-negative number whenever a vector is dotted with itself (only the zero-vector dotted with itself gives the number zero).[20] A set of vectors $\{\mathbf{A}_n; n = 1, ..., N\}$ is said to be linearly independent if (note that a sum over n is implied)

$$C_n \mathbf{A}_n = 0 \;\Rightarrow\; C_n = 0 \;\forall n. \tag{1}$$

In general, if N is the maximum allowed number of linearly independent vectors, then the space is said to be N-dimensional and, in this case, N linearly independent vectors are said to form a *basis*—any vector may be written as a linear combination of the basis. Our ordinary space is Euclidean because the Pythagorean theorem holds, and it is 3D because the maximum value of N is $N = 3$; that is, there exist sets of three vectors $\{\mathbf{e}_a\,; a = 1, 2, 3\}$ such that *any* vector $\mathbf{A}$ may be written as a linear combination of the vectors $\mathbf{e}_a$.

In our ordinary 3D space a basis of vectors may always be chosen to be *orthonormal*. By this we mean that if $\{\hat{x}_a\,; a = 1, 2, 3\}$ is such a basis, their mutual dot products obey

$$\hat{x}_a \cdot \hat{x}_b = \delta_{ab}, \tag{2}$$

where δ_{ab} is the Kronecker delta, a quantity whose 9 components are equal to 0 if $a \neq b$ or equal to 1 if $a = b$. Each vector may be written as

$$\mathbf{A} = A_a \hat{x}_a, \tag{3}$$

where A_a is the ath component of the vector $\mathbf{A}$ in the basis $\{\hat{x}_a\}$. A direct consequence of Eqs. (2) and (3) and the stated properties of the dot product is that

$$A_a = \hat{x}_a \cdot \mathbf{A} \tag{4}$$

and

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} &= (A_a \hat{x}_a) \cdot (B_b \hat{x}_b) = (A_a B_b)(\hat{x}_a \cdot \hat{x}_b) \\ &= (A_a B_b)\delta_{ab} = A_a B_a. \end{aligned} \tag{5}$$

As done here, and in several calculations that follow, we shall redundantly collect some factors within parentheses with the aim of clarifying the algebra involved without the need for any other explanation.

In the 3D space, a *cross* (or *vector*) product is defined as

$$\mathbf{A} \times \mathbf{B} \equiv \det \begin{bmatrix} \hat{x}_1 & \hat{x}_2 & \hat{x}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{bmatrix} = \varepsilon_{abc} A_a B_b \hat{x}_c, \tag{6}$$

where ε_{abc} is the Levi-Civita symbol, a quantity whose 27 components are equal to $+1$ or -1 according to whether (a, b, c) forms an even or odd permutation of the sequence $(1, 2, 3)$, and equal to 0 otherwise (if two or more indices take on the same value).

A *field* is a function of the space coordinates $\mathbf{x} \equiv (x_1, x_2, x_3)$, and possibly of time as well, and may be a

scalar or a vector field depending on whether the function is a scalar or a vector. In this section, we denote the differential operator by $\partial_a \equiv \partial/\partial x_a$. Given a vector field $\mathbf{A} = \mathbf{A}(\mathbf{x}) = A_a(\mathbf{x})\hat{x}_a$, its *divergence* is the scalar field given by

$$\nabla \cdot \mathbf{A} = \partial_a A_a \qquad (7)$$

and its *curl* is the vector field given by

$$\nabla \times \mathbf{A} \equiv \det \begin{bmatrix} \hat{x}_1 & \hat{x}_2 & \hat{x}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ A_1 & A_2 & A_3 \end{bmatrix} = \varepsilon_{abc}\partial_a A_b \hat{x}_c. \qquad (8)$$

Given a scalar field $\phi(\mathbf{x})$, its *gradient* is the vector field defined by

$$\nabla\phi(\mathbf{x}) = \hat{x}_a \partial_a \phi, \qquad (9)$$

and its *Laplacian* is defined as the divergence of its gradient, namely

$$\Delta\phi \equiv \nabla^2\phi \equiv \nabla \cdot \nabla\phi = \partial_a \partial_a \phi. \qquad (10)$$

Why do we need scalars, vectors, and, in general, tensors? What we need in defining a physical quantity is for it to have some character of objectivity in the sense that it does not have to depend on the coordinate system used. For instance, a one-component function of position, such as a temperature field specifying the temperature T at each point $\mathbf{P}$ of space, must provide a *unique* value T when the point $\mathbf{P}$ is considered, regardless of the coordinate system used.

Not every quantity specified by a single number is a scalar. For example, in our 3D space whose points are parameterized by a Cartesian coordinate system, all *displacements* from a given point, such as the origin, will be specified by three quantities (x_1, x_2, x_3). Now, the displacement that in one coordinate system is specified, for instance, as $(1, 0, 0)$, would be specified as $(0, 1, 0)$ in a coordinate system rotated clockwise by $\pi/2$ around the third axis. Thus, *each component* of the displacement is indeed a single-component quantity, yet these components depend on the coordinate system used (they are not scalars) and do not have that objectivity character we desire. An objectivity-preserving, single-component quantity must transform in a specific way under a coordinate transformation—it must be an invariant.

Likewise, multi-component quantities must have an objectivity character (and are thereby called vectors and tensors), a circumstance that translates into specific rules about how their components must transform in order to preserve that objectivity. As not all single-component quantities are scalars, similarly, not all multi-component quantities are vectors or tensors. We will specify the transformation rules in Sec. IV. For the time being, we only remark that the above definitions for dot and cross products, divergence, curl, and gradient are what they are because they assure that the objectivity we are requiring is preserved. For instance, we will see that the above-defined displacements are vectors, and the dot product [as defined in Eq. (5)] of each displacement with itself is a scalar. Had we defined, for instance, the single-component quantity $(x_1)^2 - (x_2)^2 + (x_3)^2$, we would not have had an invariant. (For the displacement example considered in the previous paragraph, this quantity is equal to $+1$ in the first coordinate system and -1 in the other.)

Another reason to require in physics the objectivity we are considering here is that if we believe our space to be homogeneous and isotropic (invariant under translations and rotations) and if we believe in the law of inertia (that invariance under a Lorentz transformation holds), then the *fundamental* laws of physics must be accordingly invariant. In other words, they must be expressed by equations involving single- and multi-component quantities that under translation, rotation, and Lorentz transformations preserve the objectivity character we are after. Quantities with such a character and defined in an N-dimensional space are called *tensors* and, more specifically, rth-rank tensors if they have N^r components. Scalars, being single-component quantities, are then zero-rank tensors, while vectors, being N-component quantities, are first-rank tensors. Higher-rank tensors are…well, tensors!

To see how higher-rank quantities naturally emerge, a classic example is the case of the stress tensor defined on each point of a solid body. We consider a small volume V surrounding a point within the solid and let F_a denote the ath Cartesian component of the total internal stress force in V. This force may be due, for instance, to external forces being applied to the surface of the body, or because V feels on its surface the gravitational force due to the rest of the surrounding body. Whatever its cause, the internal stress force can be written $F_a = \int_V dV \phi_a$, where ϕ_a is the ath Cartesian component of the stress force density. Stress forces act within a few molecular layers (their action range is negligible when compared to the linear dimensions of the body) so they act through the surface S enclosing V. The above volume integral must then be expressible as a surface integral over S, and making use of the divergence (Gauss's) theorem, we have $F_a = \int_V \partial_b \sigma_{ab} dV = \int_S \sigma_{ab} n_b ds$, where n_b is the bth component of the unit vector perpendicular to the surface element ds of S, $\phi_a = \partial_b \sigma_{ab}$, and σ_{ab} is a 9-component quantity called the *stress tensor*.

III. ARBITRARY BASES: DUALITY

For reasons that will soon become apparent, for the remainder of the paper, we apply the Einstein summation rule as stated in Sec. I: summation over a repeated index is implied *only* if it appears once in an upper and once in a lower position. Let us now see what happens as soon as we relax our very first assumption, expressed in Eq. (2), according to which the basis used is orthonormal. In contrast to Eq. (2), we then consider a set of basis vectors $\{\mathbf{e}_a\, ; a = 1, 2, 3\}$ for which

$$g_{ab} \equiv \mathbf{e}_a \cdot \mathbf{e}_b \neq \delta_{ab}. \qquad (11)$$

It is appropriate to remark at this point that the (symmetric) matrix $[g_{ab}]$, whose elements are g_{ab}, has the determinant $G \equiv \det[g_{ab}] \neq 0$. In fact, given $\{\mathbf{e}_a\}$ as a basis its vectors are linearly independent [Eq. (1) holds] so that

$$C^a \mathbf{e}_a = 0 \;\Rightarrow\; C^a = 0 \;\forall a. \qquad (12)$$

But this equation implies that

$$(C^a \mathbf{e}_a) \cdot \mathbf{e}_b = C^a(\mathbf{e}_a \cdot \mathbf{e}_b) = C^a g_{ab} = 0 \;\Rightarrow\; C^a = 0 \;\forall a. \qquad (13)$$

However, a homogeneous linear equation such as $C^a g_{ab} = 0$ possesses the trivial solution ($C^a = 0$) if and only if $\det[g_{ab}] \neq 0$. (Incidentally, this implies that $[g_{ab}]$ is a

non-singular matrix; in other words, it admits an inverse.) We shall denote the elements of the inverse matrix by g^{ab}, so that

$$[g_{ab}]^{-1} \equiv [g^{ab}], \tag{14}$$

which is equivalent to writing

$$[g_{ab}][g^{ab}] = \mathbf{I} \quad \text{or} \quad g_{ac}g^{cb} = \delta_a^b, \tag{15}$$

where $\mathbf{I}$ is the identity matrix and the indices in the Kronecker delta have been written according to our Rule 3. Clearly, $[g^{ab}]$ is also symmetric and $\det[g^{ab}] = 1/\det[g_{ab}] \equiv 1/G$, a result that follows from Eq. (15) and the fact that the determinant of a product of matrices is equal to the product of their determinants.

Any vector can be expressed as a linear combination of the basis $\{\mathbf{e}_a\}$ so that

$$\mathbf{A} = A^a \mathbf{e}_a, \tag{16}$$

which is equivalent to Eq. (3) except for the change in notation, where the components of the vector $\mathbf{A}$ in the basis $\{\mathbf{e}_a\}$ have been labelled with an *upper* index. However, if one asks if $A^a = \mathbf{e}_a \cdot \mathbf{A}$ or $\mathbf{A} \cdot \mathbf{B} = \sum_a A^a B^a$ still hold as in Eqs. (4) and (5), the answer is negative:

$$\mathbf{e}_a \cdot \mathbf{A} = \mathbf{e}_a \cdot (A^b \mathbf{e}_b) = (\mathbf{e}_a \cdot \mathbf{e}_b) A^b = g_{ab} A^b \neq \delta_b^a A^b = A^a \tag{17}$$

and

$$\mathbf{A} \cdot \mathbf{B} = (A^a \mathbf{e}_a) \cdot (B^b \mathbf{e}_b) = (\mathbf{e}_a \cdot \mathbf{e}_b) A^a B^b$$
$$= g_{ab} A^a B^b \neq \delta_{ab} A^a B^b = \sum_a A^a B^a. \tag{18}$$

In conclusion, $A^a \neq \mathbf{e}_a \cdot \mathbf{A}$ and $\mathbf{A} \cdot \mathbf{B} \neq \sum_a A^a B^a$. Please note that stating $A^a \neq \mathbf{e}_a \cdot \mathbf{A}$ is in keeping with Rule 3. In some sense, Rule 3 warns us that this cannot be a *bona fide* equality. The reader might wonder whether this is a consequence of the adopted rule and our having arbitrarily set the labeling index in the upper position. This is not the case. The inequality in Eq. (17) says that the ath component of a vector $\mathbf{A}$ in a given basis $\{\mathbf{e}_a\}$ is *not* given by the dot product of $\mathbf{A}$ with $\mathbf{e}_a$, no matter where we choose to locate the index to label that component; Rule 3 simply certifies that this is so.

A natural question to ask is whether there exists some *other* basis, which we provisionally denote by $\{\mathbf{e}^a\}$, such that

$$A^a = \mathbf{e}^a \cdot \mathbf{A}. \tag{19}$$

Although this provisional notation is motivated by our desire to follow Rule 3, we will see that such a basis does exist and is *unique*. Moreover, if we call $\{\mathbf{e}^a\}$ the *dual basis* of the original basis $\{\mathbf{e}_a\}$, it turns out that the dual of $\{\mathbf{e}^a\}$ is simply $\{\mathbf{e}_a\}$. To find such a basis, we note that Eqs. (16) and (19) lead to

$$A^a = \mathbf{e}^a \cdot \mathbf{A} = \mathbf{e}^a \cdot (A^b \mathbf{e}_b) = (\mathbf{e}^a \cdot \mathbf{e}_b) A^b, \tag{20}$$

which holds provided the vectors $\{\mathbf{e}^a\}$ are solutions of the equation

$$\mathbf{e}^a \cdot \mathbf{e}_b = \delta_b^a. \tag{21}$$

If the vectors $\{\mathbf{e}^a\}$ exist they must be expressible as a linear combination of the basis $\{\mathbf{e}_a\}$. Therefore, $\mathbf{e}^a = C^{ac} \mathbf{e}_c$ and the left-hand side of Eq. (21) becomes $(C^{ac} \mathbf{e}_c) \cdot \mathbf{e}_b = C^{ac} (\mathbf{e}_c \cdot \mathbf{e}_b) = C^{ac} g_{cb}$, which means

$$C^{ac} g_{cb} = \delta_b^a. \tag{22}$$

Because the matrix $[g_{ab}]$ is non-singular, Eq. (22) has the solution $[C^{ab}] = [g_{ab}]^{-1} = [g^{ab}]$, which means that the unique solution of Eq. (21) is given by

$$\mathbf{e}^a = g^{ab} \mathbf{e}_b. \tag{23}$$

The basis $\{\mathbf{e}^a\}$ is called the *dual basis* of $\{\mathbf{e}_a\}$ and is obtained from the latter by transforming it via the matrix *inverse* of $[\mathbf{e}_a \cdot \mathbf{e}_b]$. Similarly, the dual basis of the basis $\{\mathbf{e}^a\}$ is obtained from the latter by transforming it via the matrix *inverse* of $[\mathbf{e}^a \cdot \mathbf{e}^b]$. To understand this, we note that Eqs. (23) and (21) lead to

$$\mathbf{e}^a \cdot \mathbf{e}^b = (g^{ac} \mathbf{e}_c) \cdot \mathbf{e}^b = g^{ac} (\mathbf{e}_c \cdot \mathbf{e}^b) = g^{ac} \delta_c^b = g^{ab}, \tag{24}$$

so that

$$g^{ab} = \mathbf{e}^a \cdot \mathbf{e}^b \tag{25}$$

whereby $[\mathbf{e}^a \cdot \mathbf{e}^b]^{-1} \equiv [g^{ab}]^{-1} = [g_{ab}]$. Thus, the vector dual to $\mathbf{e}^a$ is

$$g_{ab} \mathbf{e}^b = g_{ab} (g^{bc} \mathbf{e}_c) = (g_{ab} g^{bc}) \mathbf{e}_c = \delta_a^c \mathbf{e}_c = \mathbf{e}_a. \tag{26}$$

In other words, if the basis dual to $\{\mathbf{e}_a\}$ is $\{\mathbf{e}^a\}$, then the basis dual to the latter is $\{\mathbf{e}_a\}$:

$$\{\mathbf{e}^a = g^{ab} \mathbf{e}_b\} \iff \{\mathbf{e}_a = g_{ab} \mathbf{e}^b\}. \tag{27}$$

We therefore see that once a non-orthonormal basis is considered, another basis—its dual—naturally emerges. Accordingly, any vector $\mathbf{A}$ may then be written as

$$\mathbf{A} = A^b \mathbf{e}_b = A_b \mathbf{e}^b, \tag{28}$$

where the components A^a can be found by dotting $\mathbf{e}^a$ with Eq. (28) and using Eq. (21) to get

$$A^a = \mathbf{A} \cdot \mathbf{e}^a. \tag{29}$$

Similarly, the components A_a can be found by dotting $\mathbf{e}_a$ with Eq. (28) to get

$$A_a = \mathbf{A} \cdot \mathbf{e}_a. \tag{30}$$

For reasons that will become apparent in the following section, the components labeled by an *upper* index are called the *contravariant* components of $\mathbf{A}$, and those labeled by a *lower* index are called the *covariant* components of $\mathbf{A}$. A relation between the contravariant and covariant components may be readily obtained from the equality $A^b \mathbf{e}_b = A_b \mathbf{e}^b$: dotting it with either $\mathbf{e}_a$ or $\mathbf{e}^a$ and using Eq. (21) together with either Eq. (11) or Eq. (25) gives

$$A_a = g_{ab} A^b \tag{31}$$

and

$$A^a = g^{ab} A_b. \tag{32}$$

What about the dot product between two vectors? Using Eqs. (28), (21), (11), and (25), we now have four ways to write it [compare with Eq. (5)]:

$$\mathbf{A} \cdot \mathbf{B} = g_{ab}A^a B^b = g^{ab}A_a B_b = A^a B_a = A_a B^a. \tag{33}$$

If we consider the two sets of components as vectors of two different spaces, one dual of the other, we would say that the dot product is actually performed between the vectors of a space and those of its dual. This distinction is hidden when an orthonormal basis is used because $g_{ab} = \delta_{ab}$ and from Eqs. (27), (31), and (32) we see that such a basis is self-dual so the contravariant and covariant components coincide; this is why there was no need in Sec. II to distinguish between them by using upper and lower indices.

We notice here that in our 3D space, a ready way to obtain the dual vectors $\mathbf{e}^a$ of a given basis $\{\mathbf{e}_a\}$ is by using the relation

$$\mathbf{e}^a = \frac{\mathbf{e}_b \times \mathbf{e}_c}{V}. \tag{34}$$

Here (a, b, c) form an even permutation of the triplet $(1, 2, 3)$ and $V = \mathbf{e}_a \cdot (\mathbf{e}_b \times \mathbf{e}_c)$ is the volume of the parallelepiped spanned by the vectors of $\{\mathbf{e}_a\}$. Indeed, the vectors defined by Eq. (34) satisfy Eq. (21), whose unique solution guarantees that these vectors are the wanted duals.[21]

To find the contravariant and covariant components of the cross product of two vectors, let us first show that for *any* given six vectors, which shall here be denoted as $\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3$, $\mathbf{T}_1, \mathbf{T}_2$, and $\mathbf{T}_3$, one has

$$(\mathbf{S}_1 \cdot \mathbf{S}_2 \times \mathbf{S}_3)(\mathbf{T}_1 \cdot \mathbf{T}_2 \times \mathbf{T}_3) = \det[\mathbf{S}_a \cdot \mathbf{T}_b]. \tag{35}$$

Using Eqs. (5) and (6), we see that the left-hand side becomes $\det[S_{aA}]\det[T_{bB}] = \det[S_{aA}]\det[T_{Bb}] = \det[S_{aA}T_{Ab}]$, thus proving the assertion. Here, S_{aA} is the Ath component of the vector $\mathbf{S}_a$ (with similar meaning for T_{aA}) and we have made use of the facts that determinants are invariant when rows and columns are interchanged (first equality) and the determinant of a product of matrices is equal to the product of their determinants (second equality). Applying Eq. (35) to the basis $\{\mathbf{e}_a\}$ we get

$$V^2 = (\mathbf{e}_1 \cdot \mathbf{e}_2 \times \mathbf{e}_3)^2 = \det[\mathbf{e}_a \cdot \mathbf{e}_b] = \det[g_{ab}] \equiv G, \tag{36}$$

where again, $V = \mathbf{e}_a \cdot (\mathbf{e}_b \times \mathbf{e}_c)$ is the volume of the parallelepiped spanned by the basis vectors $\mathbf{e}_a$. To avoid any unlikely confusion between an upper index and a power exponent, we have inserted within parentheses all quantities raised to a power. We note that

$$(\mathbf{e}^1 \cdot \mathbf{e}^2 \times \mathbf{e}^3)^2 = \det[\mathbf{e}^a \cdot \mathbf{e}^b] = \det[g^{ab}] = 1/G, \tag{37}$$

due to the fact that $\det[g^{ab}] = 1/\det[g_{ab}]$. Armed with Eqs. (36)–(37) and (28)–(30), we are now ready to give an expression for the cross product of two vectors $\mathbf{C} = \mathbf{A} \times \mathbf{B}$. For the covariant components, we have

$$C_c = \mathbf{C} \cdot \mathbf{e}_c = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{e}_c = A^a B^b(\mathbf{e}_a \times \mathbf{e}_b) \cdot \mathbf{e}_c$$
$$= V\varepsilon_{abc}A^a B^b = \sqrt{G}\varepsilon_{abc}A^a B^b, \tag{38}$$

and similarly, for the contravariant components, we have

$$C^c = \mathbf{C} \cdot \mathbf{e}^c = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{e}^c = A_a B_b(\mathbf{e}^a \times \mathbf{e}^b) \cdot \mathbf{e}^c$$
$$= \frac{1}{\sqrt{G}}\varepsilon^{abc}A_a B_b, \tag{39}$$

where $\varepsilon^{abc} = \varepsilon_{abc}$ with the indices located to be consistent with our Rule 0. In conclusion we have

$$\mathbf{A} \times \mathbf{B} = \sqrt{G}\varepsilon_{abc}A^a B^b\mathbf{e}^c = \frac{1}{\sqrt{G}}\varepsilon^{abc}A_a B_b\mathbf{e}_c. \tag{40}$$

From Eqs. (11) and (25), we see that the quantities g_{ab} and g^{ab} depend only on the basis vectors (the way the space has been parameterized). In addition, from Eqs. (27), (31), and (32), we see that these quantities determine the dual basis from the original basis and relate the covariant and contravariant components to each other or, as usually phrased, that g_{ab} and g^{ab} allow the lowering and rising of the indices. Finally, these quantities determine the dot product of two vectors [Eq. (33)] and, via $G \equiv \det[g_{ab}]$, their cross product as well [Eq. (40)]. Because of the central role played by the quantities g_{ab} and g^{ab}, they have been given a special name: the *metric*. In Sec. IV, we will see that the metric is indeed a second-rank tensor.

IV. CHANGING BASES: TENSORS

We now wish to move from a given basis $\{\mathbf{e}_a\}$ to another basis that, in keeping with Shouten's notation, will be denoted as $\{\mathbf{e}_{a'}\}$. From the requirement that $\{\mathbf{e}_{a'}\}$ be a basis, it follows that each vector $\mathbf{e}_a$ of the unprimed basis can be written as a linear combination of the primed basis vectors

$$\mathbf{e}_a = R_a^{b'}\mathbf{e}_{b'} \tag{41}$$

and, given $\{\mathbf{e}_a\}$ as a basis, each vector $\mathbf{e}_{a'}$ of the primed basis can be written as a linear combination of the unprimed basis vectors

$$\mathbf{e}_{a'} = R_{a'}^{b}\mathbf{e}_b. \tag{42}$$

This equation acts as the definition of the primed basis. Note that we might think of $R_a^{b'}$ (which is a 2-index quantity) as the elements of a matrix. If so, we have to decide which index refers to the rows and which to the columns; we freely opt for the lower index for the rows and the upper for the columns.

Inserting Eq. (42) into (41), we get $\mathbf{e}_a = R_a^{b'}(R_{b'}^{c}\mathbf{e}_c)$ $= (R_a^{b'}R_{b'}^{c})\mathbf{e}_c$, whereby

$$R_a^{b'}R_{b'}^{c} = \delta_a^c, \quad \text{or} \quad [R_a^{b'}][R_{a'}^{b}] = \mathbf{I}. \tag{43}$$

The matrix $[R_a^{b'}]$ is invertible and its inverse is the matrix $[R_{a'}^{b}]$. It is crucial at this point to have no misunderstandings about the notation: $R_a^{b'}$ and $R_{a'}^{b}$ are matrix elements of *different* matrices—in fact, they are inverse to each other—and it is the position of the prime (on the lower or on the upper index) that tells us that this is so. In particular, $\det[R_a^{b'}] \neq \det[R_{a'}^{b}]$ and these nonzero determinants are reciprocal to each other.

Any vector can now be written as

$$\mathbf{A} \equiv A_a\mathbf{e}^a \equiv A_{a'}\mathbf{e}^{a'} \equiv A^{a'}\mathbf{e}_{a'} \equiv A^a\mathbf{e}_a, \tag{44}$$

where the equalities have been denoted as identities to stress the objectivity character of the quantity we are dealing with. Our task now is to find how to go from the direct and dual bases (contravariant and covariant components) in the unprimed system to the direct and dual bases (contravariant and covariant components) in the primed system, and *vice versa*. Equation (42) is the starting point as it defines the primed basis (we assume the matrix $\mathbf{R} \equiv [R_{a'}^b]$ is given). Equation (41) gives the unprimed direct basis from the primed direct basis (requiring inversion of the matrix $\mathbf{R}$).

The metric of the primed system is

$$g_{a'b'} \equiv \mathbf{e}_{a'} \cdot \mathbf{e}_{b'} = (R_{a'}^c \mathbf{e}_c) \cdot (R_{b'}^d \mathbf{e}_d) = R_{a'}^c R_{b'}^d (\mathbf{e}_c \cdot \mathbf{e}_d), \quad (45)$$

or

$$g_{a'b'} = R_{a'}^c R_{b'}^d g_{cd}. \quad (46)$$

From Eq. (41) and the last identity in Eq. (44), and making a judicious use of our rules, we have

$$A^{a'} \mathbf{e}_{a'} \equiv A^a \mathbf{e}_a = A^a (R_a^{b'} \mathbf{e}_{b'}) = (R_a^{b'} A^a) \mathbf{e}_{b'} = (R_b^{a'} A^b) \mathbf{e}_{a'}, \quad (47)$$

whereby

$$A^{a'} = R_b^{a'} A^b, \quad (48)$$

which are the primed contravariant components from the unprimed ones. To obtain the primed covariant components from the unprimed ones, we may lower the index in the contravariant components by means of the metric. From Eqs. (31), (43), (46), and (48), we obtain

$$A_{a'} = g_{a'c'} A^{c'} = (R_{a'}^b R_{c'}^d g_{bd})(R_e^{c'} A^e) = R_{a'}^b g_{bd}(R_{c'}^d R_e^{c'}) A^e$$
$$= R_{a'}^b g_{bd}(\delta_e^d A^e) = R_{a'}^b (g_{bd} A^d), \quad (49)$$

or

$$A_{a'} = R_{a'}^b A_b. \quad (50)$$

Finally, we want to obtain the primed dual vectors from the unprimed ones, which we find as

$$\mathbf{e}^{a'} = g^{a'c'} \mathbf{e}_{c'} = (R_b^{a'} R_d^{c'} g^{bd})(R_{c'}^e \mathbf{e}_e) = R_b^{a'} g^{bd}(R_d^{c'} R_{c'}^e) \mathbf{e}_e$$
$$= R_b^{a'} g^{bd}(\delta_d^e \mathbf{e}_e) = R_b^{a'}(g^{bd} \mathbf{e}_d), \quad (51)$$

or

$$\mathbf{e}^{a'} = R_b^{a'} \mathbf{e}^b. \quad (52)$$

From this and Eq. (25) we see that

$$g^{a'b'} = R_c^{a'} R_d^{b'} g^{cd}. \quad (53)$$

Let us now look at Eqs. (42), (48), (50), and (52). We see that the covariant components transform exactly as does the direct basis—compare Eqs. (50) and (42), where the transformation is governed by the matrix $\mathbf{R} \equiv [R_{a'}^b]$—whereas the contravariant-component transformation (48) is governed by the matrix $\mathbf{R}^{-1} \equiv [R_a^{a'}]$. These components transform in a manner "contrary" (so to speak) to that adopted by the direct

basis; in fact, they transform as does the dual basis as can be seen by comparing Eqs. (48) and (52). This is the origin of the names *covariant* and *contravariant* given to the components labeled with indices in, respectively, the lower and upper positions.

The dot product defined in Eq. (33) is indeed a scalar (an invariant). For instance, from the last form of Eq. (33), we have

$$A_{a'} B^{a'} = (R_{a'}^c A_c)(R_d^{a'} B^d) = (R_{a'}^c R_d^{a'}) A_c B^d$$
$$= \delta_d^c A_c B^d = A_c B^c = A_a B^a. \quad (54)$$

Thus, when going from one basis to another in our 3D space as dictated by the transformation (42), scalar quantities are, by definition, invariant. Or, if we like, quantities that are invariant under transformation (42) are legitimate one-component physical quantities. On the other hand, the components of a legitimate 3-component physical quantity (vector) must transform according to either Eq. (48) (contravariant components) or Eq. (50) (covariant components).

Composing two vectors as in $A_a B^a$ has yielded a scalar, but two vectors might be composed as in $A_a B^b$, $A_a B_b$ or $A^a B^b$, producing 9-component quantities whose transformation rules, readily obtained from those of the single vectors, guarantee the objectivity character we are requiring for a physical quantity. However, such higher-rank quantities need not be obtained by composing two vectors. We are then led to define a *second-rank tensor* as a 9-component quantity $\mathbf{T}$ whose covariant, contravariant, and mixed components transform, respectively, as

$$T_{a'b'} = R_{a'}^c R_{b'}^d T_{cd}, \quad (55)$$

$$T^{a'b'} = R_c^{a'} R_d^{b'} T^{cd}, \quad (56)$$

$$T_{a'}^{b'} = R_{a'}^c R_d^{b'} T_c^d. \quad (57)$$

Equations (55)–(57) may also be considered as the defining relations of, respectively, a $(0, 2)$-type, $(2, 0)$-type, and $(1, 1)$-type second-rank tensor. Contravariant and covariant vectors are $(1, 0)$-type and $(0, 1)$-type first-rank tensors, and scalars are $(0, 0)$-type zero-rank tensors. Meanwhile, (p, q)-type tensors of rank $(p + q)$ may be easily defined by generalizing the relations (55)–(57).

From Eqs. (46) and (53) we see that g_{ab} and g^{ab} are the covariant and contravariant components of a second-rank tensor, the *metric tensor*. We can obtain its mixed components by either lowering one index of g^{ab} or raising one index of g_{ab}; in both cases, we have $g^{ca} g_{ab} = \delta_b^c$, showing that the metric-tensor mixed components are given by the Kronecker delta δ_a^b. To show that δ_a^b is indeed a tensor we should check that, according to Eq. (57), we have $\delta_{a'}^{b'} = R_{a'}^c R_d^{b'} \delta_c^d$. Sure enough, we find that $R_{a'}^c (R_d^{b'} \delta_c^d) = R_{a'}^c R_c^{b'} = \delta_{a'}^{b'}$. This result also shows that the Kronecker delta δ_a^b has the same components in all coordinate systems, a property not shared by either δ_{ab} or δ^{ab}. For instance, $R_{a'}^c R_{b'}^d \delta_{cd} = \sum_c R_{a'}^c R_{b'}^c$, which, in general, is *not* equal to $\delta_{a'b'}$ unless the transformation matrix $\mathbf{R} \equiv [R_{a'}^b]$ happens to be orthogonal (its inverse is equal to its transpose $\mathbf{R}^{-1} = \mathbf{R}^T$).

In an obvious way, a tensor may be multiplied by a real number by multiplying each component by the number, and tensors of the same type may be added by summing their

homonymous components. Tensors of any type may be multiplied: the *outer product* of a tensor of type (p, q) by a tensor of type (s, t) produces a tensor of type $(p + s, q + t)$. From a tensor of type (p, q), we can get another tensor, of type $(p - 1, q - 1)$, by setting two indices—one covariant and the other contravariant—equal to each other, thereby summing over the resulting repeated index; this operation is called *contraction*. The outer product of two tensors followed by a contraction is an *inner product*, and a fully contracted inner product of two equal-rank tensors can be called a *scalar product*, an example of which is the dot product in Eq. (33). That the operations just defined produce new tensors is easy to check from the defining transformation properties (55)–(57).

It is possible to test the tensor character of a multi-component quantity also by using the *quotient rule*. To explain this rule and how it works, let us start with a certain notation. Given two tensors $\mathbf{T}_1$ and $\mathbf{T}_2$, let $\mathbf{T}_1 * \mathbf{T}_2$ denote any resulting tensor arising from taking their outer product, possibly followed by a contraction. Then let $\mathbf{X}$ be a multi-component quantity whose tensor character (if any) we wish to explore. The quotient rule says the following. By treating $\mathbf{X}$ provisionally as a tensor and considering a specific product $\mathbf{X} * \mathbf{A} \equiv \mathbf{T}$ between the components of $\mathbf{X}$ and $\mathbf{A}$, as defined above (i.e., a product that does not violate our Rules 0-3), if it can be shown that *for any* tensor $\mathbf{A}$ the quantity $\mathbf{T}$ is also a tensor, then $\mathbf{X}$ is a tensor.

Before seeing how the rule arises, we can readily apply it. Suppose we do not know that the metric was a tensor, but we do know that it is a 2-index quantity such that, *for any* tensor A^b, Eq. (31) holds so that $g_{ab}A^b$ is a tensor. Then by the quotient rule, g_{ab} are the covariant components of a tensor or, in short, g_{ab} is a tensor. Similarly, because Eq. (32) holds for any tensor A_b, then g^{ab} is also a tensor. Moreover, because we know that for any tensor A^a the quantity $\delta_a^b A^a = A^b$ is also a tensor, we can deduce that δ_a^b is a tensor. On the other hand, to show from the quotient rule that δ_{ab} is a tensor we would need to show that, for instance, for any tensor A^a the quantity $\delta_{ab}A^a$ also is a tensor T_b (in this case, a covariant vector). However, from the definition of δ_{ab}, we have $\delta_{ab}A^a \neq A_b$ (as previously discussed, the equality would hold only within orthonormal bases). Therefore, δ_{ab} is not a tensor.

The quotient rule should be applied with some care; the key words in the rule's statement are *for any*. Instead of proving the rule, we shall see how the proof would proceed in a case where it is not legitimate to invoke it, thereby catching with one token both the reasoning to follow for a proof and its pitfalls. Say we have proved that $X_{ab}A^a A^b \equiv T$ is a scalar (a zero-rank tensor) *for any* tensor A^a. Then

$$X_{a'b'}A^{a'}A^{b'} = X_{cd}A^c A^d = X_{cd}(R_{a'}^c A^{a'})(R_{b'}^d A^{b'})$$
$$= R_{a'}^c R_{b'}^d X_{cd}A^{a'}A^{b'}, \tag{58}$$

where in the second equality we have used the tensor nature of A^a. From the first and last terms of this result, it follows that

$$(X_{a'b'} - R_{a'}^c R_{b'}^d X_{cd})A^{a'}A^{b'} = 0, \tag{59}$$

supposedly true *for any* A^a. This circumstance, however, does not allow us to infer that $X_{a'b'} - R_{a'}^c R_{b'}^d X_{cd} = 0$, i.e., that

X_{ab} is a tensor. This would be true if we had had an equation such as $(X_{a'b'} - R_{a'}^c R_{b'}^d X_{cd})A^{a'b'} = 0$ holding for any A^{ab}. But $A^a A^b$ is not *any* tensor—it is a special (a symmetric) tensor. Therefore, from Eq. (59), we cannot infer that X_{ab} is a tensor.

A tensor is *symmetric* if it is invariant under the exchange of two *equal-variance* (both upper or lower) indices, such as $T_{ab} = T_{ba}$, whereas a tensor is *antisymmetric* if it changes sign under the exchange of two equal-variance indices, such as $T_{ab} = -T_{ba}$. The importance of these properties is due to the fact that they hold in *any* coordinate system. If $T_{ab} = \pm T_{ba}$ then

$$T_{a'b'} = R_{a'}^c R_{b'}^d T_{cd} = \pm R_{a'}^c R_{b'}^d T_{dc} = \pm R_{b'}^d R_{a'}^c T_{dc}$$
$$= \pm T_{b'a'}, \tag{60}$$

and similarly for two contravariant indices. From any tensor with two equal-variance indices, such as T_{ab}, we may construct a symmetric and an antisymmetric tensor, namely $T_{ab} \pm T_{ba}$. Likewise, any such tensor may be written as a sum of a symmetric and an antisymmetric tensor as $T_{ab} = \frac{1}{2}[(T_{ab} + T_{ba}) + (T_{ab} - T_{ba})]$. Symmetry/antisymmetry is not defined for two indices of different variance because in this case the property would not be coordinate-independent.

V. FIELDS

A *field* is a function of the space coordinates (and possibly of time as well). To represent an objectivity-preserving physical quantity the field function must have a tensorial character. Specifically, in our ordinary 3D space, it must be a 3^r-component tensor field of rank r, with $r = 0, 1, 2, ...$: a $(0, 0)$-type tensor (scalar field), $(1, 0)$-type tensor (contravariant vector field), $(0, 1)$-type tensor (covariant vector field) or, in general, a (p, q)-type tensor field of rank $r = p + q$.

Let us then consider our ordinary Euclidean 3D space parameterized by a set of *arbitrary* coordinates, which we shall denote by $\mathbf{x} \equiv (x^1, x^2, x^3)$. We assume that they are related to the Cartesian coordinates—hereafter denoted by $\mathbf{x}' \equiv (x^{1'}, x^{2'}, x^{3'})$—by an invertible relation, so that

$$x^{a'} = x^{a'}(x^1, x^2, x^3) \equiv x^{a'}(\mathbf{x}) \tag{61a}$$

and

$$x^a = x^a(x^{1'}, x^{2'}, x^{3'}) \equiv x^a(\mathbf{x}'). \tag{61b}$$

The Jacobian determinants are nonzero

$$J \equiv \det\left[\frac{\partial x^{a'}}{\partial x^b}\right] \equiv \det[R_b^{a'}] \neq 0 \tag{62a}$$

and

$$J^{-1} \equiv \det\left[\frac{\partial x^a}{\partial x^{b'}}\right] \equiv \det[R_{b'}^a] \neq 0, \tag{62b}$$

where we have set

$$\frac{\partial x^{a'}}{\partial x^b} \equiv R_b^{a'} \text{ and } \frac{\partial x^a}{\partial x^{b'}} \equiv R_{b'}^a. \tag{63}$$

These are elements of matrices inverse to each other

$$R^a_{b'}R^{b'}_c = \sum_b \frac{\partial x^a}{\partial x^{b'}}\frac{\partial x^{b'}}{\partial x^c} = \frac{\partial x^a}{\partial x^c} = \delta^a_c, \qquad (64)$$

which is why we wrote the Jacobian as J^{-1} in Eq. (62b).

We now proceed to show how two bases of vectors emerge from Eqs. (61). From Eq. (61a), upon fixing *one* coordinate at a time, three surfaces (the *coordinate surfaces*) are specified; by fixing $x^a = x^a_0$, say, one specifies the surface of the points of space with constant x^a, whose parametric equations are given by Eq. (61a) upon setting $x^a = x^a_0$ (the other two coordinates play the role of parameters). For instance, the coordinate surfaces arising by parameterizing the space with spherical polar coordinates are spheres centered at the origin, cones with the vertex at the origin, and half-planes originating from the z-axis.

Similarly, upon fixing *two* coordinates at a time, $x^a = x^a_0$ and $x^b = x^b_0$, say, one specifies the points lying on both the ath and bth coordinate surfaces, i.e., the points lying on their intersection curve—a *coordinate curve*—whose parametric equations are given by Eq. (61a) upon setting $x^a = x^a_0$ and $x^b = x^b_0$ (the remaining coordinate plays the role of the parameter).

Of course, upon fixing all three coordinates a single point of space is specified. Each point is at the intersection of three coordinate curves, whose tangent vectors at that point—readily evaluated from the parametric equation of the curves—are

$$\mathbf{e}_a = \sum_b \frac{\partial x^{b'}}{\partial x^a}\mathbf{e}_{b'} = R^{b'}_a \mathbf{e}_{b'}. \qquad (65)$$

Here $\mathbf{e}_{b'}$ is the bth Cartesian basis vector and Eq. (63) has been used. Recall that, being orthonormal, a Cartesian basis is self-dual ($\mathbf{e}_{b'} = \mathbf{e}^{b'}$) and it is precisely this fact that allows us to write Eq. (65) in a manner suitable to make use of our Rule 0.

Are the vectors $\{\mathbf{e}_a\}$ in Eq. (65) a basis? That is, does $C^a\mathbf{e}_a = 0$ imply $C^a = 0$, $\forall a$? The answer is yes; the orthonormal vectors $\{\mathbf{e}_{a'}\}$ are a basis, whereby $C^a(R^{b'}_a \mathbf{e}_{b'}) = (C^a R^{b'}_a)\mathbf{e}_{b'} = 0$ implies $C^a R^{b'}_a = 0$, $\forall b$. However, $C^a R^{b'}_a = 0$ implies $C^a = 0, \forall a$ because, by Eq. (62a), $\det[R^{b'}_a] \neq 0$.

As seen above, the matrices whose elements are given in Eq. (63) are inverses of each other, and using Eqs. (65), (41), and (52), we could at once write the vectors dual to those in Eq. (65). However, it is instructive to see how they emerge from Eq. (61b) by computing, from each of them, the gradient vectors as prescribed by Eq. (9)

$$\mathbf{e}^a \equiv \nabla x^a = \sum_b \frac{\partial x^a}{\partial x^{b'}}\mathbf{e}_{b'} = R^a_{b'}\mathbf{e}^{b'}. \qquad (66)$$

Here again, thanks to the fact that $\mathbf{e}_{b'} = \mathbf{e}^{b'}$, we have written the last term of Eq. (66) in a manner suitable for applying Rule 0. Also, the left-hand side has been written in keeping with Rule 3, although there is more to it than just the desire of not breaking rules that, as appealing as they might look, are arbitrary. Indeed, the vectors in Eqs. (65) and (66) are dual to each other, as the notation chosen suggests. In fact,

$$\begin{aligned}\mathbf{e}^a \cdot \mathbf{e}_b &= (R^a_{c'}\mathbf{e}^{c'}) \cdot (R^{d'}_b\mathbf{e}_{d'}) = R^a_{c'}R^{d'}_b(\mathbf{e}^{c'} \cdot \mathbf{e}_{d'}) \\ &= R^a_{c'}R^{d'}_b\delta^{c'}_{d'} = R^a_{c'}R^{c'}_b = \delta^a_b \end{aligned} \qquad (67)$$

so that they satisfy Eq. (21), the defining equation of dual vectors.

As promised, two bases have indeed emerged from the arbitrary coordinate system $\mathbf{x} \equiv (x^1, x^2, x^3)$: the basis $\{\mathbf{e}_a\}$, *defined at each point in space*, of the vectors tangent to the coordinate curves, and the basis $\{\mathbf{e}^a\}$, *defined at each point in space*, of the vectors normal to the coordinate surfaces. The components of the basis $\{\mathbf{e}_a\}$ are referred to as *contravariant*, whereas the components of the basis $\{\mathbf{e}^a\}$ are referred to as *covariant*. For instance, if $\mathbf{A}(\mathbf{x}) = A^a(\mathbf{x})\mathbf{e}_a = A_a(\mathbf{x})\mathbf{e}^a$ is a vector field, its contravariant and covariant components would transform, respectively, as prescribed by Eqs. (48) and (50), namely

$$A^a = R^a_{b'}A^{b'} \qquad (68)$$

and

$$A_a = R^{b'}_a A_{b'}. \qquad (69)$$

The tensor algebra of the preceding sections remains unchanged, except that $R^a_{b'}$ and $R^{b'}_a$ now have an explicit expression, given by Eq. (63).

Which basis should be used? It just so happens that some tensors are naturally expressed in one basis rather than in another. For instance, taking the differential of Eq. (61b), we have

$$dx^a = \sum_b \frac{\partial x^a}{\partial x^{b'}}dx^{b'} \equiv R^a_{b'}dx^{b'}. \qquad (70)$$

The differentials dx^a do indeed transform as the components of a contravariant vector. However, the differentials $dx^{a'}$ are the Cartesian components of the displacement vector so that we may conclude that dx^a are the *contravariant* components of the displacement vector in *any* coordinate system: $d\mathbf{x} = dx^a\mathbf{e}_a$.

Let us now consider a scalar field $\phi \equiv \phi(\mathbf{x}) = \phi(\mathbf{x}')$ (the equality arises from the definition of a scalar as an invariant under coordinate transformations) and compute $\partial_a\phi(\mathbf{x})$, where we denote $\partial_a \equiv \partial/\partial x^a$, to get

$$\partial_a\phi \equiv \frac{\partial\phi(\mathbf{x})}{\partial x^a} = \frac{\partial\phi(\mathbf{x}')}{\partial x^a} = \sum_b \frac{\partial\phi(\mathbf{x}')}{\partial x^{b'}}\frac{\partial x^{b'}}{\partial x^a} \equiv R^{b'}_a\partial_{b'}\phi. \qquad (71)$$

Here, the second equality follows from the fact that ϕ is a scalar, the third from the chain rule, and the fourth from Eq. (63). From Eq. (71), we have $\partial_a\phi(\mathbf{x}) = R^{b'}_a\partial_{b'}\phi(\mathbf{x}')$, which tells us that the partial derivatives $\partial_a\phi$ do transform as the components of a covariant vector. However, from Eq. (9), we have $\partial_{b'}\phi \equiv (\nabla\phi)_{b'}$, and hence we may conclude that $\partial_a\phi$ are the *covariant* components of the gradient of the scalar field ϕ in *any* coordinate system:

$$\nabla\phi \equiv (\nabla\phi)_a\mathbf{e}^a = (\partial_a\phi)\mathbf{e}^a. \qquad (72)$$

The partial derivative of a scalar field with respect to any coordinate component readily gives the homonymous *covariant* component of the gradient of the field [this is the meaning of the word "naturally" in the sentence above Eq. (70)]. Wishing to compute the contravariant components of the gradient, Eq. (32) gives a recipe: $(\nabla\phi)^a = g^{ab}(\nabla\phi)_b = g^{ab}\partial_b\phi$, where g^{ab} are the contravariant components of the metric tensor.

In principle, the metric tensor may be evaluated from Eqs. (61), (65), (66), (11), (14), and (25). An easier way to evaluate it, however, is from the invariance of $d\mathbf{x} \cdot d\mathbf{x}$. This is $g_{ab}dx^a dx^b$ in arbitrary coordinates, whereas in Cartesian coordinates it is

$$g_{b'c'}dx^{b'}dx^{c'} = \delta_{b'c'}dx^{b'}dx^{c'} = \sum_c dx^{c'}dx^{c'}$$
$$= \sum_c (R_a^{c'}dx^b)(R_b^{c'}dx^b)$$
$$= \left(\sum_c R_a^{c'}R_b^{c'}\right)dx^a dx^b. \tag{73}$$

By comparison with $g_{ab}dx^a dx^b$ we see that

$$g_{ab} = \sum_c R_a^{c'}R_b^{c'}, \tag{74}$$

directly computable from Eqs. (61a) and (63). For instance, for the well-known spherical polar coordinates—for which we replace (x^1, x^2, x^3) and $(x^{1'}, x^{2'}, x^{3'})$ by the more common (r, ϑ, φ) and (x, y, z)—one would get $g_{ab}dx^a dx^b = (dr)^2 + r^2(d\vartheta)^2 + r^2\sin^2\vartheta(d\varphi)^2$, whereby the metric is the diagonal matrix $[g_{ab}] = \text{diag}[1, r^2, r^2\sin^2\vartheta]$.

The partial derivatives of a scalar field $\partial_a\phi$ are then the (covariant) components of a (first-rank) tensor (the gradient of the field), which gives valuable (and, most importantly, coordinate-independent or objectivity-preserving) information on the spatial variation rate of the field. The question naturally arises as to how to secure such important information on the components of a vector field. Do the partial derivatives $\partial_b A_a$ or $\partial_b A^a$ of the components of a vector field, or the second partial derivatives $\partial_{ba}\phi \equiv \partial_b\partial_a\phi$ of a scalar field, have a similar appeal as coordinate-independent, objectivity-preserving quantities? In our language, are $\partial_b A_a$, $\partial_b A^a$, or $\partial_{ba}\phi$ (second-rank) tensors? This is an easy question to answer because we have just to check Eqs. (55)–(57). For instance,

$$\partial_{b'}A_{a'} = \partial_{b'}(R_{a'}^c A_c) = R_{a'}^c(\partial_{b'}A_c) + A_c\partial_{b'}R_{a'}^c$$
$$= R_{a'}^c(R_{b'}^d\partial_d A_c) + A_c R_{b'a'}^c = R_{b'}^d R_{a'}^c\partial_d A_c + A_c R_{b'a'}^c, \tag{75a}$$

where we have used the chain rule in the second equality and have set $\partial_{b'}R_{a'}^c \equiv R_{b'a'}^c$. Due to the extra term $A_c R_{b'a'}^c$, Eq. (75a) shows that $\partial_{b'}A_{a'} \neq R_{b'}^d R_{a'}^c\partial_d A_c$ so that $\partial_b A_a$ is *not* a tensor. Similar reasoning, and reference to Eq. (71), shows that $\partial_{ba}\phi \equiv \partial_b(\partial_a\phi)$ is not a tensor either, and nor is $\partial_b A^a$:

$$\partial_{b'}A^{a'} = \partial_{b'}(R_c^{a'}A^c) = R_c^{a'}(\partial_{b'}A^c) + A^c\partial_{b'}R_c^{a'}$$
$$= R_c^{a'}(R_{b'}^d\partial_d A^c) + A^c R_{b'}^d\partial_d R_c^{a'}$$
$$= R_{b'}^d R_c^{a'}\partial_d A^c + A^c R_{b'}^d R_{dc}^{a'}. \tag{75b}$$

Although $\partial_{ba}\phi \equiv \partial_b(\partial_a\phi)$ and $\partial_b A^a$ turn out not to be tensors, we still have the desire, if not the need, of a physical quantity providing coordinate-independent information on the rate of spatial variation of arbitrary-rank tensors, not just scalars.

Let us compute the bth spatial derivative of a vector, as given by Eq. (28), keeping in mind that in non-Cartesian coordinates the basis vectors are not constant:

$$\partial_b\mathbf{A} = \partial_b(A^c\mathbf{e}_c) = (\partial_b A^c)\mathbf{e}_c + A^c(\partial_b\mathbf{e}_c). \tag{76}$$

In the last term, $\partial_b\mathbf{e}_c$ may be rewritten as a linear combination of the basis vectors to give

$$\partial_b\mathbf{e}_c = (\partial_b\mathbf{e}_c)^a\mathbf{e}_a = [\mathbf{e}^a \cdot (\partial_b\mathbf{e}_c)]\mathbf{e}_a \equiv \Gamma_{bc}^a\mathbf{e}_a, \tag{77}$$

where we have used Eq. (19) in the second equality and introduced the *Christoffel symbol* (of the second kind) as

$$\Gamma_{bc}^a \equiv \mathbf{e}^a \cdot (\partial_b\mathbf{e}_c). \tag{78}$$

This is just the ath contravariant component of $\partial_b\mathbf{e}_c$ and it is symmetric under the exchange of the lower indices

$$\partial_b\mathbf{e}_c = \partial_b(R_c^{a'}\mathbf{e}_{a'}) = (\partial_b R_c^{a'})\mathbf{e}_{a'} = R_{bc}^{a'}\mathbf{e}_{a'} = R_{cb}^{a'}\mathbf{e}_{a'}$$
$$= (\partial_c R_b^{a'})\mathbf{e}_{a'} = \partial_c(R_b^{a'}\mathbf{e}_{a'}) = \partial_c\mathbf{e}_b, \tag{79}$$

where denoting by primes a Cartesian coordinate system (in which the basis vectors are constant), $\partial_b\mathbf{e}_{c'} = 0$, and use has been made of the fact that $R_{bc}^{a'} = R_{cb}^{a'}$ (the mixed second partial derivatives may be exchanged). Therefore, we have

$$\partial_b\mathbf{e}_c = \partial_c\mathbf{e}_b \tag{80}$$

and

$$\Gamma_{bc}^a = \Gamma_{cb}^a. \tag{81}$$

Equation (76) can then be written as

$$\partial_b\mathbf{A} = (\partial_b A^c)\mathbf{e}_c + \Gamma_{bc}^a A^c\mathbf{e}_a = (\partial_b A^a + \Gamma_{bc}^a A^c)\mathbf{e}_a \equiv A_{;b}^a\mathbf{e}_a, \tag{82}$$

where we have defined the covariant derivative of the contravariant components of the vector $\mathbf{A}$ (or more concisely, the covariant derivative of the contravariant vector A^a) as

$$A_{;b}^a \equiv \partial_b A^a + \Gamma_{bc}^a A^c, \tag{83}$$

where the covariant derivative is denoted by a subscript semicolon. Notice that, rather than Eq. (76), we could have instead written

$$\partial_b\mathbf{A} = \partial_b(A_c\mathbf{e}^c) = (\partial_b A_c)\mathbf{e}^c + A_c(\partial_b\mathbf{e}^c) \tag{84}$$

with

$$\partial_b\mathbf{e}^c = (\partial_b\mathbf{e}^c)_a\mathbf{e}^a = [\mathbf{e}_a \cdot (\partial_b\mathbf{e}^c)]\mathbf{e}^a. \tag{85}$$

However,

$$\mathbf{e}_a \cdot (\partial_b\mathbf{e}^c) = \partial_b(\mathbf{e}_a \cdot \mathbf{e}^c) - \mathbf{e}^c \cdot \partial_b\mathbf{e}_a = \partial_b\delta_a^c - \mathbf{e}^c \cdot \partial_b\mathbf{e}_a = -\Gamma_{ba}^c, \tag{86}$$

where we have used Eq. (21) in the second equality and Eq. (78) in the last. Equation (84) finally becomes

$$\partial_b\mathbf{A} = (\partial_b A_c)\mathbf{e}^c - A_c\Gamma_{ba}^c\mathbf{e}^a = (\partial_b A_a - \Gamma_{ba}^c A_c)\mathbf{e}^a \equiv A_{a;b}\mathbf{e}^a, \tag{87}$$

where we have defined the covariant derivative of the covariant (components of the) vector A_a as

$$A_{a;b} \equiv \partial_b A_a - \Gamma_{ba}^c A_c. \tag{88}$$

To compute the covariant derivatives, the needed Christoffel symbols are provided by the precious metric

$$
\begin{aligned}
\Gamma^a_{bc} &= \frac{1}{2}(\Gamma^a_{bc} + \Gamma^a_{cb}) = \frac{1}{2}[\mathbf{e}^a \cdot (\partial_b \mathbf{e}_c) + \mathbf{e}^a \cdot (\partial_c \mathbf{e}_b)] \\
&= \frac{1}{2}[(g^{ad}\mathbf{e}_d) \cdot (\partial_b \mathbf{e}_c) + (g^{ad}\mathbf{e}_d) \cdot (\partial_c \mathbf{e}_b)] \\
&= \frac{1}{2}g^{ad}[\partial_b(\mathbf{e}_d \cdot \mathbf{e}_c) + \partial_c(\mathbf{e}_d \cdot \mathbf{e}_b) - \mathbf{e}_c \cdot (\partial_b \mathbf{e}_d) \\
&\quad - \mathbf{e}_b \cdot (\partial_c \mathbf{e}_d)] \\
&= \frac{1}{2}g^{ad}[\partial_b g_{dc} + \partial_c g_{db} - \mathbf{e}_c \cdot (\partial_d \mathbf{e}_b) - \mathbf{e}_b \cdot (\partial_d \mathbf{e}_c)] \\
&= \frac{1}{2}g^{ad}[\partial_b g_{dc} + \partial_c g_{db} - \partial_d(\mathbf{e}_b \cdot \mathbf{e}_c)] \\
&= \frac{1}{2}g^{ad}[\partial_b g_{dc} + \partial_c g_{db} - \partial_d g_{bc}], \tag{89}
\end{aligned}
$$

where, along with the derivation, we have used, in order, Eqs. (81), (78), (27), (11), (80), and (11) again. The important relation obtained is

$$
\Gamma^a_{bc} = \frac{1}{2}g^{ad}[\partial_b g_{dc} + \partial_c g_{db} - \partial_d g_{bc}]. \tag{90}
$$

From Eqs. (82) and (87), we then have

$$
\partial_b \mathbf{A} = A^a_{;b}\mathbf{e}_a = A_{a;b}\mathbf{e}^a, \tag{91}
$$

where the covariant derivatives are given by Eqs. (83) and (88) and the Christoffel symbol (of the second kind) by Eq. (90).

As we have already seen, the partial-derivative term in Eqs. (83) and (88) is not a tensor and neither is the other term because the Christoffel symbols are not tensors. However, the combination of the two terms *does* make the covariant derivative a tensor, as we will soon see. To demonstrate that the Christoffel symbols are not tensors, it is convenient to start by introducing the Christoffel symbol of the first kind (please pay attention to the relative position of the indices), defined by

$$
\begin{aligned}
\Gamma_{abc} &\equiv g_{ad}\Gamma^d_{bc} = \frac{1}{2}g_{ad}g^{de}[\partial_b g_{ec} + \partial_c g_{eb} - \partial_e g_{bc}] \\
&= \frac{1}{2}\delta^e_a[\partial_b g_{ec} + \partial_c g_{eb} - \partial_e g_{bc}], \tag{92}
\end{aligned}
$$

which gives

$$
\Gamma_{abc} \equiv g_{ad}\Gamma^d_{bc} = \frac{1}{2}[\partial_b g_{ac} + \partial_c g_{ab} - \partial_a g_{bc}]. \tag{93}
$$

From $\Gamma_{abc} \equiv g_{ad}\Gamma^d_{bc}$, we have $g^{ae}\Gamma_{abc} = g^{ae}g_{ad}\Gamma^d_{bc} = \delta^e_d\Gamma^d_{bc} = \Gamma^e_{bc}$, or

$$
\Gamma^a_{bc} = g^{ad}\Gamma_{dbc}. \tag{94}
$$

Incidentally, it should be clear from Eqs. (90) and (93) that all components of the Christoffel symbols are equal to zero in Cartesian coordinates, where the metric-tensor components are given by the (constant) Kronecker delta.

The transformation rule for the Christoffel symbol of the first kind is obtained from Eqs. (93) and (46) as

$$
\begin{aligned}
2\Gamma_{a'b'c'} &= \partial_{b'}g_{a'c'} + \partial_{c'}g_{a'b'} - \partial_{a'}g_{b'c'} = \partial_{b'}(R^d_{a'}R^e_{c'}g_{de}) + \partial_{c'}(R^d_{a'}R^e_{b'}g_{de}) - \partial_{a'}(R^d_{b'}R^e_{c'}g_{de}) \\
&= (R^d_{a'}R^e_{c'})\partial_{b'}g_{de} + (R^d_{a'}R^e_{b'})\partial_{c'}g_{de} - (R^d_{b'}R^e_{c'})\partial_{a'}g_{de} + g_{de}[\partial_{b'}(R^d_{a'}R^e_{c'}) + \partial_{c'}(R^d_{a'}R^e_{b'}) - \partial_{a'}(R^d_{b'}R^e_{c'})] \\
&= R^d_{a'}R^e_{c'}R^f_{b'}\partial_f g_{de} + R^d_{a'}R^e_{b'}R^f_{c'}\partial_f g_{de} - R^d_{b'}R^e_{c'}R^f_{a'}\partial_f g_{de} \\
&\quad + g_{de}[(R^d_{b'a'}R^e_{c'} + R^d_{a'}R^e_{b'c'}) + (R^d_{c'a'}R^e_{b'} + R^d_{a'}R^e_{c'b'}) - (R^d_{a'b'}R^e_{c'} + R^d_{b'}R^e_{a'c'})] \\
&= R^d_{a'}R^e_{b'}R^f_{c'}(\partial_e g_{df} + \partial_f g_{de} - \partial_d g_{ef}) + g_{de}(2R^d_{a'}R^e_{b'c'} + R^d_{c'a'}R^e_{b'} - R^d_{b'}R^e_{a'c'}) \\
&= 2R^d_{a'}R^e_{b'}R^f_{c'}\Gamma_{def} + 2g_{de}R^d_{a'}R^e_{b'c'}. \tag{95}
\end{aligned}
$$

Here, we point out the steps in the next-to-last equality, where extensive use of Rule 2 has been made, and in the last equality, where the cancellation of the last two terms occurs because we can freely exchange the dummy indices d and e so that the symmetry of g_{de} allows the cancellation. We then see that

$$
\Gamma_{a'b'c'} = R^d_{a'}R^e_{b'}R^f_{c'}\Gamma_{def} + g_{de}R^d_{a'}R^e_{b'c'}. \tag{96}
$$

The transformation rule for the Christoffel symbol of the second kind follows from Eqs. (94), (96), and (53) to give

$$
\begin{aligned}
\Gamma^{a'}_{b'c'} &= g^{a'd'}\Gamma_{d'b'c'} = (R^{a'}_e R^{d'}_f g^{ef})(R^p_{d'}R^q_{b'}R^r_{c'}\Gamma_{pqr} + g_{pq}R^p_{d'}R^q_{b'c'}) = R^{a'}_e R^q_{b'}R^r_{c'}(R^{d'}_f R^p_{d'})g^{ef}\Gamma_{pqr} + g^{ef}g_{pq}R^{a'}_e(R^{d'}_f R^p_{d'})R^q_{b'c'} \\
&= R^{a'}_e R^q_{b'}R^r_{c'}(\delta^p_f g^{ef})\Gamma_{pqr} + g^{ef}(g_{pq}\delta^p_f)R^{a'}_e R^q_{b'c'} = R^{a'}_e R^q_{b'}R^r_{c'}(g^{ep}\Gamma_{pqr}) + (g^{ef}g_{fq})R^{a'}_e R^q_{b'c'} \\
&= R^{a'}_e R^q_{b'}R^r_{c'}\Gamma^e_{qr} + (\delta^e_q R^{a'}_e)R^q_{b'c'} = R^{a'}_e R^q_{b'}R^r_{c'}\Gamma^e_{qr} + R^{a'}_q R^q_{b'c'}, \tag{97}
\end{aligned}
$$

or

$$
\Gamma^{a'}_{b'c'} = R^{a'}_d R^e_{b'}R^f_{c'}\Gamma^d_{ef} + R^{a'}_d R^d_{b'c'} = R^{a'}_d R^e_{b'}R^f_{c'}\Gamma^d_{ef} - R^{a'}_{de}R^d_{b'}R^e_{c'}, \tag{98}
$$

where the last equality follows from the fact that

$$R_d^{a'} R_{b'c'}^d = R_d^{a'} \partial_{b'} R_{c'}^d = \partial_{b'}(R_d^{a'} R_{c'}^d) - R_{c'}^d \partial_{b'} R_d^{a'}$$
$$= \partial_{b'} \delta_{c'}^{a'} - R_{c'}^d (R_{b'}^e \partial_e R_d^{a'})$$
$$= -R_{c'}^d R_{b'}^e R_{ed}^{a'} = -R_{ed}^{a'} R_{b'}^e R_{c'}^d = -R_{de}^{a'} R_{b'}^d R_{c'}^e. \tag{99}$$

We then see that, because of the presence of the second terms in Eqs. (96) and (98), the Christoffel symbols do not transform as a tensor. As noted, the partial derivatives do not transform as a tensor either. However, their combination in the covariant derivatives (83) and (88) *does* transform as a tensor

$$A_{a';b'} \equiv \partial_{b'} A_{a'} - \Gamma_{b'a'}^{c'} A_{c'} = (R_{b'}^d R_{a'}^e \partial_d A_e + A_e R_{b'a'}^e) - (R_{b'}^d R_{a'}^e R_f^{c'} \Gamma_{de}^f + R_e^{c'} R_{b'a'}^e) R_{c'}^p A_p$$
$$= R_{b'}^d R_{a'}^e (\partial_d A_e - \Gamma_{de}^f (R_f^{c'} R_{c'}^p) A_p) + A_e R_{b'a'}^e - (R_e^{c'} R_{c'}^p) A_p R_{b'a'}^e = R_{b'}^d R_{a'}^e (\partial_d A_e - \Gamma_{de}^f \delta_f^p A_p) + A_e R_{b'a'}^e - \delta_e^p A_p R_{b'a'}^e$$
$$= R_{b'}^d R_{a'}^e (\partial_d A_e - \Gamma_{de}^f A_f) + A_e R_{b'a'}^e - A_e R_{b'a'}^e = R_{b'}^d R_{a'}^e A_{d;e}, \tag{100}$$

where Eqs. (88), (75a), (50), and the second equality in Eq. (98) have been used. Similarly,

$$A_{;b'}^{a'} \equiv \partial_{b'} A^{a'} + \Gamma_{b'c'}^{a'} A^{c'} = (R_{b'}^d R_e^{a'} \partial_d A^e + A^e R_{b'}^d R_{de}^{a'}) + (R_e^{a'} R_{b'}^d R_{c'}^f \Gamma_{df}^e - R_{de}^{a'} R_{b'}^d R_{c'}^e) R_p^{c'} A^p$$
$$= R_{b'}^d R_e^{a'} (\partial_d A^e + R_{c'}^f \Gamma_{df}^e R_p^{c'} A^p) + A^e R_{b'}^d R_{de}^{a'} - R_{de}^{a'} R_{b'}^d (R_{c'}^e R_p^{c'}) A^p$$
$$= R_{b'}^d R_e^{a'} A_{;d}^e + A^e R_{b'}^d R_{de}^{a'} - R_{de}^{a'} R_{b'}^d \delta_p^e A^p = R_{b'}^d R_e^{a'} A_{;d}^e + A^e R_{b'}^d R_{de}^{a'} - R_{de}^{a'} R_{b'}^d A^e = R_{b'}^d R_e^{a'} A_{;d}^e, \tag{101}$$

where Eqs. (83), (75b), (48), and the first equality in Eq. (98) have been used.

It is worth mentioning that the covariant derivative could have been obtained in the following alternative way, where we only sketch the procedure and leave the details to the reader. Starting from the first equality in Eq. (98), solve for $R_{b'c'}^d$ and insert the result into Eq. (75a). Rearranging the equation obtained, Eq. (101) follows, showing that the covariant derivative defined in Eq. (88) is indeed a tensor. Similarly, starting from the second equality in Eq. (98), solve for $R_{de}^{a'}$ and insert the result into Eq. (75b). Rearranging the equation obtained, Eq. (100) follows, showing that the covariant derivative defined in Eq. (83) is indeed a tensor.

Covariant derivatives of higher-rank tensors are readily written. For instance

$$A_{;c}^{ab} \equiv \partial_c A^{ab} + \Gamma_{cd}^a A^{db} + \Gamma_{cd}^b A^{ad}, \tag{102}$$

$$A_{ab;c} \equiv \partial_c A_{ab} - \Gamma_{ac}^d A_{db} - \Gamma_{ad}^d A_{ad}, \tag{103}$$

and

$$A_{a;c}^b \equiv \partial_c A_a^b + \Gamma_{cd}^b A_a^d - \Gamma_{ac}^d A_d^b, \tag{104}$$

with a scheme that should be self-evident. Just recall that the sign is positive when computing the covariant derivative of a contravariant component and negative when computing the covariant derivative of a covariant component. Rules 1–3 automatically take care of the indices and if any rule is broken there is a mistake. (Regrettably it does not work the other way—satisfying the rules does not guarantee mistake-free algebra.)

The covariant derivative of a product follows the same rules as the ordinary derivative. For instance

$$(A_a B^b)_{;c} = A_{a;c} B^b + B_{;c}^b A_a, \tag{105}$$

as can be easily verified by applying Eq. (104) to $A_a B^b$. (A word of caution: unlike the case of ordinary differentiation,

where the mixed second derivatives may be exchanged $\partial_{bc} A^a = \partial_{cb} A^a$, in covariant differentiation the order is important, $A_{;bc}^a \neq A_{;cb}^a$).

The covariant derivative of the metric is zero. Indeed,

$$g_{ab;c} = \partial_c g_{ab} - \Gamma_{ac}^d g_{db} - \Gamma_{bc}^d g_{ad} = \partial_c g_{ab} - \Gamma_{bac} - \Gamma_{abc}$$
$$= \partial_c g_{ab} - \frac{1}{2}[(\partial_a g_{bc} + \partial_c g_{ab} - \partial_b g_{ac})$$
$$+ (\partial_b g_{ac} + \partial_c g_{ab} - \partial_a g_{bc})] = 0, \tag{106}$$

where Eqs. (103) and (93) have been used. Likewise, from Eq. (104),

$$\delta_{a;c}^b = \partial_c \delta_a^b + \Gamma_{cd}^b \delta_a^d - \Gamma_{ac}^d \delta_d^b = \Gamma_{ca}^b - \Gamma_{ac}^b = 0, \tag{107}$$

and from Eqs. (105) and (107),

$$0 = \delta_{a;c}^b = (g_{ad} g^{db})_{;c} = g_{ad;c} g^{db} + g_{;c}^{db} g_{ad} = g_{;c}^{db} g_{ad}, \tag{108}$$

whereby $g_{;c}^{ab} = 0$. The metric-tensor components are "transparent" to covariant differentiation, a result known as the *Ricci theorem*.

Now that we have a new tensor—the covariant derivative of a field—we may treat it as such. For instance, the second-rank tensor in Eq. (83) may be contracted to a scalar, as explained in Sec. IV. If we do so, we obtain $A_{;a}^a = \partial_a A^a + \Gamma_{ac}^a A^c$. We have already noted that in Cartesian coordinates the Christoffel symbols are zero, and the covariant and contravariant components of a vector are equal to each other. Thus, in Cartesian coordinates, the above contraction becomes $A_{;a}^a = \partial_a A^a = \sum_a \partial_a A_a$. But this is precisely Eq. (7), the definition of the divergence of a vector, which is then a scalar (an invariant). We are then led to *define* the divergence of a vector in an arbitrary coordinate system as

$$\nabla \cdot \mathbf{A} \equiv A_{;a}^a \equiv \partial_a A^a + \Gamma_{ac}^a A^c. \tag{109}$$

Notice that the divergence of a vector is defined in terms of its contravariant components. If the covariant components are available, then $\nabla \cdot \mathbf{A} \equiv A^a_{;a} = (g^{ab}A_b)_{;a} = g^{ab}A_{b;a}$; we define the divergence of (the components of) a covariant vector to be the divergence of the associated contravariant (components of the) vector.

Remarkably, to evaluate the divergence of a vector, it is not necessary to compute the Christoffel symbols. In fact, the ones appearing in Eq. (109) are

$$\Gamma^a_{ac} = \frac{1}{2} g^{ad}[\partial_a g_{dc} + \partial_c g_{da} - \partial_d g_{ac}] = \frac{1}{2} g^{ad}\partial_c g_{ad},$$
(110)

where the cancellation between the first and third terms on the right-hand side of the first equality arises upon exchanging the dummy indices a and d and taking advantage of the symmetry of the metric tensor. However, for the matrix whose elements are g^{ad}, we may write

$$[g^{ad}] = [g^{da}] = [g_{da}]^{-1} = \left[\frac{G_{ad}}{G}\right],$$
(111)

where the first equality follows from the symmetry of the metric, the second from Eq. (14), and the last equality—in which G_{ad} and G are, respectively, the cofactor of the element g_{ad} and the determinant of the matrix $[g_{ab}]$—follows from the definition of the inverse of a matrix. Then for each element of the matrix $[g_{ab}]$ we can write $g^{ab} = G_{ab}/G$. However, from the definition of a determinant we also have $G \equiv g_{ab}G_{ab}$, whereby $\partial G/\partial g_{ad} = G_{ad}$ and $g^{ad} = G_{ad}/G = (1/G)(\partial G/\partial g_{ad})$. Equation (110) becomes

$$\Gamma^a_{ac} = \frac{1}{2} g^{ad}\partial_c g_{ad} = \frac{1}{2G}\frac{\partial G}{\partial g_{ad}}\frac{\partial g_{ad}}{\partial x^c} = \frac{1}{2G}\frac{\partial G}{\partial x^c}$$
$$= \frac{1}{\sqrt{G}}\frac{\partial \sqrt{G}}{\partial x^c} \equiv \frac{\partial_c(\sqrt{G})}{\sqrt{G}},$$
(112)

which can be inserted into Eq. (109) to give

$$\nabla \cdot \mathbf{A} \equiv A^a_{;a} \equiv \partial_a A^a + \frac{\partial_a(\sqrt{G})}{\sqrt{G}}A^a = \frac{\partial_a(\sqrt{G}A^a)}{\sqrt{G}}.$$
(113)

This is our final, and most general, expression for the *divergence* of a vector field, applicable in any coordinate system. Needless to say, we may define divergences of a higher-rank tensor by taking the covariant derivative with respect to the ath coordinate of the bth contravariant component and contracting with respect to the indices a and b.

The *Laplacian* of a scalar field $\phi = \phi(\mathbf{x})$ is defined as the divergence of its gradient [see Eq. (10)]. From Eqs. (72) and (113), we have

$$\Delta\phi \equiv \nabla \cdot \nabla\phi = \frac{\partial_a[\sqrt{G}(\nabla\phi)^a]}{\sqrt{G}} = \frac{\partial_a[\sqrt{G}g^{ab}(\nabla\phi)_b]}{\sqrt{G}},$$
(114)

or

$$\Delta\phi \equiv \nabla \cdot \nabla\phi = \frac{\partial_a[\sqrt{G}g^{ab}\partial_b\phi]}{\sqrt{G}},$$
(115)

which is the general expression, applicable in any coordinate system, for the *Laplacian* of a scalar field.

Finally, as explained at the end of Sec. IV, from a covariant second-rank tensor we can construct an antisymmetric tensor (the one constructed from $A_{a;b}$ would be $A_{a;b} - A_{b;a}$, or one proportional to this). However, from Eq. (88), we have $A_{a;b} - A_{b;a} = \partial_b A_a - \partial_a A_b$ because the terms containing the Christoffel symbol cancel. Thus, even though the partial derivatives of a vector are not tensors (in fact, this is precisely what has led us to the definition of the covariant derivative), their antisymmetric difference is a tensor. A 9-component antisymmetric tensor has only 3 independent components, and this is so in any coordinate system (because antisymmetry is preserved). These three components may then be viewed as the components of a vector that, in the case of $A_{a;b} - A_{b;a}$, turn out to be proportional precisely to the *curl* of the vector **A**. This quantity is defined as the "cross product" between the gradient operator and the vector **A** [as in Eq. (40)]

$$\nabla \times \mathbf{A} = \frac{1}{\sqrt{G}}\det\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ A_1 & A_2 & A_3 \end{bmatrix} = \frac{1}{\sqrt{G}}\varepsilon^{abc}\partial_a A_b\mathbf{e}_c.$$
(116)

VI. ORTHOGONAL COORDINATES

Geometrically, orthogonal coordinates have the property that surface coordinates intersect as normal to each other, and curve coordinates do so as well. The metric matrices are diagonal,

$$[g_{ab}] = \mathrm{diag}[g_{11}, g_{22}, g_{33}] \equiv \mathrm{diag}[h^2_{(1)}, h^2_{(2)}, h^2_{(3)}],$$
(117)

with the *metric coefficients* defined by

$$h_{(a)} \equiv \sqrt{g_{aa}} = \frac{1}{\sqrt{g^{aa}}},$$
(118)

where the second equality follows from the inverse of a diagonal matrix. In orthogonal coordinates, the lengths of the coordinate basis vectors are given by $\sqrt{\mathbf{e}_a \cdot \mathbf{e}_a} = \sqrt{g_{aa}} = h_{(a)}$ and $\sqrt{\mathbf{e}^a \cdot \mathbf{e}^a} = \sqrt{g^{aa}} = h^{-1}_{(a)}$, and the coordinate basis vectors themselves may be expressed as $\mathbf{e}_a = h_{(a)}\hat{e}_a$ and $\mathbf{e}^a = h^{-1}_{(a)}\hat{e}^a$, where $\hat{e}_a = \hat{e}^a$—these basis vectors are an orthonormal set. We have written them in a manner to fulfil our Rule 3 and have inserted within parentheses the metric-coefficient label to remind us—in case of unlikely doubts—that no sum is involved over that index.

Any vector can then be written as

$$\mathbf{A} = A^a\mathbf{e}_a = h_{(a)}A^a\hat{e}_a = \tilde{A}^a\hat{e}_a = A_a\mathbf{e}^a = h^{-1}_{(a)}A_a\hat{e}^a = \tilde{A}_a\hat{e}^a,$$
(119)

where

$$\tilde{A}^a = \tilde{A}_a = h_{(a)}A^a = h^{-1}_{(a)}A_a$$
(120)

are the *physical components*, often preferred by physicists because they have the same physical dimensions of the quantity considered and because they are relative to an orthonormal basis. For instance, in terms of its physical components, the gradient of a scalar field is then

$$\nabla\phi = \left(\frac{1}{h_{(a)}}\partial_a\phi\right)\hat{e}^a.$$
(121)

Calculations should be performed using covariant and contravariant components, changing into physical components only at the end, if these are preferred.

In an orthogonal coordinate system, the Christoffel symbols take a simpler form. From Eqs. (90) and (118), we have

$$
\begin{aligned}
\Gamma^a_{bc} &= \tfrac{1}{2} g^{ad} [\partial_b g_{dc} + \partial_c g_{db} - \partial_d g_{bc}] \\
&= \frac{1}{2h^2_{(a)}} \delta^{ad} [\partial_b g_{dc} + \partial_c g_{db} - \partial_d g_{bc}] \\
&= \frac{1}{2h^2_{(a)}} [\partial_b g_{ac} + \partial_c g_{ab} - \partial_a g_{bc}].
\end{aligned}
\tag{122}
$$

We consider here the following separate cases, which are easily computed from Eq. (122):

(1) All 6 Christoffel symbols with $a \neq b \neq c \neq a$ are zero, because all relevant metric tensor elements are zero: $\Gamma^a_{bc} = 0$.
(2) The 6 Christoffel symbols with $a \neq b = c$ are given by $\Gamma^a_{bb} = -(h_{(b)}/h^2_{(a)})\partial_a h_{(b)}$.
(3) The 12 Christoffel symbols with $a = b \neq c$ are given by $\Gamma^a_{ab} = \partial_b h_{(a)}/h_{(a)}$ (no sum over a!).
(4) The 3 Christoffel symbols with $a = b = c$ are given by $\Gamma^a_{aa} = \partial_a h_{(a)}/h_{(a)}$ (no sum over a!).

For instance, for spherical polar coordinates whose metric is given in the example below Eq. (74), the metric coefficients, from Eq. (117), are $h_{(1)} = 1$, $h_{(2)} = r$, $h_{(3)} = r \sin \vartheta$, and the 27 Christoffel symbols, as computed from the relations above, are $\Gamma^1_{22} = -r$, $\Gamma^1_{33} = -r\sin^2\vartheta$, $\Gamma^2_{12} = \Gamma^2_{21} = \Gamma^3_{13} = \Gamma^3_{31} = 1/r$, $\Gamma^2_{33} = -\sin\vartheta\cos\vartheta$, $\Gamma^3_{23} = \Gamma^3_{32} = \cot\vartheta$; the other 18 symbols are all equal to zero.

The divergence of a vector field in orthogonal coordinates is obtained from Eq. (113) by setting $\sqrt{G} \equiv \sqrt{\det[g_{ab}]} = h_{(1)}h_{(2)}h_{(3)}$. If the vector field is given in terms of its physical components, then from Eqs. (113) and (120) its divergence is

$$
\begin{aligned}
\nabla \cdot \mathbf{A} &= \frac{1}{h_{(1)}h_{(2)}h_{(3)}} \Big[\partial_1\Big(h_{(2)}h_{(3)}\tilde{A}_1\Big) + \partial_2\Big(h_{(1)}h_{(3)}\tilde{A}_2\Big) \\
&\quad + \partial_3\Big(h_{(1)}h_{(2)}\tilde{A}_3\Big) \Big].
\end{aligned}
\tag{123}
$$

Likewise, the Laplacian of a scalar field is given by Eq. (115); in orthogonal coordinates, this equation simplifies further to

$$
\begin{aligned}
\Delta\phi &= \frac{1}{h_{(1)}h_{(2)}h_{(3)}} \Bigg[\partial_1\Bigg(\frac{h_{(2)}h_{(3)}}{h_{(1)}}\partial_1\phi\Bigg) + \partial_2\Bigg(\frac{h_{(1)}h_{(3)}}{h_{(2)}}\partial_2\phi\Bigg) \\
&\quad + \partial_3\Bigg(\frac{h_{(1)}h_{(2)}}{h_{(3)}}\partial_3\phi\Bigg) \Bigg].
\end{aligned}
\tag{124}
$$

As for the curl of a vector field, from Eqs. (116), (118), and (120), we have

$$
\nabla \times \mathbf{A} = \frac{1}{h_{(1)}h_{(2)}h_{(3)}} \det \begin{bmatrix} h_{(1)}\hat{e}_1 & h_{(2)}\hat{e}_2 & h_{(3)}\hat{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ h_{(1)}\tilde{A}_1 & h_{(2)}\tilde{A}_2 & h_{(3)}\tilde{A}_3 \end{bmatrix}.
\tag{125}
$$

We have therefore come back, after a journey along the path of tensor calculus, to expressions—Eqs. (121) and (123)–(125)—that our readers might be familiar with; a good place, as any, to stop.

VII. CONCLUSIONS

With the background from this guide, the reader is ready to tackle many applications, such as the theory of elastic bodies, involving tensors of any rank in any coordinate system. For instance, the equations in the last paragraph of Sec. II regarding the stress tensor have been written in Cartesian coordinates. If a different coordinate system were considered (because the geometrical symmetry of the material would advise us to do so), we have here learned that (i) we must deal with two types of components and (ii) we must write the divergence as the contraction of a covariant derivative. The reader who has followed the paper up to now knows how to write the above equations, namely $F_a = \int_V dV \phi_a = \int_V \sigma^b_{a;b} dV = \int_S \sigma^b_a n_b ds$. In other words, we see that the stress tensor emerges naturally as a mixed second-rank tensor, exactly as the displacement and the gradient emerged naturally as contravariant and covariant first-rank tensors.

Is this the end of the story? Well, we started by saying that we live in a Euclidean 3D space with an invariant distance given by $(dx_1)^2 + (dx_2)^2 + (dx_3)^2$. That was fine. However, we also live in a space that is neither 3D nor Euclidean. Rather, we live in a 4D space-time, where the invariant distance (the interval) is $(cdt)^2 - (dx_1)^2 - (dx_2)^2 - (dx_3)^2$, in which—by a comparison similar to that made below Eq. (74) to obtain the spherical-polar-coordinate metric—the reader will recognize that the metric is $g_{\alpha\beta} = \mathrm{diag}[1,-1,-1,-1] = g^{\alpha\beta}$. The objectivity-preserving quantities are then 4^r-component quantities, and the metric is not Euclidean. But these generalizations should not pose special problems. More surprises would come upon introducing gravity into the picture, but that is a topic we are not broaching here. The readers are now, we hope, well equipped to go into the more modern coordinate-free formulation,[22] for which Jeevanjee's book[8] provides an introduction, and Schutz's book is an excellent follow-up.[23]

[1] E. de Prunelé, "Linear strain tensor and differential geometry," Am. J. Phys. **75**, 881–887 (2007); J. D. Romano and R. H. Price, "Why no shear in 'Div, grad, curl, and all that'?," ibid. **80**, 519–524 (2012).
[2] B. S. Perkalskis and J. R. Freeman, "Examining tensors in the lab: The dielectric permittivity and electrical resistivity of wood," Am. J. Phys. **66**, 816–820 (1998); B. S. Perkalskis and J. R. Freeman, "Tensors in the lab—The thermal resistivity of wood," ibid. **67**, 452–455 (1999).
[3] T. Levi-Civita, *The Absolute Differential Calculus (Calculus of Tensors)* (Dover, USA, 1977).
[4] J. A. Schouten, *Tensor Analysis for Physicists* (Dover, USA, 1989).
[5] D. E. Bourne and P. C. Kendall, *Vector Analysis and Cartesian Tensors* (Academic Press, New York, 1977).
[6] D. Fleisch, *A Student's Guide to Vectors and Tensors* (Cambridge U.P., Cambridge, 2012).
[7] G. E. Hay, *Vector and Tensor Analysis* (Dover, USA, 1953).
[8] N. Jeevanjee, *An Introduction to Tensors and Group Theory for Physicists* (Birkhäuser Boston, Cambridge, Massachusetts, 2011).
[9] A. Lichnerowicz, *Elements of Tensor Calculus* (Methuen & Co., London, 1962).
[10] B. Spain, *Tensor Calculus* (Oliver & Boyd, Edinburgh, 1953).
[11] A. I. Borisenko and I. E. Tarapov, *Vector and Tensor Analysis* (Dover, USA, 1979).
[12] U. C. De, A. A. Shaikh, and J. Sengupta, *Tensor Calculus* (Alpha Science International, Oxford, 2008).

[13] D. C. Kay, *Schaum's Outline of Tensor Calculus* (McGraw-Hill, New York, 1988).

[14] L. P. Lebedev, M. J. Cloud, and V. A. Eremeyev, *Tensor Analysis with Applications in Mechanics* (World Scientific, Singapore, 2010).

[15] J. L. Synge and A. Schild, *Tensor Calculus* (Dover, USA, 1978).

[16] R. C. Wrede, *Introduction to Vector and Tensor Analysis* (Dover, USA, 1972).

[17] For example, see P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953); F. W. Byron, Jr. and R. W. Fuller, *Mathematics of Classical and Quantum Physics* (Dover, USA, 1970).

[18] We are also assuming that the readers have been exposed to Cramer's rule for solving linear algebraic equations, i.e., that they have some familiarity with the properties of determinants and the inversion of matrices.

[19] For completeness, we should add that multiplication by 1 leaves a vector unchanged.

[20] For vectors defined in the complex field, the scalar product would not be commutative.

[21] In the context of crystallography, the lattice space and its basis are called the direct space and direct basis, whereas the dual basis and the lattice space thereby generated are called the reciprocal basis and reciprocal lattice; V is the volume of the elementary cell of the direct lattice, whereas its inverse is the volume of the elementary cell of the reciprocal lattice. These concepts in crystallography are of utmost importance: the constructive interference which provides a diffraction spectrum occurs only when the vector differences between the wave vectors of the incident and diffracted x-rays are vectors of the reciprocal space.

[22] This is in contrast to the formulation followed here and which, as archaic as it might seem, provides nonetheless a practical tool widely used by the working physicist.

[23] B. Schutz, *Geometrical Methods of Mathematical Physics* (Cambridge U.P., Cambridge, 1980).

Steam Engine Model

In March 1882 Walker Hall at Amherst College burned, and with it Amherst's first apparatus collection was lost. The new Walker Hall was equipped with the very latest apparatus, including this sectional model of a steam engine by E. Ducretet of Paris. This is a double-acting engine, with the steam admitted alternately on either side of the piston, as in a steam locomotive. The handle controls the reversing gear that switches the sequence in which steam is admitted to the cylinder. (Notes and photograph by Thomas B. Greenslade, Jr., Kenyon College)