Cambridge University Press & Assessment
978-1-108-84461-1 — An Introduction to Continuum Physics
Steven R. Pride
Excerpt
More Information

Part I

Continuum Physics

1

An Introduction to Tensor Calculus

1.1 Overall Context

We will be developing the laws of continuum physics throughout the first part of this book. We will do so in an uncommon but pedagogic way by starting with the laws that describe the discrete movement of individual atoms and then summing over the molecular dynamics. The emergent continuum laws so obtained come in the form of partial-differential equations (PDEs) that determine how fields are changing in time at each point in space based on how the fields are varying in space in the immediate neighborhood of that point. The second part of the book will treat, for the most part, mathematical techniques for analytically solving the PDEs with a heavy dose of Fourier analysis and contour-integration methods. Students learning this material from me over the years have reported that the first part where the continuum rules are established is more difficult for them compared to the second part where math problems are solved. Perhaps this is analogous to how building a toy model is more challenging than playing with the toy once it is built.

In continuum mechanics in particular, the key field representing the underlying molecular-force interactions is a tensor (the “stress tensor”) and across all of continuum physics, the material properties and constitutive laws are often only describable using tensors. In short, it is impossible to learn continuum physics properly without a solid foundation in tensors and tensor calculus, and this is why we begin the book with this foundational topic. What you learn in this first chapter, especially the tensor-calculus product-rule identities of Section 1.7, will be used at every step throughout our development of the rules of continuum physics. Fortunately, tensors and tensor calculus are a natural, even effortless, extension from the concepts of vectors and vector calculus that I assume you are familiar with. This chapter reviews the various types of spatial derivatives employed in continuum physics (the gradient, divergence, and curl), while allowing these spatial derivatives to act upon tensor fields, which is assumed to be new to the reader. It is my experience that even more senior research scientists can benefit from this chapter’s survey of tensor calculus in preparation for the derivations in all the chapters that follow.

Throughout the book, our focus is on analytical understanding of the physics and mathematics and this involves pencil and paper work. You need to develop confidence in pushing the symbols around the page as you handle and ultimately solve the PDEs we will be deriving. The goal is to build intuition and hands-on familiarity with the physical processes being discussed. Simulating macroscopic experiments in the real world, often performed in complicated heterogeneous bodies of matter with irregular boundaries, is called the *forward problem* and usually needs to be performed numerically because analytical solutions of the governing equations are not possible. Recording the material response at places within a body during various types of experiments and minimizing the difference between the recorded data and simulations of the data with the goal of determining the physical properties throughout the body is called the *inverse problem* and is also a numerical exercise in nearly all cases. But we will not be addressing in this book numerical aspects of the forward and inverse problems posed in macroscopic bodies. Instead we content ourselves with first developing the PDEs that control basic processes of interest across many physical-science disciplines (Part I) and then solving simplified forms of the equations in simple geometries where analytical results are possible so that your physical intuition about the physics can be developed (Part II).

1.2 Some Actors

Any physical quantity continuously distributed over the space of some region is called a *field*. Continuum physics involves the study of fields. Fields can be *scalars, vectors,* or *tensors*.

Scalar Fields: A field quantity that has no intrinsic direction is called a scalar field. Examples include temperature, pressure, and various types of densities. In the nomenclature of tensors, a scalar can be called a zeroth-order tensor.

Vector Fields: A field quantity that has a direction associated with it is called a vector field. Examples include electric fields, fluid velocity, and gravitational acceleration. Vector fields are represented at each point in space by an arrow whose length denotes the amplitude of the vector field at that point. In the nomenclature of tensors, a vector can be called a first-order tensor. Vector fields as depicted in Fig. 1.1 can be written analytically in different ways:

$$\begin{aligned}\boldsymbol{r} &\mathrel{\widehat{=}} \text{position vector used to identify points in space}\\ &= x_1\hat{\boldsymbol{x}}_1 + x_2\hat{\boldsymbol{x}}_2 + x_3\hat{\boldsymbol{x}}_3 = x\,\hat{\boldsymbol{x}} + y\,\hat{\boldsymbol{y}} + z\,\hat{\boldsymbol{z}} \qquad (1.1)\\ &= (x_1, x_2, x_3) = x_i\hat{\boldsymbol{x}}_i \quad \text{(summation over repeated indices)},\end{aligned}$$

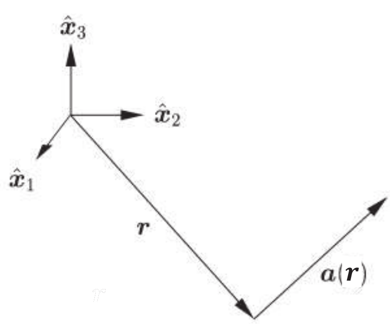

Figure 1.1 Points in space denoted by the vector $\boldsymbol{r}$ and a vector field $\boldsymbol{a}(\boldsymbol{r})$ at each point $\boldsymbol{r}$.

$$\begin{aligned}\boldsymbol{a}(\boldsymbol{r}) = \boldsymbol{a}(x_1, x_2, x_3) &\,\widehat{=}\, \text{vector field defined at each point } \boldsymbol{r} \\ &= a_1\hat{\boldsymbol{x}}_1 + a_2\hat{\boldsymbol{x}}_2 + a_3\hat{\boldsymbol{x}}_3 = a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}} \\ &= (a_1, a_2, a_3) = a_i\hat{\boldsymbol{x}}_i.\end{aligned} \tag{1.2}$$

The caret symbol ˆ placed above a vector means that vector is unitless and has an amplitude of 1, that is, $\hat{\boldsymbol{a}} \,\widehat{=}\, \boldsymbol{a}/|\boldsymbol{a}|$, where $|\boldsymbol{a}| = \sqrt{(a_1^2 + a_2^2 + a_3^2)}$ denotes the amplitude of vector $\boldsymbol{a}$.

IMPORTANT: Whenever an index appears twice in an expression, you always sum over that index. The index that is summed over is sometimes called a *dummy index* because the index does not survive the summation and could be given any name. For example, we have $a_ib_i = a_jb_j = a_nb_n = \sum_{n=1}^{3} a_nb_n = \sum_{j=1}^{3} a_jb_j = \sum_{i=1}^{3} a_ib_i = a_1b_1 + a_2b_2 + a_3b_3$, where the i, j and n are examples of the dummy indices that we sum over. The summation over repeated indices in vectorial and tensorial expressions is called the *Einstein summation convention* and simply saves us from having to write the summation sign over and over.

Another type of vector is the vector operator that we call the *gradient operator* that is defined

$$\nabla = \hat{\boldsymbol{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z} \,\widehat{=}\, \text{gradient operator (a vector operator).} \tag{1.3}$$

For example, if $\psi(\boldsymbol{r}) = \psi(x, y, z)$ is some scalar field, then the gradient of ψ is

$$\nabla\psi = \hat{\boldsymbol{x}}\,\frac{\partial\psi}{\partial x} + \hat{\boldsymbol{y}}\,\frac{\partial\psi}{\partial y} + \hat{\boldsymbol{z}}\,\frac{\partial\psi}{\partial z}. \tag{1.4}$$

and is a vector that we can also write $\nabla\psi = \hat{\boldsymbol{x}}_i\partial\psi/\partial x_i$ using the summation convention. The gradient vector $\nabla\psi$ is oriented in the direction that the scalar field ψ is increasing the most rapidly and the amplitude $|\nabla\psi|$ gives the rate of that maximum increase.

For a vector field $\boldsymbol{a} = a_i\hat{\boldsymbol{x}}_i$, we call the a_i the *scalar components of the vector* and call the unit vectors $\hat{\boldsymbol{x}}_i$ in each direction i the *base vectors*. Note that a vector at some point in space is an arrow with a length and is completely independent of the coordinate system we use to describe it. So $\boldsymbol{a}$ happily exists as the same arrow and does not change if we rotate the coordinate system. Note, however, that the scalar components of the vector a_i will change as we rotate our coordinate system (alter the orientations of the base vectors) or switch to another coordinate system such as cylindrical coordinates.

Some authors put an arrow above a symbol to denote that it is a vector field, i.e., $\vec{a}$. When working in typed text, we always use a bold-face symbol to denote a vector, i.e., $\boldsymbol{a}$. When writing by hand, we have elected not to use an arrow over a symbol but instead use a squiggly underscore, i.e., $\underset{\sim}{a}$.

You are free to develop your own vectorial and tensorial notation when writing by hand but using squiggly underscores for vectors and tensors has served me well over a long career. Note that if you do not use some type of notation to denote that a symbol is a vector or tensor, you will be in a constant state of confusion when manipulating the fields of continuum physics.

Second-Order Tensor Fields: A field quantity that acts as the proportionality between two vector fields that are related to each other at each point in space is called a *second-order tensor* field (can equivalently be called a "second-rank" tensor). Another word that is synonymous to second-order tensor is *dyad* or *dyadic*. We write a second-order tensor field as

$$\begin{aligned}\boldsymbol{T}(\boldsymbol{r}) &\widehat{=} \text{a second-order tensor field defined at each point } \boldsymbol{r}\\ &= T_{xx}\hat{\boldsymbol{x}}\hat{\boldsymbol{x}} + T_{xy}\hat{\boldsymbol{x}}\hat{\boldsymbol{y}} + T_{xz}\hat{\boldsymbol{x}}\hat{\boldsymbol{z}}\\ &\quad + T_{yx}\hat{\boldsymbol{y}}\hat{\boldsymbol{x}} + T_{yy}\hat{\boldsymbol{y}}\hat{\boldsymbol{y}} + T_{yz}\hat{\boldsymbol{y}}\hat{\boldsymbol{z}} \qquad (1.5)\\ &\quad + T_{zx}\hat{\boldsymbol{z}}\hat{\boldsymbol{x}} + T_{zy}\hat{\boldsymbol{z}}\hat{\boldsymbol{y}} + T_{zz}\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}\\ &= T_{ij}\,\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j \quad \text{(summation over repeated indices assumed).} \qquad (1.6)\end{aligned}$$

Just like the vector $\boldsymbol{a} = a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}}$ is the sum of three vectors in the three coordinate directions, so the second-order tensor $\boldsymbol{T}$ is the sum of nine second-order tensors as made explicit in Eq. (1.5). The T_{ij} are the scalar components of the second-order tensor and the various base vector pairs $\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j$ for the various possible i and j are what we call second-order tensors. And just like we can write a vector in the array format $\boldsymbol{a} = (a_x, a_y, a_z)$, so can we write a second-order tensor as

$$\boldsymbol{T} = \begin{pmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{pmatrix}. \qquad (1.7)$$

So a second-order tensor can be represented as a matrix. Much of what you learned about matrices in linear algebra applies to how we use second-order tensors. The main difference between a matrix and a second-order tensor is that although a matrix may have any dimension ($N \times M$) and corresponds to any proportionality between an M and N dimensioned vector (first-order) array, a second-order tensor is a field quantity distributed through three-dimensional space and is always a (3×3) matrix in three-dimensional space and is a physical field that is always the proportionality between two vector fields that each have clear physical meaning as will be demonstrated repeatedly throughout this book.

An example of a second-order tensor is two vector fields that sit side by side to each other in an expression without a scalar or vector product (that are defined in an upcoming section) between them:

$$\boldsymbol{ab} = \left(a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}}\right)\left(b_x\hat{\boldsymbol{x}} + b_y\hat{\boldsymbol{y}} + b_z\hat{\boldsymbol{z}}\right) \qquad (1.8)$$

$$\begin{aligned}&= a_xb_x\,\hat{\boldsymbol{x}}\hat{\boldsymbol{x}} + a_xb_y\,\hat{\boldsymbol{x}}\hat{\boldsymbol{y}} + a_xb_z\,\hat{\boldsymbol{x}}\hat{\boldsymbol{z}}\\ &\quad + a_yb_x\,\hat{\boldsymbol{y}}\hat{\boldsymbol{x}} + a_yb_y\,\hat{\boldsymbol{y}}\hat{\boldsymbol{y}} + a_yb_z\,\hat{\boldsymbol{y}}\hat{\boldsymbol{z}} \qquad (1.9)\\ &\quad + a_zb_x\,\hat{\boldsymbol{z}}\hat{\boldsymbol{x}} + a_zb_y\,\hat{\boldsymbol{z}}\hat{\boldsymbol{y}} + a_zb_z\,\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}\\ &= a_ib_j\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j \quad \text{(summation over repeated indices as always).} \qquad (1.10)\end{aligned}$$

It is convenient to construct the 3×3 matrix representing $\boldsymbol{ab}$ as the matrix product between $\boldsymbol{a}$ written as a 3×1 array and $\boldsymbol{b}$ written as a 1×3 array, which corresponds to the multiplications of Eq. (1.8):

$$\boldsymbol{ab} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \left(b_x, b_y, b_z\right) = \begin{pmatrix} a_x b_x & a_x b_y & a_x b_z \\ a_y b_x & a_y b_y & a_y b_z \\ a_z b_x & a_z b_y & a_z b_z \end{pmatrix}. \tag{1.11}$$

When two vectors sit next to each other to form a second-order tensor, it is common to call that product the *tensor product* or *dyadic product*, even if we will not employ these words outside of this paragraph. Some authors in the engineering literature introduce a special symbol $\otimes$ to denote the tensor product, i.e., $\boldsymbol{a} \otimes \boldsymbol{b} \,\widehat{=}\, \boldsymbol{ab}$. So for the tensor product between the base vectors in any second-order or higher-order tensorial expression, these authors write, for example, $\hat{\boldsymbol{x}} \otimes \hat{\boldsymbol{y}}$ to represent what most authors write more simply as $\hat{\boldsymbol{x}}\hat{\boldsymbol{y}}$. The extra symbol $\otimes$ uses space on the page without providing any needed clarification, which is why we do not use it.

Another example of a second-order tensor is the gradient of a vector field. Working in Cartesian coordinates where derivatives of base vectors are zero, we have

$$\nabla \boldsymbol{a} = \left(\hat{\boldsymbol{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z}\right)\left(a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}}\right) \tag{1.12}$$

$$= \frac{\partial a_x}{\partial x}\hat{\boldsymbol{x}}\hat{\boldsymbol{x}} + \frac{\partial a_y}{\partial x}\hat{\boldsymbol{x}}\hat{\boldsymbol{y}} + \ldots = \frac{\partial a_j}{\partial x_i}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j, \tag{1.13}$$

which can again be written in array form as

$$\nabla \boldsymbol{a} = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \left(a_x, a_y, a_z\right) = \begin{pmatrix} \frac{\partial a_x}{\partial x} & \frac{\partial a_y}{\partial x} & \frac{\partial a_z}{\partial x} \\ \frac{\partial a_x}{\partial y} & \frac{\partial a_y}{\partial y} & \frac{\partial a_z}{\partial y} \\ \frac{\partial a_x}{\partial z} & \frac{\partial a_y}{\partial z} & \frac{\partial a_z}{\partial z} \end{pmatrix}. \tag{1.14}$$

We emphasize that we get these simple expressions for the components of $\nabla \boldsymbol{a}$ only in Cartesian coordinates where derivatives of the base vectors are zero because the base vectors in Cartesians are spatially uniform. When the components of $\nabla \boldsymbol{a}$ are written out in curvilinear coordinates (cylindrical, spherical, etc.) in which the base vectors themselves vary with position in space and thus have nonzero spatial derivatives, the result of performing $\nabla \boldsymbol{a}$ is more complicated. The expressions for $\nabla \boldsymbol{a}$ in arbitrary orthogonal curvilinear coordinates, cylindrical coordinates, and spherical coordinates are all given in Section 1.8.6.

Just like with a matrix, we can talk about the transpose of a second-order tensor $\boldsymbol{T} = T_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j$ and write

$$\boldsymbol{T}^T \,\widehat{=}\, \text{ the transpose of } \boldsymbol{T}$$
$$= T_{ij}\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i = T_{ji}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j. \tag{1.15}$$

Thus to perform the transpose, we can either flip the indices on the scalar components $T_{ij} \rightarrow T_{ji}$ of the tensor or flip the position of the two base vectors as they sit side by side.

Note that like with a vector, a tensor $\boldsymbol{T}$ exists at a point and is independent of the coordinate system. If we rotate or change coordinate systems, $\boldsymbol{T}$ does not change. However, the scalar components of the tensor T_{ij} will change as we change the coordinates because the base vectors $\hat{\boldsymbol{x}}_i$ are changing. When working in typed text, we always denote a second-order tensor with bold type. When we write a second-order tensor by hand, we use two squiggly underscores $\boldsymbol{T} = \underset{\approx}{T}$.

Higher-Order Tensor Fields: The generalization to higher-order tensors is straightforward. A third-order tensor is written

$$_3\boldsymbol{P} = P_{ijk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k \tag{1.16}$$

a fourth-order tensor as

$$_4\boldsymbol{Q} = Q_{ijkl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l \tag{1.17}$$

and so on for still higher-order tensors. Summation over each index is again assumed.

If, for example, a second-order tensor $\boldsymbol{A}$ happens to sit next to two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ we would have the fourth-order tensor

$$\boldsymbol{Aab} = A_{ij}a_k b_l\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l. \tag{1.18}$$

In general, we have $\boldsymbol{Aab} \neq \boldsymbol{Aba} \neq \boldsymbol{aAb} \neq \boldsymbol{bAa} \neq \boldsymbol{abA} \neq \boldsymbol{baA}$, so the order, from left to right, in which tensorial expressions sit next to each other to form higher-order tensors is very important.

The transpose of higher-order tensors must be specified by the way in which the base vectors are moved around relative to each other in the desired transpose operation. So, for example, for the fourth-order tensor $_4\boldsymbol{Q} = Q_{ijkl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l$, we can define transpose operations such as

$$_4\boldsymbol{Q}^{\overset{T}{2134}} = Q_{ijkl}\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l = Q_{jikl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l \tag{1.19}$$

$$_4\boldsymbol{Q}^{\overset{T}{1243}} = Q_{ijkl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_k = Q_{ijlk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l \tag{1.20}$$

$$_4\boldsymbol{Q}^{\overset{T}{3412}} = Q_{ijkl}\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j = Q_{klij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l \tag{1.21}$$

and so on. There are $4! - 1 = 23$ such transposes for a fourth-order tensor, that is, there are $4! - 1$ different ways of placing the four base vectors next to each other that are different than in the nontransposed form. Similarly, an nth-order tensor would have $n! - 1$ different possible transpose operations; so a second-order tensor has only one way to write the transpose.

We write an nth-order tensor $_n\boldsymbol{Q}$ by hand as $_n\underset{\sim}{Q}$ for $n > 2$.

1.3 Some Acts

In tensor calculus, just like in vector calculus, we define two commonly employed types of products between vectors and tensors called the *scalar product* and the *vector product*.

Scalar Products: A scalar product between two vector fields $\boldsymbol{a}$ and $\boldsymbol{b}$ that have an angle θ between them at each point as depicted in Fig. 1.2 is the product of the amplitude of the two vectors after one of the two vectors is projected into the direction of the other vector. The scalar product $\boldsymbol{a}\cdot\boldsymbol{b}$ between two vectors is a scalar and is denoted with a dot sitting between the vectors and is defined by the following rule

$$\boldsymbol{a}\cdot\boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|\cos\theta, \quad \text{where} \;\; |\boldsymbol{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}. \tag{1.22}$$

So $\boldsymbol{a}\cdot\boldsymbol{b} = 0$ if $\boldsymbol{a}\perp\boldsymbol{b}$, which means that $\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{y}} = 0$ and $\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{z}} = 0$, but $\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{x}} = 1$, etc. Using these rules, we thus have

$$\begin{aligned}\boldsymbol{a}\cdot\boldsymbol{b} &= \left[a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}}\right]\cdot\left[b_x\hat{\boldsymbol{x}} + b_y\hat{\boldsymbol{y}} + b_z\hat{\boldsymbol{z}}\right]\\ &= a_xb_x + a_yb_y + a_zb_z \\ &= a_ib_i.\end{aligned} \tag{1.23}$$

The scalar product is also called the *dot product* or the *inner product*.

What if vector field $\boldsymbol{a}$ is related to vector field $\boldsymbol{b}$ at some point in space? How do you obtain $\boldsymbol{a}$ given $\boldsymbol{b}$? That is what a second-order tensor such as $\boldsymbol{T} = T_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j$ does for us once we introduce the scalar product:

$$\begin{aligned}\boldsymbol{a} &= \boldsymbol{T}\cdot\boldsymbol{b}\\ &= \begin{pmatrix} & T_{xx}\hat{\boldsymbol{x}}\hat{\boldsymbol{x}} & + & T_{xy}\hat{\boldsymbol{x}}\hat{\boldsymbol{y}} & + & T_{xz}\hat{\boldsymbol{x}}\hat{\boldsymbol{z}} \\ + & T_{yx}\hat{\boldsymbol{y}}\hat{\boldsymbol{x}} & + & T_{yy}\hat{\boldsymbol{y}}\hat{\boldsymbol{y}} & + & T_{yz}\hat{\boldsymbol{y}}\hat{\boldsymbol{z}} \\ + & T_{zx}\hat{\boldsymbol{z}}\hat{\boldsymbol{x}} & + & T_{zy}\hat{\boldsymbol{z}}\hat{\boldsymbol{y}} & + & T_{zz}\hat{\boldsymbol{z}}\hat{\boldsymbol{z}} \end{pmatrix}\cdot\left(b_x\hat{\boldsymbol{x}} + b_y\hat{\boldsymbol{y}} + b_z\hat{\boldsymbol{z}}\right)\end{aligned} \tag{1.24}$$

$$\begin{aligned}&= \left(T_{xx}\,b_x + T_{xy}\,b_y + T_{xz}\,b_z\right)\hat{\boldsymbol{x}} + \left(T_{yx}\,b_x + T_{yy}\,b_y + T_{yz}\,b_z\right)\hat{\boldsymbol{y}}\\ &\quad + \left(T_{zx}\,b_x + T_{zy}\,b_y + T_{zz}\,b_z\right)\hat{\boldsymbol{z}}\end{aligned} \tag{1.25}$$

$$= \left(T_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\right)\cdot\left(b_k\hat{\boldsymbol{x}}_k\right) = T_{ij}b_k\hat{\boldsymbol{x}}_i\left(\hat{\boldsymbol{x}}_j\cdot\hat{\boldsymbol{x}}_k\right) = T_{ij}b_j\hat{\boldsymbol{x}}_i, \tag{1.26}$$

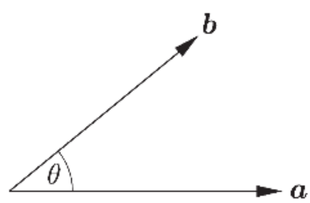

Figure 1.2 Two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ with an angle θ between them.

where in the last line we used that $\hat{\boldsymbol{x}}_j \cdot \hat{\boldsymbol{x}}_k$ requires $k = j$. Using the familiar matrix multiplication for the scalar product, this can be written

$$\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix}. \tag{1.27}$$

> IMPORTANT: Second-order tensor fields are always maps between two vectors that are physically related to each other at a point $\boldsymbol{r}$. You cannot visualize directly a second-order tensor (or higher-order tensors) using your 3D sense of perception. But you can picture in your mind's eye the two vectors (arrows) that are related to each other at a point and thus imagine there is a mapping (second-order tensor) that takes the one vector to the other.

Note that throughout this entire book, we work exclusively in orthogonal coordinates where dot products are zero between the different base vectors of a coordinate system. It is possible, for example, in crystallography, to want to work in *skew* coordinate systems where the base vectors are not orthogonal to each other. Complicating ideas such as covariant and contravariant base vectors arise and the reader interested in tensor calculus in skew coordinates is directed toward specialized texts (e.g., Lebedev et al., 2010).

We can also speak of the *double-dot product* : between tensors, that in this book is defined

$$\boldsymbol{ab} : \boldsymbol{cd} = (\boldsymbol{a} \cdot \boldsymbol{d})(\boldsymbol{b} \cdot \boldsymbol{c}) \tag{1.28}$$
$$= \left(a_i \hat{\boldsymbol{x}}_i \cdot d_j \hat{\boldsymbol{x}}_j\right) \left(b_k \hat{\boldsymbol{x}}_k \cdot c_l \hat{\boldsymbol{x}}_l\right) \tag{1.29}$$
$$= a_i d_j b_k c_l \left(\hat{\boldsymbol{x}}_i \cdot \hat{\boldsymbol{x}}_j\right) \left(\hat{\boldsymbol{x}}_k \cdot \hat{\boldsymbol{x}}_l\right). \tag{1.30}$$

Other authors define the double-dot product as $(\boldsymbol{a} \cdot \boldsymbol{c})(\boldsymbol{b} \cdot \boldsymbol{d})$. Either definition works if used consistently. We choose the convention of Eq. (1.28) so that when you see the : between vectors or base vectors, you perform the first dot product between the vectors that reside immediately on either side of the dot symbol and once that is done, perform the second dot product between the remaining vectors. This convention is the easiest to remember and is highly recommended. In writing out a tensorial expression such as given in Eq. (1.29), always use a different index for each base vector and associated coefficient. Because of the nature of the scalar product in orthogonal coordinate systems, we thus have $l = k$ and $j = i$ in Eq. (1.30) or

$$\boldsymbol{ab} : \boldsymbol{cd} = a_i b_k c_k d_i \quad \text{with summation over repeated indices} \tag{1.31}$$
$$= a_1 b_1 c_1 d_1 + a_2 b_1 c_1 d_2 + a_1 b_2 c_2 d_1 + a_2 b_2 c_2 d_2 \quad \text{in 2D.} \tag{1.32}$$

Note that for two second-order tensors $\boldsymbol{S}$ and $\boldsymbol{T}$, we have $\boldsymbol{S} \cdot \boldsymbol{T} = \left(\boldsymbol{T}^T \cdot \boldsymbol{S}^T\right)^T$ and that $\boldsymbol{S} \cdot \boldsymbol{T} \neq \boldsymbol{T} \cdot \boldsymbol{S}$ in general. For the double-dot product, however, we do have $\boldsymbol{S} : \boldsymbol{T} = \boldsymbol{T} : \boldsymbol{S}$ for any $\boldsymbol{S}$ and $\boldsymbol{T}$, where

$$\boldsymbol{S}:\boldsymbol{T} = S_{ij}\,\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j : T_{kl}\,\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l \tag{1.33}$$

$$= S_{ij}T_{kl}\left(\hat{\boldsymbol{x}}_j\cdot\hat{\boldsymbol{x}}_k\right)\left(\hat{\boldsymbol{x}}_i\cdot\hat{\boldsymbol{x}}_l\right) \quad \text{which requires } l=i \text{ and } k=j \tag{1.34}$$

$$= S_{ij}T_{ji} \quad \text{with summation over repeated indices.} \tag{1.35}$$

Renaming the dummy indices gives $\boldsymbol{S}:\boldsymbol{T} = S_{ij}T_{ji} = S_{ji}T_{ij} = T_{ij}S_{ji} = \boldsymbol{T}:\boldsymbol{S}$.

The second-order *identity tensor* $\boldsymbol{I}$ is defined $\boldsymbol{I} = \delta_{ij}\,\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j$, where the δ_{ij} are called the *Kronecker coefficients* and are defined

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad \text{so that} \quad \boldsymbol{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.36}$$

Upon summing over the indices, we have $\boldsymbol{I} = \hat{\boldsymbol{x}}_1\hat{\boldsymbol{x}}_1 + \hat{\boldsymbol{x}}_2\hat{\boldsymbol{x}}_2 + \hat{\boldsymbol{x}}_3\hat{\boldsymbol{x}}_3 = \hat{\boldsymbol{x}}\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}}\hat{\boldsymbol{y}} + \hat{\boldsymbol{z}}\hat{\boldsymbol{z}}$. The identity tensor works as follows: $\boldsymbol{A}\cdot\boldsymbol{I} = \boldsymbol{I}\cdot\boldsymbol{A} = \boldsymbol{A}$ for any second-order tensor $\boldsymbol{A}$. We further have that if the position vector is written $\boldsymbol{r} = x_j\hat{\boldsymbol{x}}_j$ in Cartesian coordinates, then $\boldsymbol{I} = \nabla\boldsymbol{r} = (\partial x_j/\partial x_i)\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j = \delta_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j$.

A double-dot product with the identity tensor results in $\boldsymbol{A}:\boldsymbol{I} = A_{ij}\,\delta_{ji} = A_{ii} = \text{tr}\,\{\boldsymbol{A}\} = A_{11} + A_{22} + A_{33}$, which is called the *trace* of second-order tensor $\boldsymbol{A}$. The trace is the sum of the second-order tensor components along the diagonal, for example, $\boldsymbol{I}:\boldsymbol{I} = 3$ (in 3D). The double-dot product between two second-order tensors is the trace of the scalar (matrix) product of the two tensors, that is, $\boldsymbol{A}:\boldsymbol{B} = A_{ij}B_{ji} = \text{tr}\,\{\boldsymbol{A}\cdot\boldsymbol{B}\} = \text{tr}\,\{\boldsymbol{B}\cdot\boldsymbol{A}\}$.

We can extend the number of dot products we take between two higher-order tensors to as many as desired. So the *triple-dot product* $\overset{3}{\vdots}$ between, say, two third-order tensors ${}_3\boldsymbol{S}$ and ${}_3\boldsymbol{T}$ can be defined

$$\begin{aligned} {}_3\boldsymbol{S} \overset{3}{\vdots} {}_3\boldsymbol{T} &= S_{ijk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k \overset{3}{\vdots} T_{lmn}\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n \\ &= S_{ijk}T_{lmn}(\hat{\boldsymbol{x}}_k\cdot\hat{\boldsymbol{x}}_l)(\hat{\boldsymbol{x}}_j\cdot\hat{\boldsymbol{x}}_m)(\hat{\boldsymbol{x}}_i\cdot\hat{\boldsymbol{x}}_n), \end{aligned}$$

which tells us that $n = i$, $m = j$ and $l = k$ so that the triple-dot product between two third-order tensors comes out to be

$${}_3\boldsymbol{S} \overset{3}{\vdots} {}_3\boldsymbol{T} = S_{ijk}T_{kji} = \text{tr}\,\{{}_3\boldsymbol{S} : {}_3\boldsymbol{T}\} \tag{1.37}$$

with summation over repeated indices. We can extend such notation and definition to still higher-order dot products between still higher-order tensors.

Note that each dot product removes two base vectors from a tensorial expression. So without writing anything out, we know that a tensorial expression like ${}_8\boldsymbol{A} \overset{5}{\vdots} {}_6\boldsymbol{B}$ is a fourth-order tensor, that is, the eighth-order tensor ${}_8\boldsymbol{A}$ contributes 8 base vectors to this expression and the sixth-order tensor ${}_6\boldsymbol{B}$ contributes 6 more base vectors but the 5 dot products remove 10 of those base vectors so that the result is a fourth-order tensor. As practice, we can write this lengthy example out to give

$$\begin{aligned} &{}_8\boldsymbol{A} \overset{5}{\vdots} {}_6\boldsymbol{B} = \\ &\quad A_{ijklmnop}B_{qrstuv}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k(\hat{\boldsymbol{x}}_p\cdot\hat{\boldsymbol{x}}_q)(\hat{\boldsymbol{x}}_o\cdot\hat{\boldsymbol{x}}_r)(\hat{\boldsymbol{x}}_n\cdot\hat{\boldsymbol{x}}_s)(\hat{\boldsymbol{x}}_m\cdot\hat{\boldsymbol{x}}_t)(\hat{\boldsymbol{x}}_l\cdot\hat{\boldsymbol{x}}_u)\hat{\boldsymbol{x}}_v, \end{aligned} \tag{1.38}$$