Lecture 2
Review of Scalars, Vectors, Tensors, and Dyads

Our vice always lies in the direction of our virtues, and in their best estate are but plausible imitations of the latter.
Henry David Thoreau

In MHD, we will deal with relationships between quantities such as the magnetic field and the velocity that have both magnitude and direction. These quantities are examples of vectors (or, as we shall soon see, pseudovectors). The basic concepts of scalar and vector quantities are introduced early in any scientific education. However, to formulate the laws of MHD precisely, it will be necessary to generalize these ideas and to introduce the less familiar concepts of matrices, tensors, and dyads. The ability to understand and manipulate these abstract mathematical concepts is essential to learning MHD. Therefore, for the sake of both reference and completeness, this lecture is about the mathematical properties of scalars, vectors, matrices, tensors, and dyads. If you are already an expert, or think you are, please skip class and go on to Lecture 3. You can always refer back here if needed!

A *scalar* is a quantity that has *magnitude*. It can be written as

$$S \quad \alpha \quad 9 \tag{2.1}$$

It seems self-evident that such a quantity is independent of the coordinate system in which it is measured. However, we will see later in this lecture that this is somewhat naïve, and we will have to be more careful with definitions. For now, we say that the magnitude of a scalar is independent of coordinate transformations that involve translations or rotations.

A *vector* is a quantity that has both *magnitude* and *direction*. It is often printed with an arrow over it (as in $\vec{V}$) or in bold-face type (as in $\mathbf{V}$, which is my preference). When handwritten, I use an underscore (as in $\underline{V}$, although many prefer the arrow notation here, too). It can be geometrically represented as an arrow. A vector has a tail and a head (where the arrowhead is). Its *magnitude* is represented by its length. We emphasize that the vector has an "absolute" orientation in space, i.e., it exists independent of any particular coordinate system. Vectors are therefore "coordinate-free" objects, and expressions involving vectors are true in any coordinate system. Conversely, if an expression involving vectors is true in one coordinate system, it is true in all coordinate systems. (As with the scalar, we will be more careful with our statements in this regard later in this lecture.)

Schnack, D.D.: *Review of Scalars, Vectors, Tensors, and Dyads*. Lect. Notes Phys. **780**, 5–18 (2009)
DOI 10.1007/978-3-642-00688-3_2

Vectors are added with the parallelogram rule. This is shown geometrically in Fig. 2.1.

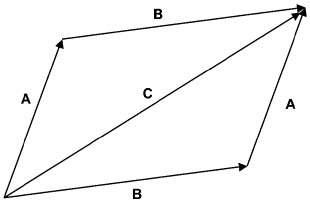

Fig. 2.1 Illustration of the parallelogram rule for adding vectors

This is represented algebraically as $\mathbf{C} = \mathbf{A} + \mathbf{B}$.

We define the *scalar product* of two vectors $\mathbf{A}$ and $\mathbf{B}$ as

$$\mathbf{A} \cdot \mathbf{B} = AB \cos\theta, \tag{2.2}$$

where A and B are the magnitudes of $\mathbf{A}$ and $\mathbf{B}$ and θ is the angle (in radians) between them, as in Fig. 2.2:

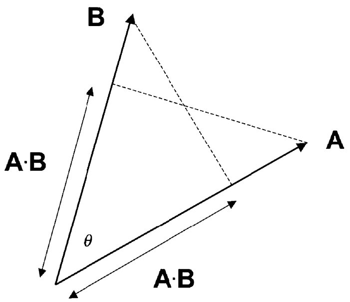

Fig. 2.2 Illustration of the scalar product of two vectors **A** and **B** as the projection of one on the other

The quantity $S = \mathbf{A} \cdot \mathbf{B}$ is the projection of $\mathbf{A}$ on $\mathbf{B}$, and vice versa. Note that it can be negative or zero. We will soon prove that S is a scalar.

It is sometimes useful to refer to a vector $\mathbf{V}$ with respect to some coordinate system (x_1, x_2, x_3), as shown in Fig. 2.3. Here the coordinate system is orthogonal. The vectors $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, and $\hat{\mathbf{e}}_3$ have unit length and point in the directions of x_1, x_2, and x_3, respectively. They are called *unit basis vectors*. The *components* of $\mathbf{V}$ with respect to (x_1, x_2, x_3) are then defined as the scalar products

$$V_1 = \mathbf{V} \cdot \hat{\mathbf{e}}_1,\ V_2 = \mathbf{V} \cdot \hat{\mathbf{e}}_2,\ V_3 = \mathbf{V} \cdot \hat{\mathbf{e}}_3. \tag{2.3a,b,c}$$

The three numbers (V_1, V_2, V_3) also define the vector $\mathbf{V}$.

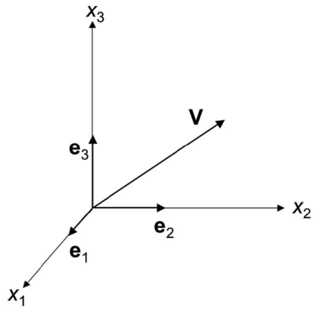

Fig. 2.3 A vector **V** and the basis vectors in a three-dimensional Cartesian coordinate system

Of course, a vector can be referred to another coordinate system (x_1', x_2', x_3') by means of a *coordinate transformation*. This can be expressed as

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3, \end{aligned} \tag{2.4}$$

where the nine numbers a_{ij} are independent of position; it is a *linear transformation*. Equation (2.4) can be written as

$$x_i' = \sum_{j=1}^{3} a_{ij}x_j,\, i = 1, 2, 3. \tag{2.5}$$

We will often use the shorthand notation

$$x_i' = a_{ij}x_j, \tag{2.6}$$

with an implied summation over the repeated index (in this case j). This is called the Einstein summation convention. Since the repeated index j does not appear in the result (the left-hand side), it can be replaced by any other symbol. It is called a *dummy index*. The economy of the notation of (2.6) over (2.4) is self-evident.

Equation (2.6) is often written as

$$\mathbf{x}' = \mathsf{\mathbf{A}} \cdot \mathbf{x}, \tag{2.7}$$

where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{2.8}$$

is called a *column vector*. The transpose of $\mathbf{x}$ is the *row vector*

$$\mathbf{x}^T = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}. \tag{2.9}$$

The nine numbers arranged in the array

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \tag{2.10}$$

form a *matrix*. (In this case the matrix is 3×3.) The "dot product" in Eq. (2.7) implies summation over the neighboring indices, as in Eq. (2.6). Note that $\mathbf{x}^T \cdot \mathsf{A} \equiv x_j a_{ji} \neq \mathsf{A} \cdot \mathbf{x}$ (unless A is *symmetric*, i.e., $a_{ij} = a_{ji}$).

Differentiating Eq. (2.6) with respect to x_k, we find

$$\frac{\partial x_i'}{\partial x_k} = a_{ij} \frac{\partial x_j}{\partial x_k} = a_{ij} \delta_{jk} = a_{ik}, \tag{2.11}$$

which defines the transformation coefficients a_{ik}.

For reference, we give some matrix definitions and properties:

1. The *identity matrix* is defined as

$$\mathsf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \delta_{ij}. \tag{2.12}$$

2. The *inverse matrix* A^{-1}, is defined by $\mathsf{A}^{-1} \cdot \mathsf{A} = \mathsf{I}$.
3. If a_{ij} are the components of A, then a_{ji} are the components of A^T, the *transpose* of A. (If $\mathsf{A} = \mathsf{A}^T$, A is *symmetric*.)
4. The *adjoint matrix* is defined by $\mathsf{A}^\dagger = \mathsf{A}^{*T}$, where $(..)^*$ is the complex conjugate; i.e., $a_{ij}^\dagger = a_{ji}^*$.
5. If $\mathsf{A} = \mathsf{A}^\dagger$, then A is said to *self-adjoint*. (This is the generalization of a symmetric matrix to the case where the components are complex numbers.)
6. Matrix multiplication is defined by $\mathsf{A} \cdot \mathsf{B} = A_{ij} B_{jk}$.
7. $(\mathsf{A} \cdot \mathsf{B})^T = \mathsf{B}^T \cdot \mathsf{A}^T$.
8. $(\mathsf{A} \cdot \mathsf{B})^\dagger = \mathsf{B}^\dagger \cdot \mathsf{A}^\dagger$.

The prototypical vector is the *position vector*

$$\mathbf{r} = x_1 \hat{\mathbf{e}}_1 + x_2 \hat{\mathbf{e}}_2 + x_3 \hat{\mathbf{e}}_3 \equiv (x_1, x_2, x_3)\,. \tag{2.13}$$

It represents a vector from the origin of coordinates to the point $P\,(x_1, x_2, x_3)$. We say that the three numbers (V_1, V_2, V_3) are the components of a vector if they transform like the components of the position vector $\mathbf{r}$ under *coordinate rotations*. Vectors are defined by their transformation properties.

We require that the *length* of the position vector, defined by $l^2 = \mathbf{x}^T \cdot \mathbf{x}$, be invariant under coordinate rotations, i.e., $l^2 = \mathbf{x}^T \cdot \mathbf{x} = \mathbf{x}'^T \cdot \mathbf{x}'$. Then

$$
\begin{aligned}
l^2 = \mathbf{x}^T \cdot \mathbf{x} = \mathbf{x}'^T \cdot \mathbf{x}' \\
&= (\mathsf{A} \cdot \mathbf{x})^T \cdot (\mathsf{A} \cdot \mathbf{x}) \\
&= \left(\mathbf{x}^T \cdot \mathsf{A}^T\right) \cdot (\mathsf{A} \cdot \mathbf{x}) \\
&= \mathbf{x}^T \cdot \left(\mathsf{A}^T \cdot \mathsf{A}\right) \cdot \mathbf{x},
\end{aligned}
$$

so that $\mathsf{A}^T \cdot \mathsf{A} = \mathsf{I}$, or $\mathsf{A}^T = \mathsf{A}^{-1}$. Matrices with this property are called *orthogonal matrices*, and the rotation matrix A is an orthogonal matrix, i.e.,

$$
a_{ij}^{-1} = a_{ji}, \tag{2.14}
$$

or, since $\mathbf{x} = \mathsf{A}^{-1} \cdot \mathbf{x}'$,

$$
\frac{\partial x_i'}{\partial x_j} = \frac{\partial x_j}{\partial x_i'}. \tag{2.15}
$$

Then the components of the rotation matrix A have the property

$$
a_{ij}a_{ik} = \frac{\partial x_i'}{\partial x_j}\frac{\partial x_i'}{\partial x_k} = \frac{\partial x_i'}{\partial x_j}\frac{\partial x_k}{\partial x_i'} = \frac{\partial x_k}{\partial x_j} = \delta_{jk}. \tag{2.16}
$$

We now say that *the three numbers V_1, V_2, and V_3 are the components of a vector if they transform like the position vector* $\mathbf{r}$ *under coordinate rotations*, i.e.,

$$
V_i' = a_{ij} V_j, \tag{2.17}
$$

where the a_{ij} are the components of an orthogonal matrix. (Note that *not* all triplets are components of vectors.)

Suppose that $\mathbf{A}$ and $\mathbf{B}$ are vectors. As an illustration of the algebraic manipulation of vector quantities, we now prove that the product defined in Eq. (2.2) is a scalar. To do this, we must show that $S' = \mathbf{A}' \cdot \mathbf{B}'$, the value of the product in the primed coordinate system, is the same as $S = \mathbf{A} \cdot \mathbf{B}$, the value in the unprimed system:

$$
\begin{aligned}
S' &= \mathbf{A}' \cdot \mathbf{B}' \\
&= a_{ij} A_j a_{ik} B_k \\
&= a_{ij} a_{ik} A_j B_k \\
&= \delta_{jk} A_j B_k \\
&= A_j B_j = S,
\end{aligned}
$$

where the property of orthogonal matrices defined in Eq. (2.16) has been used. Further, if $S = \mathbf{A} \cdot \mathbf{B}$ is a scalar, and $\mathbf{B}$ is a vector, then $\mathbf{A}$ is also a vector.

In addition to the scalar product of two vectors, we can also define the *vector product* of two vectors. The result is another vector. This operation is written symbolically as $\mathbf{C} = \mathbf{A} \times \mathbf{B}$. The magnitude of $\mathbf{C}$, C, is given by $C = AB \sin\theta$, in analogy with Eq. (2.2). By definition, the direction of **C** is perpendicular to the plane defined by **A** and **B** along with the "right-hand rule." Note that it can point either "above" or "below" the plane, and may be zero.

In index notation, the vector product is written as

$$C_i = \varepsilon_{ijk} A_j B_k. \tag{2.18}$$

The quantity ε_{ijk} is called the *Levi-Civita tensor density*. It will prove to be quite important and useful in later analysis. It has 27 components, most of which vanish. These are defined as

$$\left.\begin{array}{l} \varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1 \text{ (even permutation of the indices)} \\ \varepsilon_{132} = \varepsilon_{213} = \varepsilon_{321} = -1 \text{ (odd permutation of the indices)} \\ \varepsilon_{ijk} = 0 \text{ if } i = j, \text{ or } i = k, \text{ or } j = k \end{array}\right\}. \tag{2.19}$$

We should really prove that **C** so defined is a vector, but I will leave it as an exercise for the student.

The Levi-Civita symbol satisfies the very useful identity

$$\varepsilon_{ijk}\varepsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}. \tag{2.20}$$

The expression can be used to derive a wide variety of formulas and identities involving vectors and tensors. For example, consider the "double cross product" $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$. We write this in a Cartesian coordinate system as

$$\begin{aligned} \mathbf{A} \times (\mathbf{B} \times \mathbf{C}) &\Rightarrow \varepsilon_{ijk} A_j (\mathbf{B} \times \mathbf{C})_k = \varepsilon_{ijk} A_j (\varepsilon_{klm} B_l C_m) \\ &= \varepsilon_{ijk}\varepsilon_{klm} A_j B_l C_m = \varepsilon_{ijk} \underbrace{\varepsilon_{lmk}}_{\text{Even permutation of indices}} A_j B_l C_m \\ &= \underbrace{\left(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}\right)}_{\text{From Equation (2.20)}} A_j B_l C_m = \delta_{il} A_j B_l C_j - \delta_{im} A_j B_j C_m \\ &= A_j B_i C_j - A_j B_j C_i = B_i \left(A_j C_j\right) - C_i \left(A_j B_j\right) \\ &\Rightarrow \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}), \end{aligned} \tag{2.21}$$

which is commonly known as the "BAC-CAB rule." The first step in the derivation was to translate the vector formula into components in some convenient Cartesian coordinate system, then turn the crank. We recognized the formula in the line preceding Eq. (2.21) as another vector formula expressed in the same Cartesian system. However, *if a vector formula is true on one system, it is true in all systems* (even generalized, non-orthogonal, curvilinear coordinates), so we are free to translate it back into vector notation. This is a very powerful technique for simplifying and manipulating vector expressions.

We define the *tensor product* of two vectors **B** and **C** as $\mathbf{A} = \mathbf{BC}$ or

$$A_{ij} = B_i C_j. \tag{2.22}$$

How do the nine numbers A_{ij} transform under rotations? Since **B** and **C** are vectors, we have

$$A'_{ij} = B'_i C'_j = (a_{ik} B_k)\left(a_{jl} C_l\right) = a_{ik} a_{jl} B_k C_l,$$

or

$$A'_{ij} = a_{ik} a_{jl} A_{kl}. \tag{2.23}$$

Equation (2.23) is the *tensor transformation law*. Any set of nine numbers that transform like this under rotations form the components of a tensor.

The *rank* of the tensor is the number of indices. We notice that a scalar is a tensor of rank zero, a vector is a first-rank tensor, the 3×3 array just defined is a second-rank tensor, etc. In general, a tensor transforms according to

$$A'_{ijkl.....} = a_{ip} a_{jq} a_{kr} a_{ls} \ldots\ldots A_{pqrs.....}. \tag{2.24}$$

We can also write A in *dyadic notation*:

$$\begin{aligned} \mathsf{A} = \mathbf{BC} &= (B_1 \hat{\mathbf{e}}_1 + B_2 \hat{\mathbf{e}}_2 + B_3 \hat{\mathbf{e}}_3)(C_1 \hat{\mathbf{e}}_1 + C_2 \hat{\mathbf{e}}_2 + C_3 \hat{\mathbf{e}}_3) \\ &= B_1 C_1 \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1 + B_1 C_2 \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 + B_1 C_3 \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_3 \\ &\quad + B_2 C_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_1 + B_2 C_2 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2 + B_2 C_3 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 \\ &\quad + B_3 C_1 \hat{\mathbf{e}}_3 \hat{\mathbf{e}}_1 + B_3 C_2 \hat{\mathbf{e}}_3 \hat{\mathbf{e}}_2 + B_3 C_3 \hat{\mathbf{e}}_3 \hat{\mathbf{e}}_3. \end{aligned} \tag{2.25}$$

The quantities $\hat{\mathbf{e}}_i \hat{\mathbf{e}}_j$ are called *unit dyads*. Note that

$$\hat{\mathbf{e}}_1 \cdot \mathsf{A} = B_1 C_1 \hat{\mathbf{e}}_1 + B_1 C_2 \hat{\mathbf{e}}_2 + B_1 C_3 \hat{\mathbf{e}}_3 \tag{2.26}$$

is a vector, while

$$\mathsf{A} \cdot \hat{\mathbf{e}}_1 = B_1 C_1 \hat{\mathbf{e}}_1 + B_2 C_1 \hat{\mathbf{e}}_2 + B_3 C_1 \hat{\mathbf{e}}_3 \tag{2.27}$$

is a *different* vector. In general, $\mathbf{BC} \neq \mathbf{CB}$.

We could similarly define higher-rank tensors and dyads as $\mathbf{D} = \mathsf{A}\mathbf{E}$, or $D_{ijk} = A_{ij} E_k$, etc.

Contraction is defined as summation over a pair of indices, e.g., $D_i = A_{ij} E_j$. Contraction reduces the rank by 2. We have also used the notation $\mathbf{D} = \mathsf{A} \cdot \mathbf{E}$ to indicate contraction over "neighboring" indices. (Note that $\mathsf{A} \cdot \mathbf{E} \neq \mathbf{E} \cdot \mathsf{A}$.) The "double-dot" notation $(\mathbf{ab}) : (\mathbf{cd})$ is often used, but is ambiguous. We define $\mathsf{A} : \mathsf{B} \equiv A_{ij} B_{ij}$, a scalar.

We now define a *differential operator* in our Cartesian coordinate system[1]

$$\nabla \equiv \hat{\mathbf{e}}_1 \frac{\partial}{\partial x_1} + \hat{\mathbf{e}}_2 \frac{\partial}{\partial x_2} + \hat{\mathbf{e}}_3 \frac{\partial}{\partial x_3} \equiv \hat{\mathbf{e}}_i \frac{\partial}{\partial x_i} \equiv \hat{\mathbf{e}}_i \partial_i . \tag{2.28}$$

The symbol ∇ is sometimes called "nabla," and more commonly, "grad," which is short for "gradient." So far, it is just a linear combination of partial derivatives; it needs something more. What happens when we let it "operate" on a scalar function $f(x_1, x_2, x_3)$? We have

$$\nabla f = \hat{\mathbf{e}}_1 \frac{\partial f}{\partial x_1} + \hat{\mathbf{e}}_2 \frac{\partial f}{\partial x_2} + \hat{\mathbf{e}}_3 \frac{\partial f}{\partial x_3} . \tag{2.29}$$

What kind of a "thing" is ∇f? Consider the quantity $g = d\mathbf{x} \cdot \nabla f$, where $d\mathbf{x} = \hat{\mathbf{e}}_1 dx_1 + \hat{\mathbf{e}}_2 dx_2 + \hat{\mathbf{e}}_3 dx_3$ is a vector defining the differential change in the position vector:

$$\begin{aligned} g &= d\mathbf{x} \cdot \nabla f, \\ &= dx_1 \frac{\partial f}{\partial x_1} + dx_2 \frac{\partial f}{\partial x_2} + dx_3 \frac{\partial f}{\partial x_3} = df, \end{aligned}$$

which we recognize as the differential change in f, and therefore a scalar. Therefore, by the argument given previously, since $d\mathbf{x} \cdot \nabla f$ is a scalar, and $d\mathbf{x}$ is a vector, the three quantities $\partial f/\partial x_1$, $\partial f/\partial x_2$, and $\partial f/\partial x_3$ form the components of a vector, so ∇f is a vector. It measures the magnitude and direction of the rate of change of the function f at any point in space.

Now form the dyad $\mathbf{D} = \nabla \mathbf{V}$, where $\mathbf{V}$ is a vector. Then the nine quantities $D_{ij} = \partial_i V_j$ are the components of a second-rank tensor. If we contract over the indices i and j we have

$$D = \partial_i V_i \equiv \nabla \cdot \mathbf{V}, \tag{2.30}$$

which is a scalar. It is called the *divergence* of $\mathbf{V}$.

We can take the vector product of ∇ and $\mathbf{V}$, $\mathbf{D} = \nabla \times \mathbf{V}$, or

$$D_i = \varepsilon_{ijk} \partial_j V_k . \tag{2.31}$$

This is called the *curl* of $\mathbf{V}$. For example, in Cartesian coordinates, the x_1 component is

[1] The discussion of the vector nature of the gradient operator follows that of the Feynman's lectures: R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* Vol. 1, Addison-Wesley, Reading, MA (1963).

$$D_1 = \varepsilon_{123}\frac{\partial V_2}{\partial x_3} + \varepsilon_{132}\frac{\partial V_3}{\partial x_2} = \frac{\partial V_2}{\partial x_3} - \frac{\partial V_3}{\partial x_2},$$

by the properties of ε_{ijk}.

We could also have ∇ operate on a tensor or dyad: $\nabla \mathsf{A} \Rightarrow \partial_i A_{jk}$, which is a third-rank tensor. A common notation for this is $A_{jk,i}$ (the comma denotes differentiation with respect to x_i). Contracting over i and j,

$$D_k = \partial_j A_{jk} = \nabla \cdot \mathsf{A}, \tag{2.32}$$

or $A_{jk,j}$, which is the *divergence of a tensor* (it is a vector). In principle we could define the curl of a tensor, etc.

So far we have worked in Cartesian coordinates. This is because they are easy to work with, and if a vector expression is true in Cartesian coordinates it is true in *any* coordinate system.

We will now talk about *curvilinear coordinates*. Curvilinear coordinates are still orthogonal but the unit vectors $\hat{\mathbf{e}}_i$ are functions of $\mathbf{x}$, and this complicates the computation of derivatives. Examples of orthogonal curvilinear coordinates are cylindrical and spherical coordinates.

The gradient operator is

$$\nabla = \sum_i \hat{\mathbf{e}}_i \frac{\partial}{\partial x_i} \tag{2.33}$$

(the order of the unit vector and the derivative is now important), and any vector $\mathbf{V}$ is

$$\mathbf{V} = \sum_j V_j \hat{\mathbf{e}}_j, \tag{2.34}$$

where now $\hat{\mathbf{e}}_i = \hat{\mathbf{e}}_i(\mathbf{x})$. Then the tensor (or dyad) $\nabla\mathbf{V}$ is

$$\begin{aligned}
\nabla\mathbf{V} &= \sum_i \hat{\mathbf{e}}_i \frac{\partial}{\partial x_i} \sum_j V_j \hat{\mathbf{e}}_j \\
&= \sum_i \hat{\mathbf{e}}_i \sum_j \frac{\partial}{\partial x_j} V_j \hat{\mathbf{e}}_j \\
&= \sum_i \sum_j \left(\underbrace{\hat{\mathbf{e}}_i \hat{\mathbf{e}}_j \frac{\partial V_j}{\partial x_i}}_{\text{“Cartesian” part}} + \underbrace{\hat{\mathbf{e}}_i V_j \frac{\partial \hat{\mathbf{e}}_j}{\partial x_i}}_{\text{Extra terms if } \hat{\mathbf{e}}_j = \hat{\mathbf{e}}_j(x_i)} \right).
\end{aligned} \tag{2.35}$$

The first term is just the usual Cartesian derivative. The remaining terms arise in curvilinear coordinates. They must always be accounted for.

For example, in familiar cylindrical (r, θ, z) coordinates, the unit vectors $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_\theta$ are functions of space with the properties

$$\frac{\partial \hat{\mathbf{e}}_r}{\partial \theta} = \hat{\mathbf{e}}_\theta \tag{2.36}$$

and

$$\frac{\partial \hat{\mathbf{e}}_\theta}{\partial \theta} = -\hat{\mathbf{e}}_r. \tag{2.37}$$

Then in these polar coordinates,

$$\begin{aligned}
\nabla \mathbf{V} &= \left(\hat{\mathbf{e}}_r \frac{\partial}{\partial r} + \hat{\mathbf{e}}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} \right) (\hat{\mathbf{e}}_r V_r + \hat{\mathbf{e}}_\theta V_\theta) \\
&= \hat{\mathbf{e}}_r \hat{\mathbf{e}}_r \frac{\partial V_r}{\partial r} + \hat{\mathbf{e}}_r \hat{\mathbf{e}}_\theta \frac{\partial V_\theta}{\partial r} \\
&\quad + \hat{\mathbf{e}}_\theta \hat{\mathbf{e}}_r \frac{1}{r} \frac{\partial V_r}{\partial \theta} + \hat{\mathbf{e}}_\theta \frac{V_r}{r} \frac{\partial \hat{\mathbf{e}}_r}{\partial \theta} \\
&\quad + \hat{\mathbf{e}}_\theta \hat{\mathbf{e}}_\theta \frac{1}{r} \frac{\partial V_\theta}{\partial \theta} + \hat{\mathbf{e}}_\theta \frac{V_\theta}{r} \frac{\partial \hat{\mathbf{e}}_\theta}{\partial \theta} \\
&= \hat{\mathbf{e}}_r \hat{\mathbf{e}}_r \frac{\partial V_r}{\partial r} + \hat{\mathbf{e}}_r \hat{\mathbf{e}}_\theta \frac{\partial V_\theta}{\partial r} \\
&\quad + \hat{\mathbf{e}}_\theta \hat{\mathbf{e}}_r \left(\frac{1}{r} \frac{\partial V_r}{\partial \theta} - \frac{V_\theta}{r} \right) + \hat{\mathbf{e}}_\theta \hat{\mathbf{e}}_\theta \left(\frac{1}{r} \frac{\partial V_\theta}{\partial \theta} + \frac{V_r}{r} \right).
\end{aligned} \tag{2.38}$$

Expressions of the form $\mathbf{U} \cdot \nabla \mathbf{V}$ appear often in MHD. It is a vector that expresses the rate of change of $\mathbf{V}$ in the direction of $\mathbf{U}$. Then, for polar coordinates,

$$\mathbf{U} \cdot \nabla \mathbf{V} = \hat{\mathbf{e}}_r \left(U_r \frac{\partial V_r}{\partial r} + \frac{U_\theta}{r} \frac{\partial V_r}{\partial \theta} - \frac{U_\theta V_\theta}{r} \right) + \hat{\mathbf{e}}_\theta \left(U_r \frac{\partial V_\theta}{\partial r} + \frac{U_\theta}{r} \frac{\partial V_\theta}{\partial \theta} + \frac{U_\theta V_r}{r} \right). \tag{2.39}$$

The third term in each of the brackets is the new terms that arise from the differentiation of the unit vectors in curvilinear coordinates.

Of course, there is no need to insist that the bases $\hat{\mathbf{e}}_i(\mathbf{x})$ even be orthogonal. (An orthogonal system has $\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij}$.) Such systems are called *generalized curvilinear coordinates*.[2] Then the bases $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$ are not unique, because it is always possible to define equivalent, *reciprocal* basis vectors $(\hat{\mathbf{e}}^1, \hat{\mathbf{e}}^2, \hat{\mathbf{e}}^3)$ at each point in space by the process

$$\hat{\mathbf{e}}^3 = \hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_2 / J, \quad \hat{\mathbf{e}}^2 = \hat{\mathbf{e}}_3 \times \hat{\mathbf{e}}_1 / J, \quad \hat{\mathbf{e}}^1 = \hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3 / J, \tag{2.40}$$

[2] This discussion follows that of Donald H. Menzel, *Mathematic Physics*, Dover Publications, New York (1961).

where $J = \hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3$ is called the *Jacobian*. A vector $\mathbf{V}$ can be equivalently expressed as

$$\mathbf{V} = V^i \hat{\mathbf{e}}_i, \tag{2.41}$$

or

$$\mathbf{V} = V_i \hat{\mathbf{e}}^i. \tag{2.42}$$

The V^i are called the *contravariant* components of $\mathbf{V}$, and the V_i are called the *covariant* components. (Of course, the vector $\mathbf{V}$, which is invariant by definition, is neither contravariant or covariant.)

Our previous discussion of vectors, tensors, and dyads can be generalized to these non-orthogonal coordinates, *as long as extreme care is taken in keeping track of the contravariant and covariant components*. Of particular interest is the generalization of vector differentiation, previously discussed for the special case of polar coordinates. The tensor $\nabla\mathbf{V}$ can be written as

$$\nabla\mathbf{V} = \hat{\mathbf{e}}^i \hat{\mathbf{e}}_j D_i V^j, \tag{2.43}$$

where

$$D_i V^j = \partial_i V^j + V^k \Gamma^j_{ik} \tag{2.44}$$

is called the *covariant derivative*. The quantities Γ^j_{ik} are called the *Christoffel symbols* and are defined by

$$\partial_i \hat{\mathbf{e}}^k = -\Gamma^j_{ik} \hat{\mathbf{e}}^k. \tag{2.45}$$

They are the generalization of Eqs. (2.36) and (2.37). [We remark that the Γ^j_{ik} are *not* tensors, as they do not obey the transformation law, Eq. (2.24).] Expressions for the Γ^j_{ik} in any particular coordinate system are given in terms of the metric tensor components $g_{ij} = \hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j$, and $g^{ij} = \hat{\mathbf{e}}^i \cdot \hat{\mathbf{e}}^j$, as

$$\Gamma^k_{ij} = g^{kl} \left(\partial_i g_{il} + \partial_j g_{li} - \partial_l g_{ij} \right). \tag{2.46}$$

We stated previously that if an expression is true in Cartesian coordinates, it is true in all coordinate systems. In particular, expressions for generalized curvilinear coordinates can be obtained by replacing everywhere the derivative $\partial_i V^j$ with the covariant derivative $D_i V^j$, defined by Eq. (2.44). In analogy with the discussion preceding Eq. (2.32), covariant differentiation is often expressed in the shorthand notation $D_i A^{jk} \equiv A^{jk}_{\,..;i}$. Misner, Thorne and Wheeler[3] call this the "comma goes to semi-colon rule" for obtaining tensor expressions in generalized curvilinear coordinates: first get an expression in orthogonal coordinates, and then change all commas to semi-colons!

Generalized curvilinear coordinates play an essential role in the theoretical description of tokamak plasmas. The topic is so detailed and complex (and, frankly, difficult) that it will not be covered further here. I hope this short introduction will allow you to learn more about this on your own.

We now return to Cartesian coordinates.

The divergence of a tensor T has been defined as

$$\nabla \cdot \mathsf{T} = \partial_i T_{ij}. \tag{2.47}$$

It is a vector whose jth component is

$$(\nabla \cdot \mathsf{T})_j = \frac{\partial T_{1j}}{\partial x_1} + \frac{\partial T_{2j}}{\partial x_2} + \frac{\partial T_{3j}}{\partial x_3}. \tag{2.48}$$

Integrate this expression over all space:

$$\begin{aligned}\int (\nabla \cdot \mathsf{T})_j \, d^3x &= \int dx_1 dx_2 dx_3 \left(\frac{\partial T_{1j}}{\partial x_1} + \frac{\partial T_{2j}}{\partial x_2} + \frac{\partial T_{3j}}{\partial x_3} \right) \\ &= \int dx_2 dx_3 T_{1j} + \int dx_1 dx_3 T_{2j} + \int dx_1 dx_2 T_{3j} \\ &= \int dS_1 T_{1j} + \int dS_2 T_{2j} + \int dS_3 T_{3j} \\ &= \int (d\mathbf{S} \cdot \mathsf{T})_j,\end{aligned}$$

or

$$\int \nabla \cdot \mathsf{T} d^3x = \oint d\mathbf{S} \cdot \mathsf{T}. \tag{2.49}$$

This is the *generalized Gauss' theorem.*

It is also possible to derive the following integral theorems:

$$\int \nabla \mathbf{V} d^3x = \oint d\mathbf{S}\mathbf{V} \tag{2.50}$$

$$\int \nabla f d^3x = \oint d\mathbf{S} f \tag{2.51}$$

[3] C. W. Misner, K. S. Thorn, and J. A. Wheeler, *Gravitation*, W. H. Freeman and Company, San Francisco (1973).

$$\int \nabla \times \mathbf{V} d^3x = \oint d\mathbf{S} \times \mathbf{V} \tag{2.52}$$

$$\int_S d\mathbf{S} \times \nabla f d^3x = \oint_C d\mathbf{l} f \tag{2.53}$$

$$\int_S d\mathbf{S} \cdot \nabla \times \mathbf{A} d^3x = \oint_C \mathbf{A} \cdot d\mathbf{l} \tag{2.54}$$

It seems intuitive that a physically measurable quantity should not care what coordinate system it is referred to. We have shown that scalars, vectors, tensors, etc., things that we can associate with physical quantities, are invariant under rotations. This is good!

There is another important type of coordinate transformation called an *inversion*. It is also known as *parity transformation*. Mathematically, this is given by

$$x'_i = -x_i. \tag{2.55}$$

An inversion is shown in Fig. 2.4.

The first coordinate system is "right-handed"; the second coordinate system is "left-handed." Consider the position vector $\mathbf{r}$. In the unprimed coordinate system, it is given by $\mathbf{r} = x_i \hat{\mathbf{e}}_i$. In the primed (inverted) coordinate system, it is given by

$$\mathbf{r}' = x'_i \hat{\mathbf{e}}'_i = (-x_i)(-\hat{\mathbf{e}}_i) = x_i \hat{\mathbf{e}}_i = \mathbf{r}, \tag{2.56}$$

so it is invariant under inversions. Such a vector is called a*polar vector*. (It is sometimes called a *true* vector.) We remark the gradient operator $\nabla = \hat{\mathbf{e}}_i \partial_i$ transforms like a polar vector, since $\partial'_i = -\partial_i$.

Now consider the vector $\mathbf{C}$, defined by $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ (or $C_i = \varepsilon_{ijk} A_j B_k$), where $\mathbf{A}$ and $\mathbf{B}$ are polar vectors, i.e., $\mathbf{A}$ and $\mathbf{B}$ transform according to $A'_i = -A_i$ and $B'_i = -B_i$. Then under inversion, the components of $\mathbf{C}$ transform according to

$$C'_i = \varepsilon_{ijk} A'_j B'_k = \varepsilon_{ijk} \left(-A_j\right)(-B_k) = \varepsilon_{ijk} A_j B_k = +C_i. \tag{2.57}$$

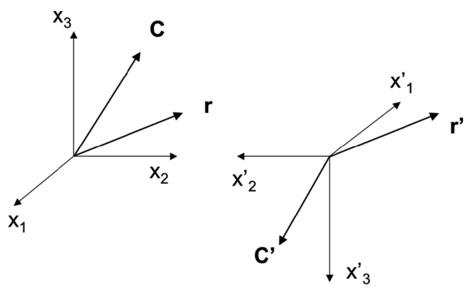

Fig. 2.4 An inversion of the Cartesian coordinate system

Then

$$\mathbf{C}' = C_i'\hat{\mathbf{e}}_i' = C_i\,(-\hat{\mathbf{e}}_i) = -\mathbf{C}, \tag{2.58}$$

so that $\mathbf{C}$ does *not* transform like a true vector under coordinate inversions. Such a vector (that changes direction in space under coordinate inversion; see Fig. 2.4) is called an *axial vector* (or *pseudovector*). Since it is defined as a vector (or cross) product, it usually describes some process involving rotation. For example, the vector area $d\mathbf{S}_k = d\mathbf{x}_i \times d\mathbf{x}_j$ is a pseudovector. However, notice that if $\mathbf{A}$ is a polar vector and $\mathbf{B}$ is an axial vector, then $C_i' = -C_i$, and $\mathbf{C}$ is a polar vector.

The elementary volume is defined as $dV = d\mathbf{x}_1 \cdot d\mathbf{x}_2 \times d\mathbf{x}_3 = \varepsilon_{ijk}dx_1dx_2dx_3$. It is easy to see that under inversions, $dV' = -dV$; the volume changes sign! Such quantities are called *pseudoscalars*: they are invariant under rotations, but change sign under inversions.

Again, it is intuitive that physical quantities should exist independent from coordinate systems. How then to account for the volume? Apparently it should be considered a "derived quantity," not directly measurable. For example, one can measure directly the true vectors (i.e., lengths) $d\mathbf{x}_i$, but one has to *compute* the volume. This renders as a pseudoscalar any quantity that expresses an amount of a scalar quantity per unit volume; these are not directly measurable. This includes the mass density ρ (mass/volume) and the pressure (internal energy/volume). (An exception is the electric charge density, which is a true scalar.) Apparently, one can measure directly mass and length (both true scalars), but must then infer the mass density.

The following is a list of some physical variables that appear in MHD, and their transformation properties:

Time is a scalar.
Temperature, which has units of energy, is a scalar.
Mass density, $\rho = M/V$, is a pseudoscalar.
Pressure, $p = \rho k_B T$, is a pseudoscalar.
Velocity, $V_i = dx_i/dt$, is a vector.
The vector potential $\mathbf{A}$ is a vector.
The magnetic flux, $\int \mathbf{A} \cdot d\mathbf{x}$, is a scalar.
The magnetic field, $\mathbf{B} = \nabla \times \mathbf{A}$, is a pseudovector.
The current density, $\mu_0\mathbf{J} = \nabla \times \mathbf{B}$, is a vector. (Note that, since J is electric charge per unit area, and area is a pseudoscalar, electric charge must be a pseudoscalar.)
The Lorentz force density, $\mathbf{f}_L = \mathbf{J} \times \mathbf{B}$, is a pseudovector.
The pressure force density, $-\nabla p$, is a pseudovector.
The acceleration density, $\rho d\mathbf{V}/dt$, is a pseudovector.

Now, it is OK to express physical relationships by using pseudovectors and pseudoscalars. What is required is that the resulting expressions be consistent, i.e., we do not end up adding scalars and pseudoscalars, or vectors and pseudovectors.

Springer

http://www.springer.com/978-3-642-00687-6

Lectures in Magnetohydrodynamics
With an Appendix on Extended MHD
Schnack, D.D.
2009, XV, 323 p. 119 illus., Hardcover
ISBN: 978-3-642-00687-6