Cambridge University Press
978-0-521-89697-9 — Modern Electrodynamics
Andrew Zangwill
Excerpt
More Information

1 Mathematical Preliminaries

The enormous usefulness of mathematics in the natural sciences is something bordering on the mysterious.
Eugene Wigner (1960)

1.1 Introduction

This chapter presents a collection of mathematical notation, definitions, identities, theorems, and transformations that play an important role in the study of electromagnetism. A brief discussion accompanies some of the less familiar topics and only a few proofs are given in detail. For more details and complete proofs, the reader should consult the books and papers listed in Sources, References, and Additional Reading at the end of the chapter. Appendix C at the end of the book summarizes the properties of Legendre polynomials, spherical harmonics, and Bessel functions.

1.2 Vectors

A vector is a geometrical object characterized by a magnitude and direction.[1] Although not necessary, it is convenient to discuss an arbitrary vector using its components defined with respect to a given coordinate system. An example is the right-handed coordinate system with orthogonal unit basis vectors $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$ shown in Figure 1.1, where

$$\hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_1 = 1 \qquad \hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_2 = 1 \qquad \hat{\mathbf{e}}_3 \cdot \hat{\mathbf{e}}_3 = 1 \tag{1.1}$$

$$\hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_2 = 0 \qquad \hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_3 = 0 \qquad \hat{\mathbf{e}}_3 \cdot \hat{\mathbf{e}}_1 = 0 \tag{1.2}$$

$$\hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_2 = \hat{\mathbf{e}}_3 \qquad \hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3 = \hat{\mathbf{e}}_1 \qquad \hat{\mathbf{e}}_3 \times \hat{\mathbf{e}}_1 = \hat{\mathbf{e}}_2. \tag{1.3}$$

We express an arbitrary vector $\mathbf{V}$ in this basis using components $V_k = \hat{\mathbf{e}}_k \cdot \mathbf{V}$,

$$\mathbf{V} = V_1\hat{\mathbf{e}}_1 + V_2\hat{\mathbf{e}}_2 + V_3\hat{\mathbf{e}}_3. \tag{1.4}$$

A vector can be decomposed in any coordinate system we please, so

$$\sum_{k=1}^{3} V_k\hat{\mathbf{e}}_k = \sum_{k=1}^{3} V'_k\hat{\mathbf{e}}'_k. \tag{1.5}$$

[1] A more precise definition of a vector is given in Section 1.8.

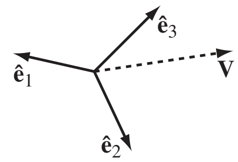

Figure 1.1: An orthonormal set of unit vectors $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3$. $\mathbf{V}$ is an arbitrary vector.

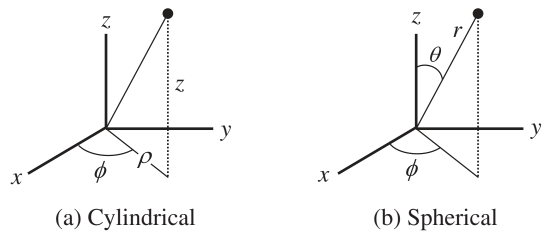

Figure 1.2: Two curvilinear coordinate systems.

1.2.1 Cartesian Coordinates

Our notation for Cartesian components and unit vectors is

$$\mathbf{V} = V_x\hat{\mathbf{x}} + V_y\hat{\mathbf{y}} + V_z\hat{\mathbf{z}}. \tag{1.6}$$

In particular, r_k always denotes the Cartesian components of the position vector,

$$\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}. \tag{1.7}$$

It is not obvious geometrically (see Example 1.7 in Section 1.8), but the gradient operator is a vector with the Cartesian representation

$$\nabla = \hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}. \tag{1.8}$$

The divergence, curl, and Laplacian operations are, respectively,

$$\nabla\cdot\mathbf{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z} \tag{1.9}$$

$$\nabla\times\mathbf{V} = \left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}\right)\hat{\mathbf{x}} + \left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right)\hat{\mathbf{y}} + \left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right)\hat{\mathbf{z}} \tag{1.10}$$

$$\nabla^2 A = \frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} + \frac{\partial^2 A}{\partial z^2}. \tag{1.11}$$

1.2.2 Cylindrical Coordinates

Figure 1.2(a) defines cylindrical coordinates (ρ, ϕ, z). Our notation for the components and unit vectors in this system is

$$\mathbf{V} = V_\rho\hat{\boldsymbol{\rho}} + V_\phi\hat{\boldsymbol{\phi}} + V_z\hat{\mathbf{z}}. \tag{1.12}$$

The transformation to Cartesian coordinates is

$$x = \rho\cos\phi \qquad y = \rho\sin\phi \qquad z = z. \tag{1.13}$$

The volume element in cylindrical coordinates is $d^3r = \rho d\rho d\phi dz$. The unit vectors $(\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\phi}}, \hat{\mathbf{z}})$ form a right-handed orthogonal triad. $\hat{\mathbf{z}}$ is the same as in Cartesian coordinates. Otherwise,

$$\hat{\boldsymbol{\rho}} = \hat{\mathbf{x}} \cos\phi + \hat{\mathbf{y}} \sin\phi \qquad \hat{\mathbf{x}} = \hat{\boldsymbol{\rho}} \cos\phi - \hat{\boldsymbol{\phi}} \sin\phi \tag{1.14}$$

$$\hat{\boldsymbol{\phi}} = -\hat{\mathbf{x}} \sin\phi + \hat{\mathbf{y}} \cos\phi \qquad \hat{\mathbf{y}} = \hat{\boldsymbol{\rho}} \sin\phi + \hat{\boldsymbol{\phi}} \cos\phi. \tag{1.15}$$

The gradient operator in cylindrical coordinates is

$$\nabla = \hat{\boldsymbol{\rho}} \frac{\partial}{\partial\rho} + \frac{\hat{\boldsymbol{\phi}}}{\rho} \frac{\partial}{\partial\phi} + \hat{\mathbf{z}} \frac{\partial}{\partial z}. \tag{1.16}$$

The divergence, curl, and Laplacian operations are, respectively,

$$\nabla \cdot \mathbf{V} = \frac{1}{\rho} \frac{\partial(\rho V_\rho)}{\partial\rho} + \frac{1}{\rho} \frac{\partial V_\phi}{\partial\phi} + \frac{\partial V_z}{\partial z} \tag{1.17}$$

$$\nabla \times \mathbf{V} = \left[\frac{1}{\rho} \frac{\partial V_z}{\partial\phi} - \frac{\partial V_\phi}{\partial z} \right] \hat{\boldsymbol{\rho}} + \left[\frac{\partial V_\rho}{\partial z} - \frac{\partial V_z}{\partial\rho} \right] \hat{\boldsymbol{\phi}} + \frac{1}{\rho} \left[\frac{\partial(\rho V_\phi)}{\partial\rho} - \frac{\partial V_\rho}{\partial\phi} \right] \hat{\mathbf{z}} \tag{1.18}$$

$$\nabla^2 A = \frac{1}{\rho} \frac{\partial}{\partial\rho} \left(\rho \frac{\partial A}{\partial\rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 A}{\partial\phi^2} + \frac{\partial^2 A}{\partial z^2}. \tag{1.19}$$

1.2.3 Spherical Coordinates

Figure 1.2(b) defines spherical coordinates (r, θ, ϕ). Our notation for the components and unit vectors in this system is

$$\mathbf{V} = V_r \hat{\mathbf{r}} + V_\theta \hat{\boldsymbol{\theta}} + V_\phi \hat{\boldsymbol{\phi}}. \tag{1.20}$$

The transformation to Cartesian coordinates is

$$x = r \sin\theta \cos\phi \qquad y = r \sin\theta \sin\phi \qquad z = r \cos\theta. \tag{1.21}$$

The volume element in spherical coordinates is $d^3r = r^2 \sin\theta dr d\theta d\phi$. The unit vectors are related by

$$\hat{\mathbf{r}} = \hat{\mathbf{x}} \sin\theta \cos\phi + \hat{\mathbf{y}} \sin\theta \sin\phi + \hat{\mathbf{z}} \cos\theta \qquad \hat{\mathbf{x}} = \hat{\mathbf{r}} \sin\theta \cos\phi + \hat{\boldsymbol{\theta}} \cos\theta \cos\phi - \hat{\boldsymbol{\phi}} \sin\phi \tag{1.22}$$

$$\hat{\boldsymbol{\theta}} = \hat{\mathbf{x}} \cos\theta \cos\phi + \hat{\mathbf{y}} \cos\theta \sin\phi - \hat{\mathbf{z}} \sin\theta \qquad \hat{\mathbf{y}} = \hat{\mathbf{r}} \sin\theta \sin\phi + \hat{\boldsymbol{\theta}} \cos\theta \sin\phi + \hat{\boldsymbol{\phi}} \cos\phi \tag{1.23}$$

$$\hat{\boldsymbol{\phi}} = -\hat{\mathbf{x}} \sin\phi + \hat{\mathbf{y}} \cos\phi \qquad \hat{\mathbf{z}} = \hat{\mathbf{r}} \cos\theta - \hat{\boldsymbol{\theta}} \sin\theta. \tag{1.24}$$

The gradient operator in spherical coordinates is

$$\nabla = \hat{\mathbf{r}} \frac{\partial}{\partial r} + \frac{\hat{\boldsymbol{\theta}}}{r} \frac{\partial}{\partial\theta} + \frac{\hat{\boldsymbol{\phi}}}{r \sin\theta} \frac{\partial}{\partial\phi}. \tag{1.25}$$

The divergence, curl, and Laplacian operations are, respectively,

$$\nabla \cdot \mathbf{V} = \frac{1}{r^2} \frac{\partial(r^2 V_r)}{\partial r} + \frac{1}{r \sin\theta} \frac{\partial(\sin\theta\, V_\theta)}{\partial\theta} + \frac{1}{r \sin\theta} \frac{\partial V_\phi}{\partial\phi} \tag{1.26}$$

$$\begin{aligned} \nabla \times \mathbf{V} = {} & \frac{1}{r \sin\theta} \left[\frac{\partial(\sin\theta\, V_\phi)}{\partial\theta} - \frac{\partial V_\theta}{\partial\phi} \right] \hat{\mathbf{r}} \\ & + \frac{1}{r} \left[\frac{1}{\sin\theta} \frac{\partial V_r}{\partial\phi} - \frac{\partial(r V_\phi)}{\partial r} \right] \hat{\boldsymbol{\theta}} + \frac{1}{r} \left[\frac{\partial(r V_\theta)}{\partial r} - \frac{\partial V_r}{\partial\theta} \right] \hat{\boldsymbol{\phi}} \end{aligned} \tag{1.27}$$

$$\nabla^2 A = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial A}{\partial r} \right) + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial A}{\partial\theta} \right) + \frac{1}{r^2 \sin^2\theta} \frac{\partial^2 A}{\partial\phi^2}. \tag{1.28}$$

1.2.4 The Einstein Summation Convention

Einstein (1916) introduced the following convention. An index which appears exactly twice in a single term of a mathematical expression is implicitly summed over all possible values for that index. The range of this *dummy index* must be clear from context and the index cannot be used elsewhere in the same expression for another purpose. In this book, the range for a roman index like i is from 1 to 3, indicating a sum over the Cartesian indices x, y, and z. Thus, $\mathbf{V}$ in (1.6) and its dot product with another vector $\mathbf{F}$ are written

$$\mathbf{V} = \sum_{k=1}^{3} V_k \hat{\mathbf{e}}_k \equiv V_k \hat{\mathbf{e}}_k \qquad \mathbf{V}\cdot\mathbf{F} = \sum_{k=1}^{3} V_k F_k \equiv V_k F_k. \tag{1.29}$$

In a Cartesian basis, the gradient of a scalar φ and the divergence of a vector $\mathbf{D}$ can be variously written

$$\nabla\varphi = \hat{\mathbf{e}}_k \nabla_k \varphi = \hat{\mathbf{e}}_k \partial_k \varphi = \hat{\mathbf{e}}_k \frac{\partial \varphi}{\partial r_k} \tag{1.30}$$

$$\nabla\cdot\mathbf{D} = \nabla_k D_k = \partial_k D_k = \frac{\partial D_k}{\partial r_k}. \tag{1.31}$$

If an $N \times N$ matrix $\mathbf{C}$ is the product of an $N \times M$ matrix $\mathbf{A}$ and an $M \times N$ matrix $\mathbf{B}$,

$$C_{ik} = \sum_{j=1}^{M} A_{ij} B_{jk} = A_{ij} B_{jk}. \tag{1.32}$$

1.2.5 The Kronecker and Levi-Cività Symbols

The Kronecker delta symbol δ_{ij} and Levi-Cività permutation symbol ϵ_{ijk} have roman indices i, j, and k which take on the Cartesian coordinate values x, y, and z. They are defined by

$$\delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j, \end{cases} \tag{1.33}$$

and

$$\epsilon_{ijk} = \begin{cases} 1 & ijk = xyz \quad yzx \quad zxy, \\ -1 & ijk = xzy \quad yxz \quad zyx, \\ 0 & \text{otherwise.} \end{cases} \tag{1.34}$$

Some useful Kronecker delta and Levi-Cività symbol identities are

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij} \qquad \delta_{kk} = 3 \tag{1.35}$$

$$\partial_k r_j = \delta_{jk} \qquad V_k \delta_{kj} = V_j \tag{1.36}$$

$$[\mathbf{V}\times\mathbf{F}]_i = \epsilon_{ijk} V_j F_k \qquad [\nabla\times\mathbf{A}]_i = \epsilon_{ijk}\partial_j A_k \tag{1.37}$$

$$\delta_{ij}\epsilon_{ijk} = 0 \qquad \epsilon_{ijk}\epsilon_{ijk} = 6. \tag{1.38}$$

A particularly useful identity involves a single sum over the repeated index i:

$$\epsilon_{ijk}\epsilon_{ist} = \delta_{js}\delta_{kt} - \delta_{jt}\delta_{ks}. \tag{1.39}$$

A generalization of (1.39) when there are no repeated indices to sum over is the determinant

$$\epsilon_{ki\ell}\,\epsilon_{mpq} = \begin{vmatrix} \delta_{km} & \delta_{im} & \delta_{\ell m} \\ \delta_{kp} & \delta_{ip} & \delta_{\ell p} \\ \delta_{kq} & \delta_{iq} & \delta_{\ell q} \end{vmatrix}. \tag{1.40}$$

Finally, let **C** be a 3 × 3 matrix with matrix elements C_{11}, C_{12}, etc. The determinant of **C** can be written using either an expansion by columns,

$$\det \mathbf{C} = \epsilon_{ijk} C_{i1} C_{j2} C_{k3}, \tag{1.41}$$

or an expansion by rows,

$$\det \mathbf{C} = \epsilon_{ijk} C_{1i} C_{2j} C_{3k}. \tag{1.42}$$

A closely related identity we will use in Section 1.8.1 is

$$\epsilon_{\ell mn} \det \mathbf{C} = \epsilon_{ijk} C_{\ell i} C_{mj} C_{nk}. \tag{1.43}$$

1.2.6 Vector Identities in Cartesian Components

The Kronecker and Levi-Cività symbols simplify the proof of vector identities. An example is

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}). \tag{1.44}$$

Using the left side of (1.37), the ith component of $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is

$$[\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]_i = \epsilon_{ijk} a_j (\mathbf{b} \times \mathbf{c})_k = \epsilon_{ijk} a_j \epsilon_{k\ell m} b_l c_m. \tag{1.45}$$

The definition (1.34) tells us that $\epsilon_{ijk} = \epsilon_{kij}$. Therefore, the identity (1.39) gives

$$[\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]_i = \epsilon_{kij} \epsilon_{k\ell m} a_j b_\ell c_m = (\delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell}) a_j b_\ell c_m = a_j b_i c_j - a_j b_j c_i. \tag{1.46}$$

The final result, $b_i(\mathbf{a} \cdot \mathbf{c}) - c_i(\mathbf{a} \cdot \mathbf{b})$, is indeed the ith component of the right side of (1.44). The same method of proof applies to gradient-, divergence-, and curl-type vector identities because the components of the ∇ operator transform like the components of a vector [see above (1.8)]. The next three examples illustrate this point.

Example 1.1 Prove that $\nabla \cdot (\nabla \times \mathbf{g}) = 0$.

Solution: Begin with $\nabla \cdot (\nabla \times \mathbf{g}) = \partial_i \epsilon_{ijk} \partial_j g_k = \frac{1}{2} \partial_i \partial_j g_k \epsilon_{ijk} + \frac{1}{2} \partial_i \partial_j g_k \epsilon_{ijk}$. Exchanging the dummy indices i and j in the last term gives

$$\nabla \cdot \nabla \times \mathbf{g} = \tfrac{1}{2} \partial_i \partial_j g_k \epsilon_{ijk} + \tfrac{1}{2} \partial_j \partial_i g_k \epsilon_{jik} = \tfrac{1}{2} \{\epsilon_{ijk} + \epsilon_{jik}\} \partial_i \partial_j g_k = 0.$$

The final zero comes from $\epsilon_{ijk} = -\epsilon_{jik}$, which is a consequence of (1.34).

Example 1.2 Prove that $\nabla \times (\mathbf{A} \times \mathbf{B}) = \mathbf{A} \nabla \cdot \mathbf{B} - (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A} - \mathbf{B} \nabla \cdot \mathbf{A}$.

Solution: Focus on the ith Cartesian component and use the left side of (1.37) to write

$$[\nabla \times (\mathbf{A} \times \mathbf{B})]_i = \epsilon_{ijk} \partial_j (\mathbf{A} \times \mathbf{B})_k = \epsilon_{ijk} \epsilon_{kst} \partial_j (A_s B_t).$$

The cyclic properties of the Levi-Cività symbol and the identity (1.39) give

$$[\nabla \times (\mathbf{A} \times \mathbf{B})]_i = \epsilon_{kij} \epsilon_{kst} \partial_j (A_s B_t) = (\delta_{is}\delta_{jt} - \delta_{it}\delta_{js})(A_s \partial_j B_t + B_t \partial_j A_s).$$

Therefore,

$$[\nabla \times (\mathbf{A} \times \mathbf{B})]_i = A_i \partial_j B_j - A_j \partial_j B_i + B_j \partial_j A_i - B_i \partial_j A_j.$$

This proves the identity because the choice of i is arbitrary.

Example 1.3 Prove the "double-curl identity" $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$.

Solution: Consider the ith Cartesian component. The identity on the left side of (1.37) and the invariance of the Levi-Cività symbol with respect to cyclic permutations of its indices give

$$[\nabla \times (\nabla \times \mathbf{A})]_i = \epsilon_{ijk}\partial_j(\nabla \times \mathbf{A})_k = \epsilon_{ijk}\partial_j \epsilon_{kpq}\partial_p A_q = \epsilon_{kij}\epsilon_{kpq}\partial_j \partial_p A_q.$$

Now apply the identity (1.39) to get

$$[\nabla \times (\nabla \times \mathbf{A})]_i = (\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp})\partial_j \partial_p A_q = \partial_i \partial_j A_j - \partial_j \partial_j A_i = \nabla_i(\nabla \cdot \mathbf{A}) - \nabla^2 A_i.$$

The double-curl identity follows because

$$\nabla^2 \mathbf{A} = \nabla^2(A_x \hat{\mathbf{x}} + A_y \hat{\mathbf{y}} + A_z \hat{\mathbf{z}}) = \hat{\mathbf{x}}\nabla^2 A_x + \hat{\mathbf{y}}\nabla^2 A_y + \hat{\mathbf{z}}\nabla^2 A_z.$$

1.2.7 Vector Identities in Curvilinear Components

Care is needed to interpret the vector identities in Examples 1.2 and 1.3 when the vectors in question are decomposed into spherical or cylindrical components such as $\mathbf{A} = A_r \hat{\mathbf{r}} + A_\theta \hat{\boldsymbol{\theta}} + A_\phi \hat{\boldsymbol{\phi}}$. This can be seen from Example 1.3 where the final step is no longer valid because $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\phi}}$ are not constant vectors. In other words,

$$\nabla^2 \mathbf{A} \;=\; \nabla \cdot \nabla(A_r \hat{\mathbf{r}} + A_\theta \hat{\boldsymbol{\theta}} + A_\phi \hat{\boldsymbol{\phi}}) \neq \hat{\mathbf{r}}\nabla^2 A_r + \hat{\boldsymbol{\theta}}\nabla^2 A_\theta + \hat{\boldsymbol{\phi}}\nabla^2 A_\phi. \tag{1.47}$$

One way to proceed is to work out the components of $\nabla(A_r \hat{\mathbf{r}})$, $\nabla(A_\theta \hat{\boldsymbol{\theta}})$, and $\nabla(A_\phi \hat{\boldsymbol{\phi}})$. Alternatively, we may simply *define* the meaning of the operation $\nabla^2 \mathbf{A}$ when $\mathbf{A}$ is expressed using curvilinear components. For example,

$$[\nabla^2 \mathbf{A}]_\phi \equiv \partial_\phi(\nabla \cdot \mathbf{A}) - [\nabla \times (\nabla \times \mathbf{A})]_\phi, \tag{1.48}$$

and similarly for $(\nabla^2 \mathbf{A})_r$ and $(\nabla^2 \mathbf{A})_\theta$.

Exactly the same issue arises when we examine the last step in Example 1.2, namely

$$[\nabla \times (\mathbf{A} \times \mathbf{B})]_i = A_i \nabla \cdot \mathbf{B} - (\mathbf{A} \cdot \nabla) B_i + (\mathbf{B} \cdot \nabla) A_i - B_i \nabla \cdot \mathbf{A}. \tag{1.49}$$

By construction, this equation makes sense when i stands for x, y, or z. It does *not* make sense if i stands for, say, r, θ, or ϕ. On the other hand, the full vector version of the identity is correct as long as we retain the r, θ, and ϕ variations of $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\phi}}$. For example,

$$(\mathbf{A} \cdot \nabla)\mathbf{B} = \left[A_r \frac{\partial}{\partial r} + \frac{A_\theta}{r}\frac{\partial}{\partial \theta} + \frac{A_\phi}{r \sin\theta}\frac{\partial}{\partial \phi} \right] (B_r \hat{\mathbf{r}} + B_\theta \hat{\boldsymbol{\theta}} + B_\phi \hat{\boldsymbol{\phi}}). \tag{1.50}$$

Application 1.1 Two Identities for $\nabla \times \mathbf{L}$

The $\hbar = 1$ version of the quantum mechanical angular momentum operator, $\mathbf{L} = -i\mathbf{r} \times \nabla$, plays a useful role in the analysis of classical spherical systems. In this Application, we prove two operator identities which will appear later in the text:

(A) $\quad \nabla \times \mathbf{L} = -i\mathbf{r}\nabla^2 + i\nabla(1 + \mathbf{r} \cdot \nabla)$

(B) $\quad \nabla \times \mathbf{L} = (\hat{\mathbf{r}} \times \mathbf{L})\left(\frac{1}{r}\frac{\partial}{\partial r} r\right) + \hat{\mathbf{r}}\frac{i}{r}L^2.$

Proof of Identity (A):

We use (1.37), (1.39), and the cyclic property of the Levi-Cività symbol to evaluate the kth component of $\nabla \times \mathbf{L}$ acting on a scalar function ϕ:

$$[\nabla \times \mathbf{L}]_k \phi = -i[\nabla \times (\mathbf{r} \times \nabla)]_k \phi = -i\epsilon_{mk\ell}\epsilon_{mst}\partial_\ell r_s \partial_t \phi = -i[\partial_\ell r_k \partial_\ell \phi - \partial_\ell r_\ell \partial_k \phi]. \tag{1.51}$$

Because $\partial_\ell r_\ell = 3$ and $\partial_\ell r_k = \delta_{\ell k}$,

$$\nabla \times \mathbf{L}\phi = [-i\mathbf{r}\nabla^2 + i2\nabla + i(\mathbf{r} \cdot \nabla)\nabla]\phi. \tag{1.52}$$

However,

$$\partial_k [r_\ell \partial_\ell \phi] = \partial_k \phi + r_\ell \partial_\ell \partial_k \phi, \tag{1.53}$$

which is the kth component of $\nabla(\mathbf{r} \cdot \nabla)\phi = \nabla\phi + (\mathbf{r} \cdot \nabla)\nabla\phi$. Substituting the latter into (1.53) gives Identity (A).

Proof of Identity (B):

We decompose the gradient operator into its radial and angular pieces:

$$\nabla = \hat{\mathbf{r}}(\hat{\mathbf{r}} \cdot \nabla) - \hat{\mathbf{r}} \times (\hat{\mathbf{r}} \times \nabla) = \hat{\mathbf{r}}\frac{\partial}{\partial r} - \frac{i}{r}\hat{\mathbf{r}} \times \mathbf{L}. \tag{1.54}$$

Equation (1.54) and the Levi-Cività formalism produce the intermediate result

$$\nabla \times \mathbf{L} = (\hat{\mathbf{r}} \times \mathbf{L})\frac{\partial}{\partial r} - \frac{i}{r}(\hat{\mathbf{r}} \times \mathbf{L}) \times \mathbf{L} = (\hat{\mathbf{r}} \times \mathbf{L})\frac{\partial}{\partial r} - \frac{i}{r}\left[\hat{r}_k \mathbf{L} L_k - \hat{\mathbf{r}} L^2\right]. \tag{1.55}$$

However, the angular momentum operator obeys commutation relations which can be summarized as $\mathbf{L} \times \mathbf{L} = i\mathbf{L}$. Therefore,

$$\hat{\mathbf{r}} \times (\mathbf{L} \times \mathbf{L}) = i\hat{\mathbf{r}} \times \mathbf{L} \qquad \Rightarrow \qquad \hat{r}_k \mathbf{L} L_k - \hat{r}_k L_k \mathbf{L} = i\hat{\mathbf{r}} \times \mathbf{L}. \tag{1.56}$$

On the other hand, $r_k L_k = 0$ because $\mathbf{L}$ is perpendicular to both $\mathbf{r}$ and ∇. Therefore, $\hat{r}_k \mathbf{L} L_k = i\hat{\mathbf{r}} \times \mathbf{L}$, which we can substitute into (1.55). The result is identity (B) because, for any scalar function ϕ,

$$\frac{1}{r}\left[\frac{\partial}{\partial r}(r\phi)\right] = \frac{\partial \phi}{\partial r} + \frac{\phi}{r}. \tag{1.57}$$

■

1.3 Derivatives

1.3.1 Functions of r and |r|

The position vector is $\mathbf{r} = r\hat{\mathbf{r}}$ with $r = \sqrt{x^2 + y^2 + z^2}$. If $f(r)$ is a scalar function and $f'(r) = df/dr$,

$$\nabla r = \hat{\mathbf{r}} \qquad \nabla \times \mathbf{r} = 0 \tag{1.58}$$

$$\nabla f = f'\hat{\mathbf{r}} \qquad \nabla^2 f = \frac{(r^2 f')'}{r^2} \tag{1.59}$$

$$\nabla \cdot (f\mathbf{r}) = \frac{(r^3 f)'}{r^2} \qquad \nabla \times (f\mathbf{r}) = 0. \tag{1.60}$$

Similarly, if $\mathbf{g}(r)$ is a vector function and $\mathbf{c}$ is a constant vector,

$$\nabla \cdot \mathbf{g} = \mathbf{g}' \cdot \hat{\mathbf{r}} \qquad \nabla \times \mathbf{g} = \hat{\mathbf{r}} \times \mathbf{g}' \tag{1.61}$$

$$(\mathbf{g} \cdot \nabla)\mathbf{r} = \mathbf{g} \qquad (\mathbf{r} \cdot \nabla)\mathbf{g} = r\mathbf{g}' \tag{1.62}$$

$$\nabla(\mathbf{r}\cdot\mathbf{g}) = \mathbf{g} + \frac{(\mathbf{r}\cdot\mathbf{g}')\mathbf{r}}{r} \qquad \nabla\cdot(\mathbf{g}\times\mathbf{r}) = 0 \tag{1.63}$$

$$\nabla\times(\mathbf{g}\times\mathbf{r}) = 2\mathbf{g} + r\mathbf{g}' - \frac{(\mathbf{r}\cdot\mathbf{g}')\mathbf{r}}{r} \qquad \nabla(\mathbf{c}\cdot\mathbf{r}) = \mathbf{c}. \tag{1.64}$$

1.3.2 Functions of $\mathbf{r} - \mathbf{r}'$

Let $\mathbf{R} = \mathbf{r} - \mathbf{r}' = (x - x')\hat{\mathbf{x}} + (y - y')\hat{\mathbf{y}} + (z - z')\hat{\mathbf{z}}$. Then,

$$\nabla f(R) = f'(R)\hat{\mathbf{R}} \qquad \nabla\cdot\mathbf{g}(R) = \mathbf{g}'(R)\cdot\hat{\mathbf{R}} \qquad \nabla\times\mathbf{g}(R) = \hat{\mathbf{R}}\times\mathbf{g}'(R). \tag{1.65}$$

Moreover, because

$$\nabla = \hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z} \qquad \text{and} \qquad \nabla' = \hat{\mathbf{x}}\frac{\partial}{\partial x'} + \hat{\mathbf{y}}\frac{\partial}{\partial y'} + \hat{\mathbf{z}}\frac{\partial}{\partial z'}, \tag{1.66}$$

it it straightforward to confirm that

$$\nabla' f(R) = -\nabla f(R). \tag{1.67}$$

1.3.3 The Convective Derivative

Let $\phi(\mathbf{r}, t)$ be a scalar function of space and time. An observer who repeatedly samples the value of ϕ at a fixed point in space, $\mathbf{r}$, records the time rate of change of ϕ as the partial derivative $\partial\phi/\partial t$. However, the same observer who repeatedly samples ϕ along a trajectory in space $\mathbf{r}(t)$ that moves with velocity $\boldsymbol{v}(t) = \dot{\mathbf{r}}(t)$ records the time rate of change of ϕ as the *convective derivative*,

$$\frac{d\phi}{dt} = \frac{\partial\phi}{\partial t} + \frac{dx}{dt}\frac{\partial\phi}{\partial x} + \frac{dy}{dt}\frac{\partial\phi}{\partial y} + \frac{dz}{dt}\frac{\partial\phi}{\partial z} = \frac{\partial\phi}{\partial t} + (\boldsymbol{v}\cdot\nabla)\phi. \tag{1.68}$$

For a vector function $\mathbf{g}(\mathbf{r}, t)$, the corresponding convective derivative is

$$\frac{d\mathbf{g}}{dt} = \frac{\partial\mathbf{g}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\mathbf{g}. \tag{1.69}$$

1.3.4 Taylor's Theorem

Taylor's theorem in one dimension is

$$f(x) = f(a) + (x-a)\left.\frac{df}{dx}\right|_{x=a} + \frac{1}{2!}(x-a)^2\left.\frac{d^2 f}{dx^2}\right|_{x=a} + \cdots. \tag{1.70}$$

An alternative form follows from (1.70) if $x \to x + \epsilon$ and $a \to x$:

$$f(x+\epsilon) = f(x) + \epsilon\frac{df}{dx} + \frac{1}{2!}\epsilon^2\frac{d^2 f}{dx^2} + \cdots. \tag{1.71}$$

Equivalently,

$$f(x+\epsilon) = \left[1 + \epsilon\frac{d}{dx} + \frac{1}{2!}\left(\epsilon\frac{d}{dx}\right)^2 + \cdots\right] f(x) = \exp\left(\epsilon\frac{d}{dx}\right) f(x). \tag{1.72}$$

This generalizes for a function of three variables to

$$f(x+\epsilon_x, y+\epsilon_y, z+\epsilon_z) = \exp\left(\epsilon_x\frac{\partial}{\partial x}\right)\exp\left(\epsilon_y\frac{\partial}{\partial y}\right)\exp\left(\epsilon_z\frac{\partial}{\partial z}\right) f(x, y, z), \tag{1.73}$$

or

$$f(\mathbf{r}+\boldsymbol{\epsilon}) = \exp(\boldsymbol{\epsilon}\cdot\nabla) f(\mathbf{r}) = \left[1 + \boldsymbol{\epsilon}\cdot\nabla + \frac{1}{2!}(\boldsymbol{\epsilon}\cdot\nabla)^2 + \cdots\right] f(\mathbf{r}). \tag{1.74}$$

1.4 Integrals

1.4.1 Jacobian Determinant

The determinant of the Jacobian matrix $\mathbf{J}$ relates volume elements when changing variables in an integral. For example, suppose $\mathbf{x}$ and $\mathbf{y}$ are N-dimensional space vectors in two different coordinate systems, e.g., Cartesian and spherical. The volume elements $d^N x$ and $d^N y$ are related by

$$d^N x = |\mathbf{J}(\mathbf{x}, \mathbf{y})|\, d^N y = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} & \cdots & \dfrac{\partial x_1}{\partial y_N} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \dfrac{\partial x_N}{\partial y_1} & \dfrac{\partial x_N}{\partial y_2} & \cdots & \dfrac{\partial x_N}{\partial y_N} \end{vmatrix} d^N y. \tag{1.75}$$

1.4.2 The Divergence Theorem

Let $\mathbf{F}(\mathbf{r})$ be a vector function defined in a volume V enclosed by a surface S with an outward normal $\hat{\mathbf{n}}$. If $d\mathbf{S} = dS\hat{\mathbf{n}}$, the divergence theorem is

$$\int_V d^3r\, \nabla \cdot \mathbf{F} = \int_S d\mathbf{S} \cdot \mathbf{F}. \tag{1.76}$$

Special choices for the vector function $\mathbf{F}(\mathbf{r})$ produce various integral identities based on (1.76). For example, if $\mathbf{c}$ is an arbitrary constant vector, the reader can confirm that the choices $\mathbf{F}(\mathbf{r}) = \mathbf{c}\psi(\mathbf{r})$ and $\mathbf{F}(\mathbf{r}) = \mathbf{A}(\mathbf{r}) \times \mathbf{c}$ substituted into (1.76) respectively yield

$$\int_V d^3r\, \nabla\psi = \int_S d\mathbf{S}\, \psi \tag{1.77}$$

$$\int_V d^3r\, \nabla \times \mathbf{A} = \int_S d\mathbf{S} \times \mathbf{A}. \tag{1.78}$$

1.4.3 Green's Identities

The choice $\mathbf{F}(\mathbf{r}) = \phi(\mathbf{r})\nabla\psi(\mathbf{r})$ in (1.76) leads to *Green's first identity*,

$$\int_V d^3r\, [\phi\nabla^2\psi + \nabla\phi \cdot \nabla\psi] = \int_S d\mathbf{S} \cdot \phi\nabla\psi. \tag{1.79}$$

Writing (1.79) with the roles of ϕ and ψ exchanged and subtracting that equation from (1.79) itself gives *Green's second identity*,

$$\int_V d^3r\, [\phi\nabla^2\psi - \psi\nabla^2\phi] = \int_S d\mathbf{S} \cdot [\phi\nabla\psi - \psi\nabla\phi]. \tag{1.80}$$

The choice $\mathbf{F} = \mathbf{P} \times \nabla \times \mathbf{Q}$ in (1.76) and the identity $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B}$ produces a vector analog of Green's first identity:

$$\int_V d^3r\, [\nabla \times \mathbf{P} \cdot \nabla \times \mathbf{Q} - \mathbf{P} \cdot \nabla \times \nabla \times \mathbf{Q}] = \int_S d\mathbf{S} \cdot (\mathbf{P} \times \nabla \times \mathbf{Q}). \tag{1.81}$$

Writing (1.81) with **P** and **Q** interchanged and subtracting that equation from (1.81) gives a vector analog of Green's second identity:

$$\int_V d^3r\,[\mathbf{Q}\cdot\nabla\times\nabla\times\mathbf{P}-\mathbf{P}\cdot\nabla\times\nabla\times\mathbf{Q}]=\int_S d\mathbf{S}\cdot[\mathbf{P}\times\nabla\times\mathbf{Q}-\mathbf{Q}\times\nabla\times\mathbf{P}]. \tag{1.82}$$

1.4.4 Stokes' Theorem

Stokes' theorem applies to a vector function $\mathbf{F}(\mathbf{r})$ defined on an open surface S bounded by a closed curve C. If $d\boldsymbol{\ell}$ is a line element of C,

$$\int_S d\mathbf{S}\cdot\nabla\times\mathbf{F}=\oint_C d\boldsymbol{\ell}\cdot\mathbf{F}. \tag{1.83}$$

The curve C in (1.83) is traversed in the direction given by the right-hand rule when the thumb points in the direction of $d\mathbf{S}$. As with the divergence theorem, variations of (1.83) follow from the choices $\mathbf{F}=\mathbf{c}\psi$ and $\mathbf{F}=\mathbf{A}\times\mathbf{c}$:

$$\int_S d\mathbf{S}\times\nabla\psi=\oint_C d\boldsymbol{\ell}\,\psi \tag{1.84}$$

$$\oint_C d\boldsymbol{\ell}\times\mathbf{A}=\int_S dS_k\nabla A_k-\int_S d\mathbf{S}(\nabla\cdot\mathbf{A}). \tag{1.85}$$

1.4.5 The Time Derivative of a Flux Integral

Leibniz' Rule for the time derivative of a one-dimensional integral is

$$\frac{d}{dt}\int_{x_1(t)}^{x_2(t)} dx\,b(x,t)=b(x_2,t)\frac{dx_2}{dt}-b(x_1,t)\frac{dx_1}{dt}+\int_{x_1(t)}^{x_2(t)} dx\,\frac{\partial b}{\partial t}. \tag{1.86}$$

This formula generalizes to integrals over circuits, surfaces, and volumes which move through space. Our treatment of Faraday's law makes use of the time derivative of a surface integral where the surface $S(t)$ moves because its individual area elements move with velocity $\boldsymbol{\upsilon}(\mathbf{r},t)$. In that case,

$$\frac{d}{dt}\int_{S(t)} d\mathbf{S}\cdot\mathbf{B}=\int_{S(t)} d\mathbf{S}\cdot\left[\boldsymbol{\upsilon}(\nabla\cdot\mathbf{B})-\nabla\times(\boldsymbol{\upsilon}\times\mathbf{B})+\frac{\partial\mathbf{B}}{\partial t}\right]. \tag{1.87}$$

Proof: We calculate the change in flux from

$$\delta\left[\int\mathbf{B}\cdot d\mathbf{S}\right]=\int\delta\mathbf{B}\cdot d\mathbf{S}+\int\mathbf{B}\cdot\delta(\hat{\mathbf{n}}dS). \tag{1.88}$$

The first term on the right comes from time variations of **B**. The second term comes from time variations of the surface. Multiplication of every term in (1.88) by $1/\delta t$ gives

$$\frac{d}{dt}\int\mathbf{B}\cdot d\mathbf{S}=\int\frac{\partial\mathbf{B}}{\partial t}\cdot d\mathbf{S}+\frac{1}{\delta t}\int\mathbf{B}\cdot\delta(\hat{\mathbf{n}}dS). \tag{1.89}$$

We can focus on the second term on the right-hand side of (1.89) because the first term appears already as the last term in (1.87). Figure 1.3 shows an open surface $S(t)$ with local normal $\hat{\mathbf{n}}(t)$ which moves and/or distorts to the surface $S(t+\delta t)$ with local normal $\hat{\mathbf{n}}(t+\delta t)$ in time δt.