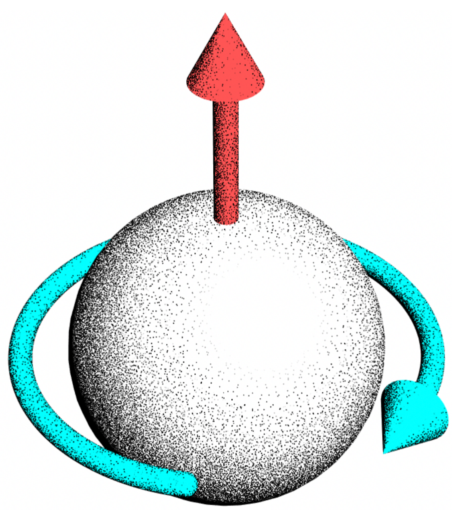
$g = 2$
A Gentle Introduction to Relativistic Quantum Mechanics

John F. Lindner
Physics Department
The College of Wooster

2020 August 15



Contents

List of Tables

List of Figures

Chapter 1

Foreword

For its charge and spin, the magnetic field of an electron is twice as large as classical physics predicts. Paul Dirac resolved this mystery by combining the twin pillars of modern physics, the theories of relativity and quantum mechanics, into a single beautiful equation.

Dirac's achievement is like Shakespeare's *Hamlet* or Beethoven's 9th Symphony. Indeed, when Dirac moved to Florida State University near the end of his career, the chair of physics defended hiring such a senior physicist by arguing, "The Physics Department hiring Dirac is like the English Department hiring Shakespeare".

But who gets to appreciate Dirac's great achievement? Not many people, as it requires a year of calculus and a year of classical physics just to get started. But you have those prerequisites and are ready for a challenging journey into the heart of modern physics culminating in the Dirac equation, which is commemorated by the Fig. 1.1 marker.

Figure 1.1: Paul Dirac's commemorative marker at Westminster Abbey includes his relativistic wave equation. (Creative Commons credit: Stanislav Kozlovskiy, 2014.)

Our journey involves the following milestones. Chapter 2 describes the classical electron using a **multipole expansion** of its electric and magnetic fields, relates charge, spin, and angular momentum, and formally states the electron $g = 2$ puzzle. Chapter 3 reduces the classical **Hamilton-Jacobi equation** for the **action**, which associates rays and wavefronts with classical particle motion, to the non-relativistic quantum-mechanical **Schrödinger equation**, when the classical action is small compared to the **action quantum** $\hbar$. Chapter 4 applies general symmetry principles to the observations of uniformly moving observers to derive the **Lorentz-Einstein** transformations and the **invariant speed** c without reference to light. Chapter 5 introduces mechanical and **field momentum** in the context of the "hidden" relativistic effect of a line charge on a parallel solenoid. Chapter 6 boldy introduces an **abstract algebra** of non-commuting numbers to devise a relativistic quantum-mechanical **Dirac equation**, generalizes it to a magnetic field, and solves the $g = 2$ electron puzzle. Chapter 7 is a teaser for the **quantum electrodynamics** sequel.

Chapter 2

Classical Electron

If an electron is a *spinning* ball of charge, it should create electric *and* magnetic fields, with the latter proportional to its angular momentum. Why is the electron's magnetic field twice as large as expected?

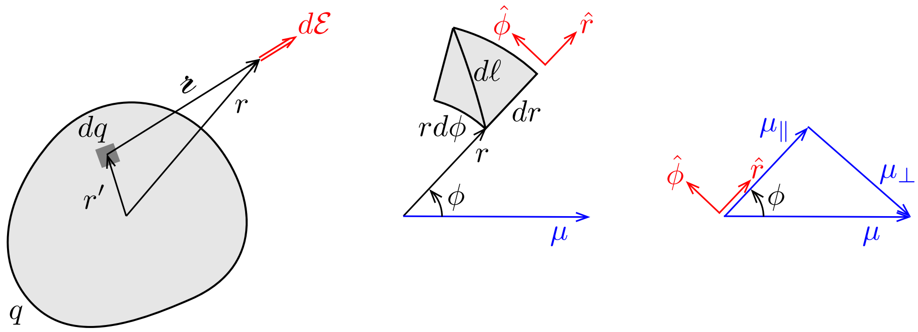


Figure 2.1: Multipole geometry (left), polar coordinate line and area element (center), dipole moment μ and projections (right).

2.1 Monopole and Dipole

The electric and magnetic fields of any charge distribution can be expanded in a series **multipole** terms. Decompose an arbitrary charge distribution into infinitesimal charges dq at positions $\vec{r}\,' = r'\hat{r}'$, as in Fig. 2.1. At a field point $\vec{r} = r\hat{r}$, the relative displacement $\vec{\boldsymbol{\imath}} = \vec{r} - \vec{r}\,'$ (pronounced "script r vector"), with square

$$\boldsymbol{\imath}^2 = r^2 - 2\vec{r}\cdot\vec{r}\,' + r'^2 = r^2\left(1 - 2\left(\frac{r'}{r}\right)\hat{r}\cdot\hat{r}' + \left(\frac{r'}{r}\right)^2\right). \tag{2.1}$$

By the **generalized binomial theorem**

$$(1+\epsilon)^{\alpha} \sim 1 + \alpha\epsilon \tag{2.2}$$

for $\epsilon \ll 1$, and so the reciprocal

$$\frac{1}{\mathscr{r}} = \frac{1}{r}\left(1 - 2\left(\frac{r'}{r}\right)\hat{r}\cdot\hat{r}' + \left(\frac{r'}{r}\right)^2\right)^{-1/2} \sim \frac{1}{r}\left(1 + \frac{\hat{r}\cdot\vec{r}'}{r}\right) \tag{2.3}$$

for $r' \ll r$. Thus, the **electric potential**

$$\begin{aligned}
\epsilon_0 V &= \int \frac{dq}{4\pi\mathscr{r}} \\
&\sim \int \frac{dq}{4\pi r} + \int \frac{dq\,\hat{r}\cdot\vec{r}'}{4\pi r^2} \\
&= \frac{1}{4\pi r}\int dq + \frac{\hat{r}}{4\pi r^2}\cdot\int dq\,\vec{r}' \\
&= \frac{q}{4\pi r} + \frac{\hat{r}\cdot\vec{\mu}}{4\pi r^2},
\end{aligned} \tag{2.4}$$

where

$$q = \int dq \tag{2.5}$$

is the **monopole moment** or total charge of the distribution, and

$$\vec{\mu} = \int dq\,\vec{r}' = \int \vec{r}'\,dq \tag{2.6}$$

is the **dipole moment** (with **quadrupole**, **octupole**, and higher order moments neglected). The monopole term decays slowly like $1/r$, but the dipole term decays quickly like $1/r^2$. If the dipole moment μ is at an angle ϕ from the field point, then the Eq. 2.4 potential

$$\epsilon_0 \mathcal{V} = \frac{q}{4\pi r} + \frac{\mu}{4\pi r^2}\cos\phi. \tag{2.7}$$

The corresponding **electric field** is the negative **gradient** of the electric potential (so positive charges move "downhill"). In polar coordinates $\{r, \phi\}$, where nearby points are separated by

$$d\vec{\ell} = dr\,\hat{r} + r\,d\phi\,\hat{\phi}, \tag{2.8}$$

as in Fig. 2.1, the total differential

$$dr\frac{\partial f}{\partial r} + d\phi\frac{\partial f}{\partial \phi} = df = \vec{\nabla} f \cdot d\vec{\ell} = \vec{\nabla} f \cdot \left(dr\,\hat{r} + r\,d\phi\,\hat{\phi}\right), \tag{2.9}$$

and so the gradient

$$\vec{\nabla} f = \frac{\partial f}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial f}{\partial \phi}\hat{\phi}, \tag{2.10}$$

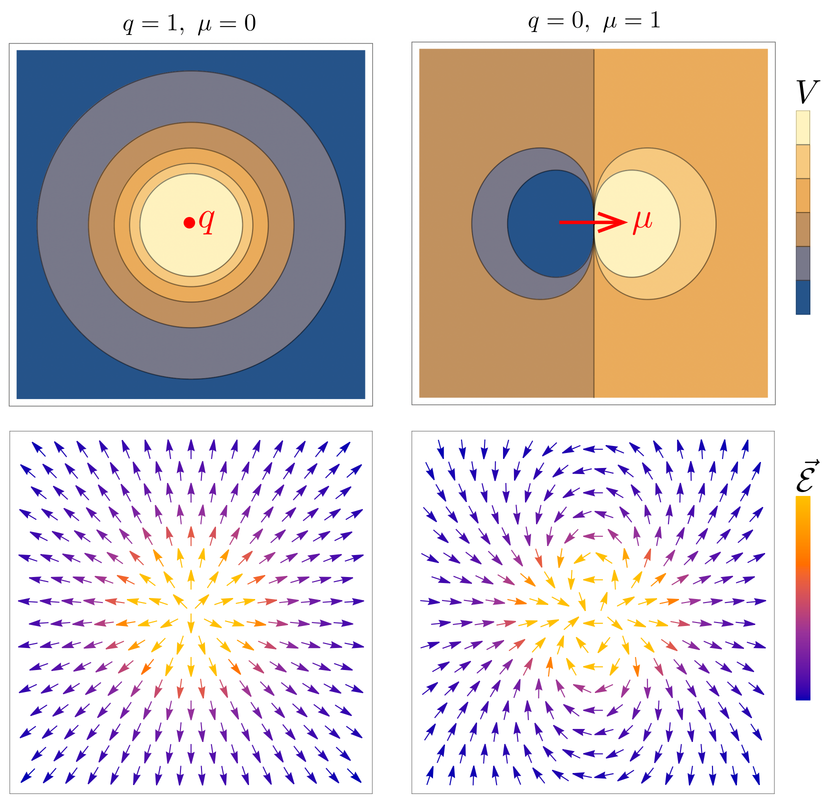


Figure 2.2: Potential contour plots (top) and field vector plots (bottom) for a spherically symmetric monopole (left) and a cylindrically symmetric dipole (right).

where the $1/r$ scale factor also makes sense dimensionally. For the Eq. 2.7 potential,

$$\epsilon_0 \vec{\mathcal{E}} = -\epsilon_0 \vec{\nabla} \mathcal{V} = \frac{q}{4\pi r^2}\hat{r} + \frac{\mu}{2\pi r^3}\cos\phi\,\hat{r} + \frac{1}{r}\frac{\mu}{4\pi r^2}\sin\phi\,\hat{\phi} \tag{2.11}$$

or

$$\begin{aligned} \epsilon_0 \vec{\mathcal{E}}[r,\phi] &= \frac{q}{4\pi r^2}\hat{r} + \frac{\mu}{4\pi r^3}\left(2\cos\phi\,\hat{r} + \sin\phi\,\hat{\phi}\right) \\ &= \frac{q}{4\pi r^2}\hat{r} + \frac{2\vec{\mu}_\parallel - \vec{\mu}_\perp}{4\pi r^3} \\ &= \frac{q}{4\pi r^2}\hat{r} + \frac{3\left(\vec{\mu}\cdot\hat{r}\right)\hat{r} - \vec{\mu}}{4\pi r^3}, \end{aligned} \tag{2.12}$$

where $\vec{\mu} - \vec{\mu}_\perp = \vec{\mu}_\parallel = (\vec{\mu}\cdot\hat{r})\hat{r}$. On the dipole axis $\phi = 0$, and the electric field

$$\epsilon_0 \mathcal{E}_z[z,0] = \frac{q}{4\pi z^2} + \frac{\mu}{2\pi z^3}. \tag{2.13}$$

For emphasis, sometimes write the electric field

$$\epsilon_0 \mathcal{E}_z[z,0] = \frac{q_\mathcal{E}}{4\pi z^2} + \frac{\mu_\mathcal{E}}{2\pi z^3}. \tag{2.14}$$

Similarly, the magnetic field

$$\mu_0^{-1} \mathcal{B}_z[z,0] = \frac{q_\mathcal{B}}{4\pi z^2} + \frac{\mu_\mathcal{B}}{2\pi z^3} = \frac{\mu_\mathcal{B}}{2\pi z^3}, \tag{2.15}$$

except $q_\mathcal{B} = 0$, as magnetic monopoles have never been observed. Figure 2.2 visualizes planar sections of monopole and dipole potentials and fields.

2.2 Dipole and Angular Momentum

A rotating charged ring generates a distant dipole magnetic field proportional to its angular moment. Consider the magnetic field $\vec{\mathcal{B}}$ at a distance z on the axis of a rotating ring of mass m and charge q generating a current loop $v dq = I d\ell$ of radius R, as in Fig. 2.3.

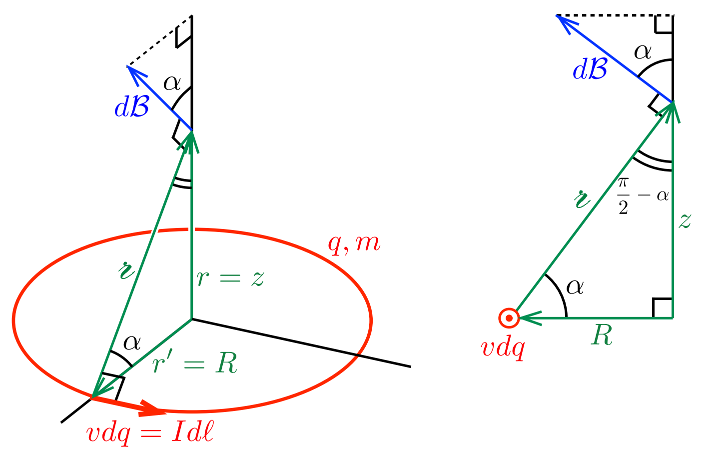


Figure 2.3: A rotating charged ring forms a current loop (left) and the geometry of the Biot-Savart triangle rotated into the plane of the page (right).

For slow speed $v \ll c$ Biot-Savart's law implies

$$d\vec{\mathcal{B}} \sim \mu_0 \epsilon_0 \vec{v} \times d\vec{\mathcal{E}} \sim \mu_0 \vec{v} \times \frac{dq}{4\pi \boldsymbol{r}^2} \hat{\boldsymbol{r}} \tag{2.16}$$

and

$$\mathcal{B}_z = \int d\mathcal{B}_z = \int d\mathcal{B} \cos\alpha = \mu_0 \int \frac{v dq \sin[\pi/2]}{4\pi \boldsymbol{r}^2} \cos\alpha. \tag{2.17}$$

Since the separation $\mathscr{r}$ and the angle α is constant around the ring,

$$\begin{aligned}
\mu_0^{-1}\mathcal{B}_z &= \frac{v}{4\pi\mathscr{r}^2}\left(\int dq\right)\cos\alpha = \frac{v}{4\pi\mathscr{r}^2}q\frac{R}{\mathscr{r}} = \frac{vqR}{4\pi\mathscr{r}^3} \\
&= \frac{vq}{4\pi}\frac{R}{(z^2+R^2)^{3/2}},
\end{aligned} \tag{2.18}$$

Far from the ring $R \ll z$ and

$$\begin{aligned}
\mu_0^{-1}\mathcal{B}_z &= \frac{vqR}{4\pi z^3}\left(1+\left(\frac{R}{z}\right)^2\right)^{-3/2} \\
&\sim \frac{q}{2m}\frac{mvR}{2\pi z^3}\left(1-\frac{3}{2}\left(\frac{R}{z}\right)^2\right) \\
&\sim \frac{q}{2m}\frac{L}{2\pi z^3} = \frac{\mu_{\mathcal{B}}}{2\pi z^3}
\end{aligned} \tag{2.19}$$

where the Eq. 2.15 magnetic dipole moment

$$\vec{\mu}_{\mathcal{B}} = \frac{q}{2m}\vec{L} \tag{2.20}$$

is proportional to the ring's $L = mvR$ angular momentum, and this relationship is generally true.

Equivalently, shrinking the ring radius R and growing the charge q, construe a point dipole as the double limit

$$\mu_{\mathcal{B}} = \lim_{\substack{R\to 0\\ q\to\infty}} \frac{qvR}{2}, \tag{2.21}$$

or interpret the rotating charged loop as a circular current

$$I = \frac{dq}{dt} = \frac{\lambda d\ell}{dt} = \frac{\lambda R d\phi}{dt} = \lambda R\omega = \frac{q}{2\pi R}R\frac{v}{R} = \frac{qv}{2\pi R}, \tag{2.22}$$

so the Eq. 2.21 dipole moment becomes

$$\mu_{\mathcal{B}} = \lim_{\substack{R\to 0\\ q\to\infty}} \frac{(I2\pi R)R}{2} = \lim I\pi R^2 = \lim_{\substack{A\to 0\\ I\to\infty}} IA. \tag{2.23}$$

When necessary, distinguish $q = q_{\mathcal{E}}$ from $q_{\mathcal{B}}$.

2.3 Dipole Force and Torque

External fields force and torque dipoles. Imagine two point charges q separated by a length ℓ, as in Fig. 2.4. Far from the dipole along the axis $\ell \ll z$ and

$$\begin{aligned}
\mathcal{E}_z &= \frac{q}{4\pi(z-\ell)^2} - \frac{q}{4\pi z^2} = \frac{q}{4\pi z^2}\left(\left(1-\frac{\ell}{z}\right)^{-2} - 1\right) \\
&\sim \frac{q}{4\pi z^2}\left(\left(1+2\frac{\ell}{z}\right) - 1\right) = \frac{q\ell}{2\pi z^3} = \frac{\mu_{\mathcal{E}}}{2\pi z^3},
\end{aligned} \tag{2.24}$$

where by comparing with Eq. 2.14, the electric dipole moment

$$\mu_{\mathcal{E}} = \lim_{\substack{\ell \to 0 \\ q \to \infty}} q\ell. \tag{2.25}$$

Call separated monopoles with $\mu \sim q\ell$ **Gilbert** dipoles and current loops with $\mu \sim IA$ **Ampère** dipoles.

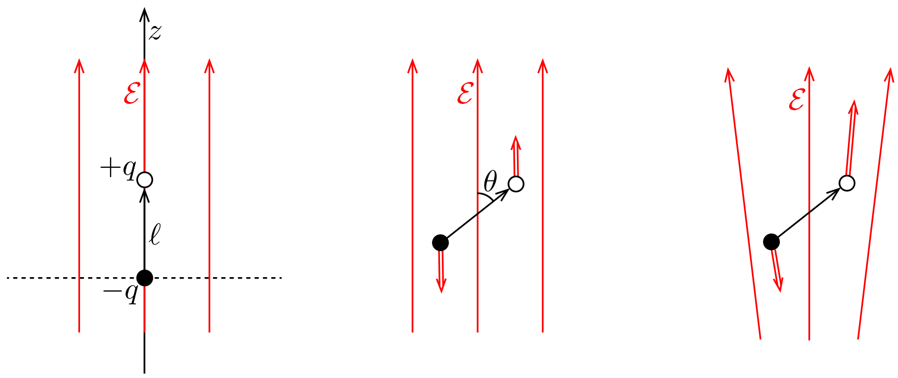


Figure 2.4: Dipole as separation of charge (left), torque due to uniform field (center), force due to nonuniform field (right).

A uniform electric field $\vec{\mathcal{E}}$ causes a torque

$$\tau = \frac{\ell}{2} q\mathcal{E} + \frac{\ell}{2} q\mathcal{E} = q\ell\,\mathcal{E} \tag{2.26}$$

or

$$\vec{\tau} = \vec{\mu} \times \vec{\mathcal{E}}. \tag{2.27}$$

The corresponding potential energy

$$U = \int d\phi\,\tau = \int d\phi\,\mu\mathcal{E}\sin\phi = -\mu\mathcal{E}\cos\phi = -\vec{\mu}\cdot\vec{\mathcal{E}} \tag{2.28}$$

A nonuniform electric field causes a force

$$\begin{aligned} \vec{F} &= q\vec{\mathcal{E}}[\vec{r} + \ell\hat{\ell}] - q\vec{\mathcal{E}}[\vec{r}] \\ &= q\ell\,\frac{\vec{\mathcal{E}}[\vec{r} + \ell\hat{\ell}] - \vec{\mathcal{E}}[\vec{r}]}{\ell} \\ &\sim \mu\,\hat{\ell}\cdot\vec{\nabla}\vec{\mathcal{E}} = \left(\vec{\mu}\cdot\vec{\nabla}\right)\vec{\mathcal{E}}. \end{aligned} \tag{2.29}$$

Analogous formulas

$$\vec{\tau} = \vec{\mu}_{\mathcal{B}} \times \vec{\mathcal{B}} \tag{2.30}$$

and

$$\vec{F} = \left(\vec{\mu}_{\mathcal{B}}\cdot\vec{\nabla}\right)\vec{\mathcal{B}} \tag{2.31}$$

hold for magnetic dipoles in magnetic fields. When necessary, distinguish $\mu = \mu_{\mathcal{E}}$ from $\mu_{\mathcal{B}}$.

2.4 Electron Puzzle

Experimentally, the electron has electric monopole moment

$$q_e = -160 \text{ zC}, \tag{2.32}$$

magnetic dipole moment

$$\mu_e = -9.28 \ \mu\text{A} \cdot \text{nm}^2, \tag{2.33}$$

mass

$$m_e = 0.000911 \text{ yg}, \tag{2.34}$$

and spin angular momentum

$$S_e = \frac{\hbar}{2} = 52.7 \text{ zJ} \cdot \text{fs}. \tag{2.35}$$

Thus, like Eq. 2.20, the electron's magnetic moment

$$\mu_e = g \frac{q_e}{2m_e} S \tag{2.36}$$

is proportional to its angular momentum but with dimensionless correction factor

$$g = \frac{2m_e \mu_e}{q_e S} = \frac{2(0.000911 \text{ yg})(9.28 \ \mu\text{A} \cdot \text{nm}^2)}{(160 \text{ zC})(52.7 \text{ zJ} \cdot \text{fs})} = 2.00. \tag{2.37}$$

Why is the electron g-factor 2? Why is it twice the expected classical value? $g \neq 1$ implies that a ball of spinning charge is at best a crude model of the electron. Combine **quantum mechanics** and **relativity**, the twin pillars of modern physics, into **relativistic quantum mechanics** to naturally derive this surprising result.

Problems

1. Use the generalized binomial theorem to simplify the following assuming $x \ll a$.

   (a) $1/(1 + x/a)$

   (b) $1/\sqrt{a^2 + x^2}$

   (c) $(a^4 + x^4)^{-3/2}$

2. Derive the following gradient identities.

   (a) $\vec{\nabla} r = \hat{r}$

   (b) $\vec{\nabla} r^2 = 2\hat{r}$

   (c) $\vec{\nabla}\dfrac{1}{r} = -\dfrac{\hat{r}}{r^2}$

3. Use *Mathematica* to visualize the potentials and fields of an monopoles and dipole in 3D. (Create a 3D version of Fig. 2.2.)

Chapter 3

Schrödinger

William Rowan Hamilton anticipated wave-particle duality and quantum mechanics by nearly a century in formulating a version of classical mechanics in analogy with the dual descriptions of **geometrical optics** as light rays and wave fronts. The resulting **Hamilton-Jacobi equation** connects classical and quantum mechanics.

3.1 Classical Wave-Particle Duality

Consider a particle moving in the x-direction. Its velocity can vary directly with time and indirectly with position like

$$\frac{dx}{dt} = v_x\big[t, x[t]\big]. \tag{3.1}$$

Its acceleration a_x is proportional to the total applied force F_x and inversely proportional to its mass m. Assume the applied force is minus the gradient of a potential energy U, so the mass tends to "roll downhill". Newton's equations and the derivative **chain rule** imply

$$-\frac{\partial U}{\partial x} = F_x = ma_x = m\frac{dv_x}{dt} = m\left(\frac{\partial v_x}{\partial t} + \frac{\partial v_x}{\partial x}\frac{\partial x}{\partial t}\right). \tag{3.2}$$

For simple systems, the momentum $p_x = mv_x$ is mass times velocity, so

$$-\frac{\partial U}{\partial x} = \frac{\partial p_x}{\partial t} + \frac{\partial p_x}{\partial x}\frac{p_x}{m} = \frac{\partial p_x}{\partial t} + \frac{1}{2m}\frac{\partial}{\partial x}p_x^2. \tag{3.3}$$

is a partial differential equation that controls the flow of momentum p_x under the potential U. If streams of constant momentum don't intersect (in shock fronts, for example), the momentum

$$p_x = \frac{\partial S}{\partial x} \tag{3.4}$$

is the gradient of a scalar function or field, the **action** $S[t, x]$. In higher dimensions and rectangular coordinates,

$$\vec{p} = +\vec{\nabla} S = \hat{x}\frac{\partial S}{\partial x} + \hat{y}\frac{\partial S}{\partial y} + \hat{z}\frac{\partial S}{\partial z}. \tag{3.5}$$

With the action Eq. 3.4, the momentum flow Eq. 3.3 rearranges to

$$\frac{\partial}{\partial t}\frac{\partial S}{\partial x} + \frac{1}{2m}\frac{\partial}{\partial x}\left(\frac{\partial S}{\partial x}\right)^2 + \frac{\partial U}{\partial x} = 0, \tag{3.6}$$

and assuming the action's second order derivatives exist and are continuous (so its partial derivatives commute),

$$\frac{\partial}{\partial x}\left(\frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + U\right) = 0. \tag{3.7}$$

Integrate over x to get a constant with respect to x,

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + U = f, \tag{3.8}$$

which is possibly a function of time $f[t]$. To eliminate it, shift the action by the time integral $S \to S + \int dt\, f$ so the derivative shifts by the function $\partial S/\partial t \to \partial S/\partial t + f$. The **Hamilton-Jacob equation** becomes

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + U = 0. \tag{3.9}$$

or

$$-\frac{\partial S}{\partial t} = \frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + U = \frac{p^2}{2m} + U = K + U = E. \tag{3.10}$$

In higher dimensions,

$$\boxed{-\frac{\partial S}{\partial t} = \frac{1}{2m}\vec{\nabla} S \cdot \vec{\nabla} S + U}. \tag{3.11}$$

3.2 Hamiltonian-Jacobi Examples

The Hamilton-Jacobi equation associates surfaces of constant action S to the momentum $\vec{p} = \vec{\nabla} S$ like geometric optics associates wave-fronts to rays, a **duality** between wave-fronts and trajectories. For simple systems, the velocity is proportional to the momentum and the path is tangent to the momenta and perpendicular to the wave-fronts.

For the motion of a $U = 0$ free particle, the action

$$\begin{aligned} S[\vec{r}, t] &= \vec{p} \cdot \vec{r} - \frac{p^2}{2m} t \\ &= p_x x + p_y y + p_z z - \frac{p_x^2 + p_y^2 + p_z^2}{2m} t. \end{aligned} \tag{3.12}$$

To check, the derivatives

$$\vec{\nabla} S = \frac{\partial S}{\partial \vec{r}} = \vec{p} - \vec{0} \tag{3.13}$$

and

$$\frac{\partial S}{\partial t} = 0 - \frac{p^2}{2m} \tag{3.14}$$

imply

$$-\frac{\partial S}{\partial t} = \frac{1}{2m}\left(\vec{\nabla} S\right)^2 = \frac{1}{2m}\vec{\nabla} S \cdot \vec{\nabla} S + 0 \tag{3.15}$$

as expected by the Hamilton-Jacobi Eq. 3.11. Constant momentum rays

$$\vec{r} = \vec{r}_0 + \frac{\vec{p}}{m} t \tag{3.16}$$

pierce plane wave-fronts like

$$S[\vec{r}, 0] = \vec{p} \cdot \vec{r}, \tag{3.17}$$

as in Fig. 3.1.

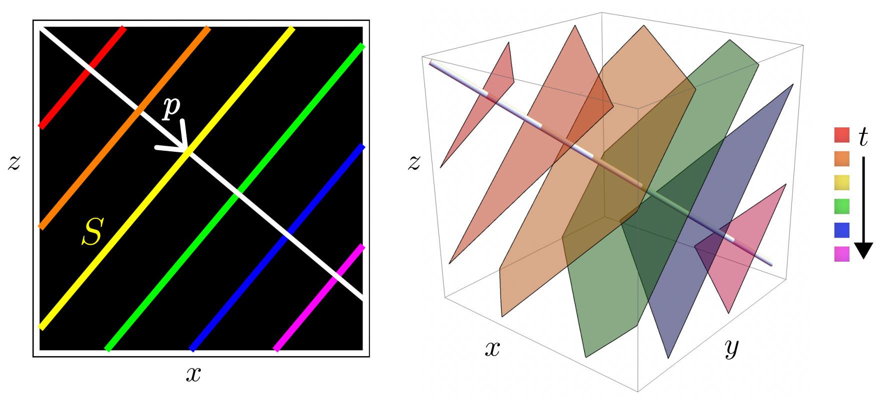


Figure 3.1: A free particle moves in a straight line at constant speed. The Hamilton-Jacobi equation associates planes of constant action S to the particle's constant momentum $\vec{p} = \vec{\nabla} S$ like geometric optics associates wave-fronts to rays in a vacuum.

Next consider a ball thrown under gravity $U = mgz$. If the initial momentum $\vec{p}_0 = \{p_{0x}, p_{0y}, p_{0z}\}$, then the action

$$S_{\pm}[x, y, z, t] = p_{0x}x + p_{0y}y \pm \frac{1}{3m^2 g}\left(p_{0z}^2 - 2m^2 gz\right)^{3/2} - \frac{p_{0x}^2 + p_{0y}^2 + p_{0z}^2}{2m} t, \tag{3.18}$$

where the positive root is for upward motion and the negative root for downward motion. Gravity breaks the symmetry in the z-direction, and variable momentum rays pierce curved wave-fronts, as in Fig. 3.2. The curved trajectory is like a light ray bent upwards by desert heat in creating the inferior mirage of a watery oasis or like light curving through variable index-of-refraction bifocal eye glasses.

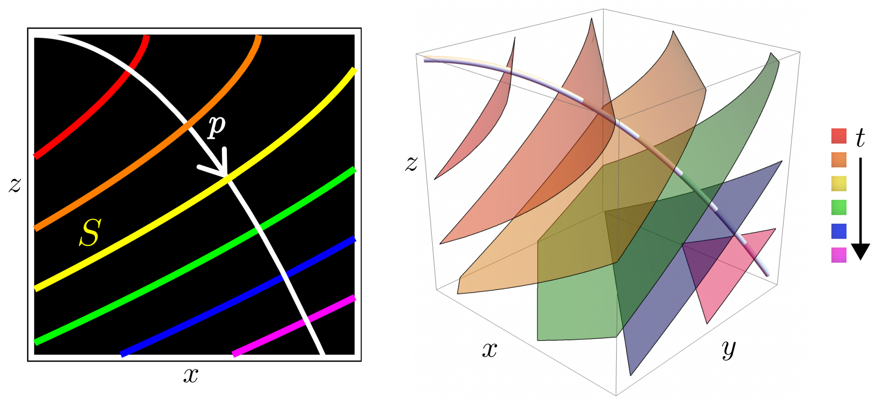


Figure 3.2: A ball thrown horizontally falls under gravity. The Hamilton-Jacobi equation associates curved surfaces of constant action S to the ball's curved trajectory $\vec{p} = \vec{\nabla} S$ like geometric optics associates wave-fronts to rays in a medium with a variable refraction index (such as air heated by asphalt producing a mirage).

3.3 Schrödinger Equation

To transition from geometric optics and mechanics to wave optics and quantum mechanics, imagine that surfaces of constant action S are indeed wavefronts of constant phase for the plane wave

$$\Psi = A e^{iS/\hbar} = A \exp\left[i\frac{S}{\hbar}\right] = A\cos\left[\frac{S}{\hbar}\right] + iA\sin\left[\frac{S}{\hbar}\right], \tag{3.19}$$

where the constant $\hbar$ (pronounced "h-bar") has the dimensions of S to enforce the dimensionlessness of the function arguments. Small $\hbar$ means small changes in S make large changes in the wave's phase, corresponding to a high frequency, geometric optics limit. Seek a partial differential equation for the wave function Ψ in this limit using the Hamilton-Jacobi equation.

First invert to solve for the action

$$S = -i\hbar \log\left[\frac{\Psi}{A}\right] = -i\hbar\log\Psi - i\hbar\log A. \tag{3.20}$$

Compute the derivatives

$$\frac{\partial S}{\partial t} = -i\hbar\frac{1}{\Psi}\frac{\partial \Psi}{\partial t}, \tag{3.21}$$

and

$$\frac{\partial S}{\partial x} = -i\hbar \frac{1}{\Psi}\frac{\partial \Psi}{\partial x}. \tag{3.22}$$

Substitute these action derivatives into the Hamilton-Jacobi Eq. 3.10 to get

$$+i\hbar \frac{1}{\Psi}\frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{1}{\Psi^2}\left(\frac{\partial \Psi}{\partial x}\right)^2 + U \tag{3.23}$$

or

$$i\hbar \Psi \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\left(\frac{\partial \Psi}{\partial x}\right)^2 + U\Psi^2. \tag{3.24}$$

Unfortunately, this is a *nonlinear* partial differential equation (doubling Ψ quadruples each term), and nonlinear equations are difficult to solve. Instead, solve Eq. 3.22 for

$$\frac{\partial \Psi}{\partial x} = \frac{i}{\hbar}\Psi \frac{\partial S}{\partial x}, \tag{3.25}$$

and compute the second derivative

$$\begin{aligned}\frac{\partial^2 \Psi}{\partial x^2} &= \frac{i}{\hbar}\frac{\partial \Psi}{\partial x}\frac{\partial S}{\partial x} + \frac{i}{\hbar}\Psi \frac{\partial^2 S}{\partial x^2} \\ &= -\frac{1}{\hbar^2}\Psi \left(\frac{\partial S}{\partial x}\right)^2 + \frac{i}{\hbar}\Psi \frac{\partial^2 S}{\partial x^2}.\end{aligned} \tag{3.26}$$

In the high frequency limit when $\hbar$ is small, $1/\hbar$ is large and $1/\hbar^2$ is *very* large, so neglect the right-side second term relative to the first term to obtain the **asymptotic** relation

$$\frac{\partial^2 \Psi}{\partial x^2} \underset{\hbar \downarrow 0}{\sim} -\frac{1}{\hbar^2}\Psi \left(\frac{\partial S}{\partial x}\right)^2, \tag{3.27}$$

and solve for

$$\left(\frac{\partial S}{\partial x}\right)^2 = -\hbar^2 \frac{1}{\Psi}\frac{\partial^2 \Psi}{\partial x^2}. \tag{3.28}$$

Finally, substitute the Eq. 3.21 and Eq. 3.28 action derivatives into the Hamilton-Jacobi Eq. 3.10 to get

$$+i\hbar \frac{1}{\Psi}\frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{1}{\Psi}\frac{\partial^2 \Psi}{\partial x^2} + U = 0 \tag{3.29}$$

or

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \Psi}{\partial x^2} + U\Psi. \tag{3.30}$$

This is a *linear* equation (doubling Ψ doubles each term), which was first discovered by **Erwin Schrödinger** [7] in 1926. In higher dimensions,

$$\boxed{i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 \Psi + U\Psi}. \tag{3.31}$$

Just as the square of the electric field in a light beam is proportional to the intensity of the light, $\mathcal{I} \propto \mathcal{E}^2$, the absolute square of the matter wave function is proportional to the probability of detecting a particle, $\mathcal{P} \propto |\Psi|^2$. Experimentally, the dimensional **reduced Planck constant**

$$\hbar = \frac{h}{2\pi} = 663 \text{ yJ/THz}, \tag{3.32}$$

and the **Planck constant**

$$h = 2\pi\hbar = 105 \text{ zJ} \cdot \text{fs}. \tag{3.33}$$

(Both are useful, just as frequency f and angular frequency $\omega = 2\pi f$ are both useful.)

3.4 Operator Formalism

Define the differential operators

$$\mathring{E} = +i\hbar\partial_t = +i\hbar\frac{\partial}{\partial t}, \tag{3.34a}$$

$$\mathring{p}_x = -i\hbar\partial_x = -i\hbar\frac{\partial}{\partial x}, \tag{3.34b}$$

read "e ring equals plus i h-bar del sub t" and "p sub x ring equals minus i h-bar del sub x", and the $1+1$ dimensional Schrödinger Eq. 3.30 simplifies to

$$\mathring{E}\Psi = \frac{\mathring{p}_x^2}{2m}\Psi + U\Psi. \tag{3.35}$$

Traveling wave functions

$$\Psi = Ae^{i(k_x x - \omega t)}, \tag{3.36}$$

are energy and momentum **eigenfunctions**

$$\mathring{E}\Psi = +i\hbar\partial_t\Psi = +i\hbar(-i\omega)\Psi = \hbar\omega\,\Psi = E\Psi, \tag{3.37a}$$

$$\mathring{p}_x\Psi = -i\hbar\partial_x\Psi = -i\hbar(+ik_x)\Psi = \hbar k_x\,\Psi = p_x\Psi, \tag{3.37b}$$

whose **eigenvalues**

$$E = \hbar\omega, \tag{3.38a}$$

$$p_x = \hbar k_x \tag{3.38b}$$

are the **Einstein-deBroglie** relations. Generalizing to $3+1$ dimensions, the Eq. 3.34 operator formalism provides the quick path

$$E = \frac{p^2}{2m} + U, \tag{3.39a}$$

$$\mathring{E}\Psi = \frac{\mathring{p}^2}{2m}\Psi + U\Psi, \tag{3.39b}$$

$$i\hbar\,\partial_t\Psi = -\frac{\hbar^2}{2m}\nabla^2\Psi + U\Psi \tag{3.39c}$$

from classical energy to quantum Schrödinger Eq. 3.31.

Problems

1. Check that the Eq. 3.18 thrown-ball action
    (a) satisfies the Hamilton-Jacobi Eq. 3.11
    (b) and has the correct Eq. 3.5 gradient.
2. Substitute the wave function Eq. 3.19 into the Schrödinger Eq. 3.31 and recover the Hamilton-Jacobi Eq. 3.11 in the limit $\hbar \downarrow 0$.

Chapter 4

Relativity

Symmetries imply an invariant speed [5].

4.1 Linear Transformation

Consider two **reference frames** or **observers** $\mathcal{O}$ and $\mathcal{O}'$, say Earth and a spaceship, in relative motion at velocity $\vec{v} = v\hat{x}$ along a common $\hat{x}$ axis, as in Fig. 4.1, where the $\hat{y}$ and $\hat{z}$ axes are suppressed for simplicity. Assume they observe the same **event**, say a supernova, which $\mathcal{O}$ locates at time t and space x and $\mathcal{O}'$ locates at

$$t' = T_v[t, x], \tag{4.1a}$$

$$x' = X_v[t, x], \tag{4.1b}$$

where the functions T and X are to be determined.

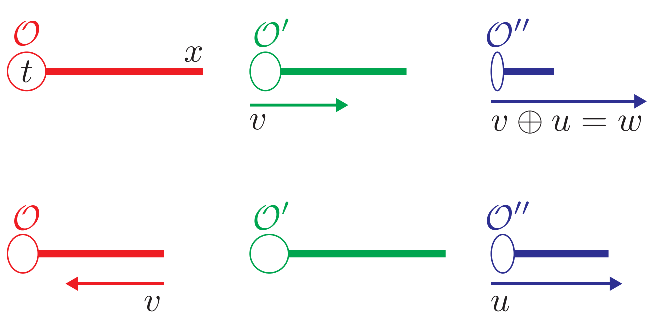


Figure 4.1: Observer $\mathcal{O}''$ moves at speed u relative to observer $\mathcal{O}'$ who moves at speed v relative to observer $\mathcal{O}$. All observers carry their own clocks t and rulers x.

Assume **space is homogeneous**, so if a ruler at rest in $\mathcal{O}$ extends from x_1 to x_2, and $\mathcal{O}'$ measures its length

$$\ell' = X_v[t, x_2] - X_v[t, x_1], \tag{4.2}$$

translating the ruler a distance h in $\mathcal{O}$ won't change its length in $\mathcal{O}'$,

$$\ell' = X_v[t, x_2 + h] - X_v[t, x_1 + h]. \tag{4.3}$$

Equate the right sides, shuffle, and divide by h,

$$\frac{X_v[t, x_2 + h] - X_v[t, x_2 + h]}{h} = \frac{X_v[t, x_1 + h] - X_v[t, x_1 + h]}{h}. \tag{4.4}$$

Assume **smoothness** and take the limit $h \to 0$ to find

$$\left.\frac{\partial X_v}{\partial x}\right|_{x_2} = \left.\frac{\partial X_v}{\partial x}\right|_{x_1} = \text{constant} = \frac{\partial X_v}{\partial x}, \tag{4.5}$$

as the first term can depend only on x_1 and the second only on x_2 only if they are the same constant. Similarly, assume **time is homogeneous** to show

$$\left.\frac{\partial X_v}{\partial t}\right|_{t_2} = \left.\frac{\partial X_v}{\partial t}\right|_{t_1} = \text{constant} = \frac{\partial X_v}{\partial t}. \tag{4.6}$$

Thus the coordinate transformation is **linear**, so write

$$t' = T_v[t, x] = A_v t + B_v x + \text{constant}, \tag{4.7a}$$
$$x' = X_v[t, x] = C_v t + D_v x + \text{constant}. \tag{4.7b}$$

By convention, choose $T_v[0, 0] = 0 = X_v[0, 0]$, so the observers' origins coincide and the constants vanish, so

$$t' = A_v t + B_v x, \tag{4.8a}$$
$$x' = C_v t + D_v x, \tag{4.8b}$$

or as the single matrix equation

$$\begin{bmatrix} t' \\ x' \end{bmatrix} = \begin{bmatrix} A_v & B_v \\ C_v & D_v \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix}. \tag{4.9}$$

4.2 Lorentz Transformation

Assume **space is isotropic**, so position and velocity invert simultaneously,

$$T_{-v}[t, -x] = +T_v[t, x], \tag{4.10a}$$
$$X_{-v}[t, -x] = -X_v[t, x], \tag{4.10b}$$

and

$$A_{-v}t - B_{-v}x = +A_v t + B_v x, \tag{4.11a}$$
$$C_{-v}t - D_{-v}x = -C_v t - V_v x. \tag{4.11b}$$

Compare term-by-term to discover the symmetry and anti-symmetry

$$A_{-v} = +A_v, \tag{4.12a}$$
$$B_{-v} = -B_v, \tag{4.12b}$$
$$C_{-v} = -C_v, \tag{4.12c}$$
$$D_{-v} = +D_v. \tag{4.12d}$$

Assume **motion is relative**, so the **inverse** transformation

$$t = T_{-v}[t', x'], \tag{4.13a}$$
$$x = X_{-v}[t', x'] \tag{4.13b}$$

inverts velocity and swaps primes and unprimes. Concatenate the transformations and their inverses to form the constraints

$$t = T_{-v}\big[T_v[t,x], X_v[t,x]\big], \tag{4.14a}$$
$$x = X_{-v}\big[T_v[t,x], X_v[t,x]\big]. \tag{4.14b}$$

or

$$t = A_{-v}(A_v t + B_v x) + B_{-v}(C_v t + D_v x), \tag{4.15a}$$
$$x = C_{-v}(A_v t + B_v x) + D_{-v}(C_v t + D_v x), \tag{4.15b}$$

or using the Eq. 4.12 symmetries and anti-symmetries,

$$t = +A_v(A_v t + B_v x) - B_v(C_v t + D_v x), \tag{4.16a}$$
$$x = -C_v(A_v t + B_v x) + D_v(C_v t + D_v x). \tag{4.16b}$$

Compare term-by-term to find

$$A_v^2 - B_v C_v = 1, \tag{4.17a}$$
$$B_v(A_v - D_v) = 0, \tag{4.17b}$$
$$C_v(A_v - D_v) = 0, \tag{4.17c}$$
$$D_v^2 - B_v C_v = 1. \tag{4.17d}$$

The middle equations imply either $A_v = 0 = D_v$, in which case the end equations imply $B_v = 1 = C_v$, which is trivial, or

$$A_v = D_v \tag{4.18}$$

and by Eq. 4.17d

$$B_v = \frac{D_v^2 - 1}{C_v}, \tag{4.19}$$

which is nontrival. Since the observers' origins coincide at $t = 0 = t'$, $x' = 0$ implies $x = vt$, and so Eq. 4.8b implies

$$C_v = -vD_v. \tag{4.20}$$

With these substitutions, the linear Eq. 4.9 transformation becomes

$$\begin{bmatrix} t' \\ x' \end{bmatrix} = \begin{bmatrix} D_v & \frac{D_v^2-1}{-vD_v} \\ -vD_v & D_v \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix}, \tag{4.21}$$

where $D_0 = 1$ and $D_{-v} = D_v$.

Consider a third observer $\mathcal{O}''$ in relative motion at velocity $\vec{u} = u\hat{x}$ relative to $\mathcal{O}'$ and at velocity $\vec{w} = w\hat{x}$ relative to $\mathcal{O}$, as in Fig. 4.1. **Concatenate** the Eq. 4.21 transformation to find

$$\begin{aligned} \begin{bmatrix} t'' \\ x'' \end{bmatrix} &= \begin{bmatrix} D_w & \frac{D_w^2-1}{-wD_w} \\ -wD_w & D_w \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix} \\ &= \begin{bmatrix} D_u & \frac{D_u^2-1}{-uD_u} \\ -uD_u & D_u \end{bmatrix} \begin{bmatrix} D_v & \frac{D_v^2-1}{-vD_v} \\ -vD_v & D_v \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix} \\ &= \begin{bmatrix} D_uD_v + (D_u^2-1)\frac{vD_v}{uD_u} & -(D_u^2-1)\frac{D_v}{uD_u} - (D_v^2-1)\frac{D_u}{vD_v} \\ -(u+v)D_uD_v & D_uD_v + (D_v^2-1)\frac{uD_u}{vD_v} \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix}. \end{aligned} \tag{4.22}$$

Since the two Eq. 4.22 primary diagonal ↘ elements must be equal, the last line implies

$$(D_u^2 - 1)\frac{vD_v}{uD_u} = (D_v^2 - 1)\frac{uD_u}{vD_v} \tag{4.23}$$

or

$$\frac{D_u^2 - 1}{u^2 D_u^2} = \frac{D_v^2 - 1}{v^2 D_v^2} = \text{constant} = \frac{1}{c^2}, \tag{4.24}$$

as the first term can depend only on v and the second only on u only if they are the same constant, here with dimensions of inverse speed squared. Thus,

$$D_v = \frac{1}{\sqrt{1 - v^2/c^2}} = \gamma \tag{4.25}$$

is the famous **relativistic stretch** or "gamma" factor, and

$$B_v = \frac{D_v^2 - 1}{-vD_v} = -\gamma v\left(\frac{1 - 1/\gamma^2}{v^2}\right) = -\gamma\frac{v}{c^2}. \tag{4.26}$$

The **Lorentz-Einstein transformation** [2] becomes

$$\begin{bmatrix} t' \\ x' \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma v/c^2 \\ -\gamma v & \gamma \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix} \tag{4.27}$$

or as two separate scalar equations

$$t' = \gamma\left(t - vx/c^2\right), \tag{4.28a}$$
$$x' = \gamma\left(x - vt\right). \tag{4.28b}$$

4.3 Velocity Addition

Also, the Eq. 4.22 top-left elements with Eq. 4.24 must satisfy

$$\begin{aligned} D_w &= D_u D_v + (D_u^2 - 1)\frac{vD_v}{uD_u} \\ &= D_u D_v + \left(\frac{u^2 D_u^2}{c^2}\right)\frac{vD_v}{uD_u} \\ &= D_u D_v \left(1 + \frac{uv}{c^2}\right). \end{aligned} \tag{4.29}$$

Square both sides and with Eq. 4.25 write

$$\frac{1}{1 - w^2/c^2} = \frac{(1 + uv/c^2)^2}{(1 - u^2/c^2)(1 - v^2/c^2)}. \tag{4.30}$$

Reciprocate and solve for

$$w^2 = c^2 - \frac{(1 - u^2/c^2)(1 - v^2/c^2)}{(1 + uv/c^2)^2} = \frac{(u+v)^2}{(1 + uv/c^2)^2} \tag{4.31}$$

and

$$w = \frac{u + v}{1 + uv/c^2} = u \oplus v, \tag{4.32}$$

which is the **velocity addition formula**.

The constant c has the dimensions of speed, but speed relative to what? If observer $\mathcal{O}''$ moves with x-velocity $u = c$ relative to observer $\mathcal{O}'$ and observer $\mathcal{O}'$ moves with x-velocity $v = c$ relative to observer $\mathcal{O}$, then observer $\mathcal{O}''$ moves with x-velocity

$$w = c \oplus c = \frac{c + c}{1 + c\,c/c^2} = c \tag{4.33}$$

relative to $\mathcal{O}$. The speed c is **invariant**, the same for all observers. Relativity and the very general assumptions of homogeneity, isotropy, smoothness, inversion, and concatenation imply its remarkable existence. Experimentally,

$$c = 299\,792\,458 \text{ m/s} = 0.3 \text{ m/ns} \approx 1 \text{ ft/ns} \approx 1 \text{ billion km/hr}. \tag{4.34}$$

Light and gravitational waves appear to travel at this unique invariant speed.

Problems

1. Verify that the velocity addition Eq. 4.32 is associative.

   (a) $(u \oplus v) \oplus w = u \oplus (v \oplus w) = u \oplus v \oplus w$

   (b) $u \oplus v \oplus (-u) = v$

Chapter 5

Potentials & Momenta

The electric and magnetic fields $\vec{\mathcal{E}}$ and $\vec{\mathcal{B}}$ are the derivatives of the electric and magnetic potentials $\mathcal{V}$ and $\vec{\mathcal{A}}$ and store energy and momentum U and $\vec{p}$.

5.1 Electric and Magnetic Potentials

A charge at rest generates an electric field. The **electric scalar potential**

$$\epsilon_0 \mathcal{V} = \int \frac{dq}{4\pi \mathscr{r}} = \iiint \frac{\rho\, dV}{4\pi \mathscr{r}}, \tag{5.1}$$

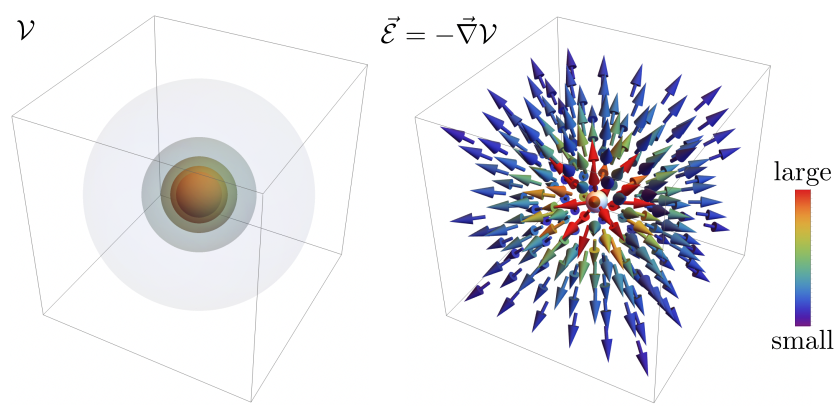


Figure 5.1: Minus the gradient of a scalar potential $\mathcal{V}$ is the (static) electric field $\vec{\mathcal{E}}$. The electric source is a static point charge.

where $\rho[\vec{r}\,']$ is the charge density and $\vec{\boldsymbol{\mathscr{r}}} = \vec{r} - \vec{r}\,'$ connects the source point $\vec{r}\,'$ to the field point $\vec{r}$. The (static) **electric field**

$$\epsilon_0 \vec{\mathcal{E}} = -\epsilon_0 \frac{\partial \mathcal{V}}{\partial \vec{r}} = -\epsilon_0 \vec{\nabla} \mathcal{V} = -\int \frac{dq}{4\pi} \vec{\nabla} \left(\frac{1}{\mathscr{r}} \right) = \int \frac{dq}{4\pi \mathscr{r}^2} \hat{\boldsymbol{\mathscr{r}}}, \tag{5.2}$$

is minus the gradient of the potential energy, which implies **Coulomb's law**. The electric field $\vec{\mathcal{E}}$ pierces surfaces of constant potential $\mathcal{V}$ from high to low, as in Fig. 5.1.

A charge in motion generates a magnetic field. The magnitude and direction of the motion suggests that the magnetic potential is a *vector* field. In analogy with the Eq. 5.1 electric potential, the **magnetic vector potential**

$$\mu_0^{-1} \vec{\mathcal{A}} = \int \frac{\vec{v}\, dq}{4\pi \mathscr{r}} = \int \frac{\vec{I} d\ell}{4\pi \mathscr{r}} = \iiint \frac{\vec{J} dV}{4\pi \mathscr{r}}, \tag{5.3}$$

where $\vec{J}[\vec{r}\,']$ is the current density. The vector nature of the magnetic potential suggests that the magnetic field is the gradient cross product of the potential. Indeed, the **magnetic field**

$$\mu_0^{-1} \vec{\mathcal{B}} = \mu_0^{-1} \vec{\nabla} \times \vec{\mathcal{A}} = \iiint \frac{1}{4\pi} \vec{\nabla} \left(\frac{1}{\mathscr{r}} \right) \times \vec{J}\, dV = \iiint \frac{\vec{J} dV \times \hat{\boldsymbol{\mathscr{r}}}}{4\pi \mathscr{r}^2} = \int \frac{\vec{I} d\ell \times \hat{\boldsymbol{\mathscr{r}}}}{4\pi \mathscr{r}^2} \tag{5.4}$$

is the **curl** of the magnetic potential, which implies **Biot-Savart's law**. The magnetic field $\vec{\mathcal{B}}$ curls around the potential $\vec{\mathcal{A}}$, as in Fig. 5.2.

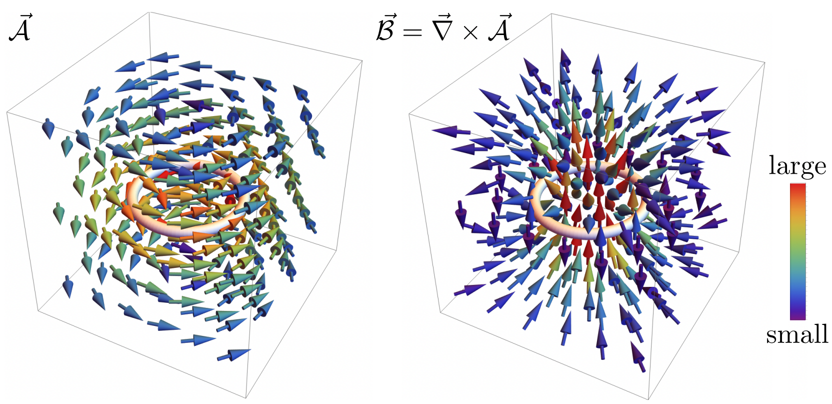


Figure 5.2: The curl of a vector potential $\vec{\mathcal{A}}$ is the magnetic field $\vec{\mathcal{B}}$. The magnetic source is a stationary ring current.

5.2 Momentum

Think of the electric potential $\mathcal{V}$ as the energy per unit charged stored in the electric field. Think of the magnetic potential $\vec{\mathcal{A}}$ as the momentum per unit charge stored in the magnetic field.

5.2.1 Mechanical & Field Momentum

A cylindrical solenoid of radius R, current I, and n turns per unit length, as in Fig. 5.3, has a uniform interior magnetic field and zero exterior magnetic field. **Ampère's law** applied to a rectangular contour parallel to the interior magnetic field and straddling the side

$$0 + 0 + 0 + \mathcal{B}_z \ell = \oint \vec{\mathcal{B}} \cdot d\vec{\ell} = \mathit{\Gamma}_{\mathcal{B}} = \mu_0 I_{\mathit{\Sigma}} = \mu_0 n \ell I \tag{5.5}$$

implies

$$\vec{\mathcal{B}} = \mu_0 n I \hat{z}. \tag{5.6}$$

Stokes' theorem implies that the magnetic flux

$$\mathit{\Phi}_{\mathcal{B}} = \iint_a \vec{\mathcal{B}} \cdot d\vec{a} = \iint_a \vec{\nabla} \times \vec{\mathcal{A}} \cdot d\vec{a} = \oint_{\ell = \partial a} \vec{\mathcal{A}} \cdot d\vec{\ell} = \mathcal{A}_\phi (2\pi s), \tag{5.7}$$

so the magnetic vector potential a perpendicular separation $s > R$ from the solenoid axis

$$\vec{\mathcal{A}} = \frac{\mathit{\Phi}_{\mathcal{B}}}{2\pi s} \hat{\phi} \tag{5.8}$$

is nonzero, where $\{s, \phi, z\}$ are cylindrical coordinates .

Next add a line charge $\lambda = dq/d\ell$ parallel to the solenoid at a distance $x > R$ from its axis [4], as in Fig. 5.3. **Gauss's law** applied to a cylinder of radius x concentric with the line charge

$$0 + \mathcal{E}(2\pi x)\ell + 0 = \oiint \vec{\mathcal{E}} \cdot d\vec{a} = \mathit{\Phi}_E = \epsilon_0^{-1} Q = \epsilon_0^{-1} \lambda \ell \tag{5.9}$$

implies

$$\epsilon_0 \mathcal{E}_x = -\frac{\lambda}{2\pi x}. \tag{5.10}$$

at the solenoid's axis.

Finally, deactivate the solenoid (slowly enough to neglect radiation), and the line charge receives an impulse, as in Fig. 5.3. Qualitatively, a changing magnetic field induces a circulating electric field that forces the line charge to move. Quantitatively, **Faraday's law** implies

$$\mathcal{E}_\phi (2\pi s) = \oint \vec{\mathcal{E}} \cdot d\vec{\ell} = \mathit{\Gamma}_{\mathcal{E}} = -\frac{d\mathit{\Phi}_{\mathcal{B}}}{dt}, \tag{5.11}$$

where the total magnetic flux $\mathit{\Phi}_{\mathcal{B}} = \mathcal{B}A = \mathcal{B}\pi R^2$. The circulating electric field

$$\mathcal{E}_\phi = -\frac{1}{2\pi s} \frac{d\mathit{\Phi}_{\mathcal{B}}}{dt} > 0, \tag{5.12}$$

and a length ℓ of charge $q = \lambda\ell$ experiences a force

$$\frac{dp_\phi}{dt} = F_\phi = q\mathcal{E}_\phi = -\frac{q}{2\pi s}\frac{d\Phi_\mathcal{B}}{dt}. \tag{5.13}$$

Assume the charge does not move far during the magnetic field decay, and integrate

$$\int_0^{p_\phi} dp_\phi = -\frac{q}{2\pi s}\int_{\Phi_\mathcal{B}}^0 d\Phi_\mathcal{B} \tag{5.14}$$

to find the final momentum

$$p_\phi = \frac{q}{2\pi s}\Phi_\mathcal{B} = q\mathcal{A}_\phi > 0. \tag{5.15}$$

On the solenoid axis, the momentum density

$$\frac{p_y}{A\ell} = \frac{\lambda\ell}{2\pi x}\frac{\mathcal{B}A}{A\ell} = \frac{\lambda}{2\pi x}\mathcal{B}_z = -\epsilon_0\mathcal{E}_x\mathcal{B}_z. \tag{5.16}$$

In general,

$$\frac{d\vec{p}_e}{dV} = \epsilon_0\vec{\mathcal{E}}\times\vec{\mathcal{B}}. \tag{5.17}$$

Like an electromagnetic wave storing momentum in its crossed electric and magnetic fields (to enable sunlight to shape a comet's dust tail, for example), the charge and solenoid store momentum in the crossed electric $\vec{\mathcal{E}}$ and magnetic $\vec{\mathcal{B}}$ fields inside the latter. But how does this momentum transfer from the solenoid to the charge, especially with no magnetic field outside the solenoid? *The momentum is stored at the charge potentially as $q\vec{\mathcal{A}}$ and actualized when the changing magnetic field induces a circulating electric field that forces the charge*, as in Fig. 5.4.

Generally, **total** or **canonical momentum**

$$\vec{P} = \vec{p} + q\vec{\mathcal{A}} \tag{5.18}$$

is the sum of the familiar **mechanical** or **kinetic momentum** $\vec{p} = \gamma m\vec{v} \sim m\vec{v}$, $v \ll c$, and the **potential** or **field momentum** $q\vec{\mathcal{A}}$. Sometimes write the mechanical momentum

$$\vec{\Pi} = \vec{P} - q\vec{\mathcal{A}} = \vec{p}. \tag{5.19}$$

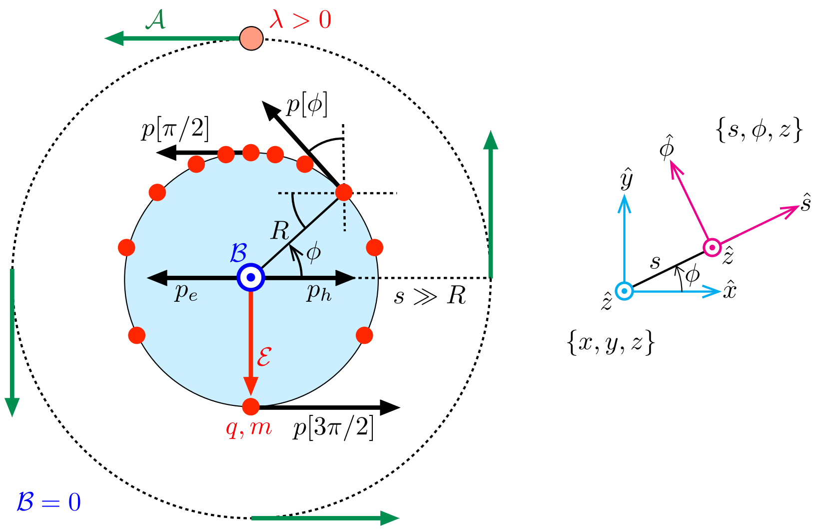


Figure 5.3: A solenoid has outward magnetic field $\vec{\mathcal{B}}$ and circulating magnetic potential $\vec{\mathcal{A}}$. A distant line charge $\lambda = dq/d\ell$ alternately accelerates and decelerates the solenoid's circulating surface charges. The electromagnetic momentum $\vec{p}_e$ stored in the solenoid's crossed electric and magnetic fields balances the hidden momentum $\vec{p}_h$ stored in the asymmetric currents.

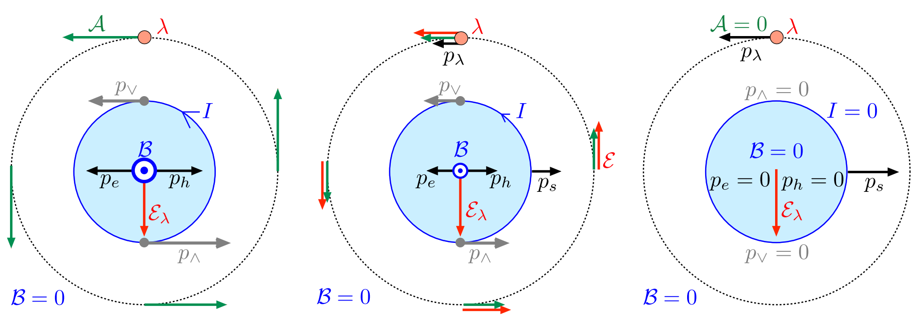


Figure 5.4: Decreasing the magnetic field $\vec{\mathcal{B}}$ induces a circulating electric field $\vec{\mathcal{E}}$, which pushes the line charge λ as it transfers field momentum $q\vec{\mathcal{A}}$ to mechanical momentum $\vec{p}_\lambda$. Sum of electromagnetic, hidden, solenoid, and line charge momenta is always zero, $\vec{p}_e + \vec{p}_h + \vec{p}_s + \vec{p}_\lambda = \vec{0}$. Finally, the line charge recoils leftward and the solenoid recoils rightward.

5.2.2 Hidden Momentum

The line charge electric field accelerates solenoid charges as they move away from it and decelerates them as they move toward it, so charges on the far side move fast and are far apart, while charges on the near side move slow and are close together. At constant current, the momenta would balance, but the nonlinear relativistic stretch $\gamma = 1/\sqrt{1 - v^2/c^2}$ makes fast faster.

The constant current

$$I = \frac{dQ}{dt} = \frac{d(Nq)}{d\ell}\frac{d\ell}{dt} = q\frac{dN}{d\ell}\frac{d\ell}{dt} = qnv = qn[\phi]v[\phi], \tag{5.20}$$

where $n[\phi]$ and $v[\phi]$ are the angle-dependent charge density and speed. In an infinitesimal arc length $d\ell = Rd\phi$, the infinitesimal momentum

$$d\vec{p} = \gamma dM\vec{v} = \gamma\frac{d(Nm)}{d\ell}\vec{v} = \gamma nm\, d\ell\, \vec{v}, \tag{5.21}$$

so the total x momentum

$$\begin{aligned} p_x = \oint dp_x = \oint \gamma nm\, d\ell\, v_x = \int_0^{2\pi} \gamma nm\,(Rd\phi)\,(-v\sin\phi) \\ = -\frac{mIR}{q}\int_0^{2\pi} d\phi\,\gamma[\phi]\sin\phi, \end{aligned} \tag{5.22}$$

which vanishes in the non-relativistic $\gamma = 1$ limit. Meanwhile, as the solenoid charges move from near to far, as in Fig. 5.3, the line charge electric field does work that change their energies by

$$q\mathcal{E}R(1 + \sin\phi) = fd = W = \Delta E = \gamma\left[\frac{\pi}{2}\right]mc^2 - \gamma[\phi]mc^2, \tag{5.23}$$

so the product

$$\gamma[\phi]\sin\phi = \left(\gamma\left[\frac{\pi}{2}\right] - \frac{q\mathcal{E}R}{mc^2}\right)\sin\phi - \frac{q\mathcal{E}R}{mc^2}\sin^2\phi. \tag{5.24}$$

Substitute Eq. 5.24 into Eq. 5.22 to get

$$p_x = -\frac{mIR}{q}\left(0 - \frac{q\mathcal{E}R}{mc^2}\pi\right) = \frac{1}{c^2}\left(I\pi R^2\right)\mathcal{E} = -\epsilon_0\mu_0\mu_z\mathcal{E}_y, \tag{5.25}$$

since the integrals over a period of $\sin\phi$ and $\sin^2\phi$ are 0 and π. More generally, the momentum **hidden** in the relativistic movement of the solenoid charges

$$\vec{p}_h = -\epsilon_0\mu_0\vec{\mu}\times\vec{\mathcal{E}} = -\vec{p}_e \tag{5.26}$$

exactly balances electromagnetic momenta stored in the crossed electric and magnetic fields

$$\vec{p}_h + \vec{p}_e = \vec{0}. \tag{5.27}$$

Problems

1. Compute the hidden momentum of a solenoid of rectangular cross section $\ell \times w$. (Hint: Similar but simpler than the Section 5.2.2 circular case, as the integrals become sums and differences.)

Chapter 6

Dirac Equation

In 1928, Paul Dirac [1] discovered a relativistic wave equation $i\hbar\partial_t\Psi = H\Psi$ that naturally predicts $g = 2$ when coupled to an electromagnetic field.

6.1 Free Electron

Recall that applying the nonrelativistic energy-momentum relation

$$E = \frac{1}{2}mv^2 = \frac{p^2}{2m} = \frac{\vec{p}\cdot\vec{p}}{2m} = H \tag{6.1}$$

with the Eq. 3.34 operator substitutions

$$E \to \mathring{E} = +i\hbar\partial_t, \tag{6.2a}$$

$$\vec{p} \to \mathring{\vec{p}} = -i\hbar\vec{\nabla} \tag{6.2b}$$

to a wave function $\Psi[t, \vec{r}]$

$$i\hbar\partial_t\Psi = -\frac{\hbar^2}{2m}\nabla^2\Psi \tag{6.3}$$

generates the free-particle Schrödinger wave equation. The corresponding relativistic energy-momentum relation

$$E^2 = (pc)^2 + (mc^2)^2 = H^2 \tag{6.4}$$

makes similarly generating a relativistic wave equation tricky. To include time and space symmetrically as first-order derivatives, boldly represent the square root of Eq. 6.4 as a linear function of the momentum

$$H = \sqrt{(pc)^2 + (mc^2)^2} = \vec{\alpha}\cdot\vec{p}c + \beta mc^2. \tag{6.5}$$

Equate the squares of both sides and expand without assuming $\alpha_x, \alpha_y, \alpha_z, \beta$ commute to get

$$\begin{aligned} H^2 = (pc)^2 + (mc^2)^2 &= (\vec{\alpha} \cdot \vec{p}c + \beta mc^2)^2 \\ &= (\vec{\alpha} \cdot \vec{p}c)^2 + (\vec{\alpha} \cdot \vec{p}c)\left(\beta mc^2\right) \\ &+ \left(\beta mc^2\right)(\vec{\alpha} \cdot \vec{p}c) + \beta^2 m^2 c^4. \end{aligned} \tag{6.6}$$

Comparing first terms

$$p^2 = (\vec{\alpha} \cdot \vec{p})^2 \tag{6.7}$$

or

$$\begin{aligned} p_x^2 + p_y^2 + p_z^2 &= (\alpha_x p_x + \alpha_y p_y + \alpha_z p_z)^2 \\ &= \alpha_x^2 p_x^2 + \alpha_x \alpha_y p_x p_y + \alpha_x \alpha_z p_x p_z \\ &+ \alpha_y \alpha_x p_y p_x + \alpha_y^2 p_y^2 + \alpha_y \alpha_z p_y p_z \\ &+ \alpha_z \alpha_x p_z p_x + \alpha_z \alpha_y p_z p_y + \alpha_z^2 p_z^2 \end{aligned} \tag{6.8}$$

implies

$$\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = 1 \tag{6.9}$$

and

$$\alpha_x \alpha_y = -\alpha_y \alpha_x, \tag{6.10a}$$
$$\alpha_y \alpha_z = -\alpha_z \alpha_y, \tag{6.10b}$$
$$\alpha_z \alpha_x = -\alpha_x \alpha_z. \tag{6.10c}$$

Comparing cross terms implies

$$\vec{\alpha}\beta = -\beta\vec{\alpha} \tag{6.11}$$

or

$$\alpha_x \beta = -\beta \alpha_x, \tag{6.12a}$$
$$\alpha_y \beta = -\beta \alpha_y, \tag{6.12b}$$
$$\alpha_z \beta = -\beta \alpha_z. \tag{6.12c}$$

Comparing last terms implies

$$\beta^2 = 1. \tag{6.13}$$

Thus, $\alpha_x, \alpha_y, \alpha_z, \beta$ are not complex numbers but realize an **abstract algebra** of anti-commuting unit squares. **Represent** this algebra most simply by the 2×2 **block matrices**

$$\vec{\alpha} = \boxed{\begin{matrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{matrix}}, \tag{6.14a}$$

$$\beta = \boxed{\begin{matrix} I & 0 \\ 0 & I \end{matrix}}, \tag{6.14b}$$

where the elements themselves are the 2×2 matrices

$$I = \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}}, \tag{6.15a}$$

$$\sigma_x = \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}}, \tag{6.15b}$$

$$\sigma_y = \boxed{\begin{matrix} 0 & -\boldsymbol{i} \\ \boldsymbol{i} & 0 \end{matrix}}, \tag{6.15c}$$

$$\sigma_z = \boxed{\begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix}}, \tag{6.15d}$$

where "$-\boldsymbol{i}$ rides high on σ_y". The **Pauli matrices** satisfy the same algebra

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = -\boldsymbol{i}\sigma_x\sigma_y\sigma_z = I. \tag{6.16}$$

as the **quaternions**. For example,

$$\sigma_x^2 = \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}}\boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}} = \boxed{\begin{matrix} 0\cdot 0 + 1\cdot 1 & 1\cdot 0 + 0\cdot 1 \\ 0\cdot 1 + 1\cdot 0 & 1\cdot 1 + 0\cdot 0 \end{matrix}} = \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} = I \tag{6.17}$$

and

$$\sigma_x\sigma_y\sigma_z = \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}}\boxed{\begin{matrix} 0 & -\boldsymbol{i} \\ \boldsymbol{i} & 0 \end{matrix}}\boxed{\begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix}} = \boxed{\begin{matrix} \boldsymbol{i} & 0 \\ 0 & -\boldsymbol{i} \end{matrix}}\boxed{\begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix}} = \boxed{\begin{matrix} \boldsymbol{i} & 0 \\ 0 & \boldsymbol{i} \end{matrix}} = \boldsymbol{i}I. \tag{6.18}$$

With the Eq. 6.14 block matrices, the Dirac equation

$$\boldsymbol{i}\hbar\partial_t\Psi = \left(\vec{\alpha}\cdot\mathring{\vec{p}}c + \beta m c^2\right)\Psi, \tag{6.19}$$

where $\alpha_x p_x c\Psi = \alpha_x(-\boldsymbol{i}\hbar\partial_x(c\Psi)) = -\boldsymbol{i}\hbar c\alpha_x\partial_x\Psi$, for example. Introduce "Large" and "Small" components

$$\Psi = \boxed{\begin{matrix} \Psi_1 \\ \Psi_2 \\ \Psi_1 \\ \Psi_2 \end{matrix}} = \boxed{\begin{matrix} \Psi_L \\ \Psi_S \end{matrix}} \tag{6.20}$$

and write

$$i\hbar\partial_t \begin{bmatrix} \Psi_L \\ \Psi_S \end{bmatrix} = \begin{bmatrix} Imc^2 & \vec{\sigma}\cdot\vec{\mathring{p}}c \\ \vec{\sigma}\cdot\vec{\mathring{p}}c & -Imc^2 \end{bmatrix} \begin{bmatrix} \Psi_L \\ \Psi_S \end{bmatrix}. \tag{6.21}$$

With the dot product

$$\begin{aligned} \vec{\sigma}\cdot\vec{p} &= \sigma_x p_x + \sigma_y p_y + \sigma_z p_z \\ &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} p_x + \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} p_y + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} p_z \\ &= \begin{bmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{bmatrix}, \end{aligned} \tag{6.22}$$

fully expand the Dirac equation

$$i\hbar\partial_t \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{bmatrix} = \begin{bmatrix} mc^2 & 0 & \mathring{p}_z c & \mathring{p}_x c - i\mathring{p}_y c \\ 0 & mc^2 & \mathring{p}_x c + i\mathring{p}_y c & -\mathring{p}_z c \\ \mathring{p}_z c & \mathring{p}_x c - i\mathring{p}_y c & -mc^2 & 0 \\ \mathring{p}_x c + i\mathring{p}_y c & -\mathring{p}_z c & 0 & -mc^2 \end{bmatrix} \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{bmatrix} \tag{6.23}$$

and as four complex equations

$$i\hbar\partial_t\Psi_1 = mc^2\Psi_1 + \mathring{p}_z c\Psi_3 + \mathring{p}_x c\Psi_4 - i\mathring{p}_y c\Psi_4, \tag{6.24a}$$
$$i\hbar\partial_t\Psi_2 = mc^2\Psi_2 + \mathring{p}_x c\Psi_3 + i\mathring{p}_y c\Psi_3 - \mathring{p}_z c\Psi_4, \tag{6.24b}$$
$$i\hbar\partial_t\Psi_3 = \mathring{p}_z c\Psi_1 + \mathring{p}_x c\Psi_2 - i\mathring{p}_y c\Psi_2 - mc^2\Psi_3, \tag{6.24c}$$
$$i\hbar\partial_t\Psi_4 = \mathring{p}_x c\Psi_1 + i\mathring{p}_y c\Psi_1 - \mathring{p}_z c\Psi_2 - mc^2\Psi_4, \tag{6.24d}$$

and with explicit spatial derivatives

$$i\hbar\partial_t\Psi_1 = \quad mc^2\Psi_1 - i\hbar c\partial_z\Psi_3 - i\hbar c\partial_x\Psi_4 + i\hbar c\partial_y\Psi_4, \tag{6.25a}$$
$$i\hbar\partial_t\Psi_2 = \quad mc^2\Psi_2 - i\hbar c\partial_x\Psi_3 - i\hbar c\partial_y\Psi_3 + i\hbar c\partial_z\Psi_4, \tag{6.25b}$$
$$i\hbar\partial_t\Psi_3 = -i\hbar c\partial_z\Psi_1 - i\hbar c\partial_x\Psi_2 + i\hbar c\partial_y\Psi_2 - mc^2\Psi_3, \tag{6.25c}$$
$$i\hbar\partial_t\Psi_4 = -i\hbar c\partial_x\Psi_1 - i\hbar c\partial_y\Psi_1 + i\hbar c\partial_z\Psi_2 - mc^2\Psi_4. \tag{6.25d}$$

In the rest frame, the four components can represent a spin “up” or spin “down” electron or anti-electron; informally,

$$\Psi = \boxed{\begin{matrix} \uparrow e^- \\ \downarrow e^- \\ \uparrow e^+ \\ \downarrow e^+ \end{matrix}}. \tag{6.26}$$

6.2 Interacting Electron

6.2.1 Pauli Equation

In an electromagnetic field, the free-particle Schrödinger equation

$$\boldsymbol{i}\hbar\partial_t\Psi = \mathring{H}\Psi = \frac{\mathring{\vec{p}}^2}{2m}\Psi \tag{6.27}$$

becomes

$$\boldsymbol{i}\hbar\partial_t\Psi = \frac{\mathring{\vec{p}}^2}{2m}\Psi + q\mathcal{V}\Psi, \tag{6.28}$$

where the Eq. 5.19 mechanical momentum

$$\mathring{\vec{p}} = \mathring{\vec{P}} - q\vec{\mathcal{A}} = -\boldsymbol{i}\hbar\vec{\nabla} - q\vec{\mathcal{A}}, \tag{6.29}$$

and $\mathcal{V}$ and $\vec{\mathcal{A}}$ are the electric and magnetic potentials. Similarly, the free-particle Dirac equation

$$\boldsymbol{i}\hbar\partial_t\Psi = \mathring{H}\Psi = \left(\vec{\alpha}\cdot\mathring{\vec{p}}c + \beta mc^2\right)\Psi \tag{6.30}$$

becomes

$$\boldsymbol{i}\hbar\partial_t\Psi = \left(\vec{\alpha}\cdot\mathring{\vec{p}}c + \beta mc^2 + q\mathcal{V}\right)\Psi. \tag{6.31}$$

Assume no electric field, so $\mathcal{V} = 0$, and seek stationary solutions of constant energy

$$\Psi[\vec{r}, t] = \psi[\vec{r}]e^{-\boldsymbol{i}Et/\hbar} \tag{6.32}$$

to get

$$E\psi = \left(\vec{\alpha}\cdot\mathring{\vec{p}}c + \beta mc^2\right)\psi. \tag{6.33}$$

Introduce “large” and “small” time-independent two-component **spinors**, as in Eq. 6.21, to write

$$E\boxed{\begin{matrix}\psi_L \\ \psi_S\end{matrix}} = \boxed{\begin{matrix} Imc^2 & \vec{\sigma}\cdot\mathring{\vec{p}}c \\ \vec{\sigma}\cdot\mathring{\vec{p}}c & -Imc^2 \end{matrix}}\,\boxed{\begin{matrix}\psi_L \\ \psi_S\end{matrix}}. \tag{6.34}$$

Move the spinors to the left side

$$\boxed{\begin{matrix} IE & 0 \\ 0 & IE \end{matrix}}\ \boxed{\begin{matrix} \psi_L \\ \psi_S \end{matrix}} - \boxed{\begin{matrix} Imc^2 & \vec{\sigma}\cdot\mathring{\vec{p}}c \\ \vec{\sigma}\cdot\mathring{\vec{p}}c & -Imc^2 \end{matrix}}\ \boxed{\begin{matrix} \psi_L \\ \psi_S \end{matrix}} = \boxed{\begin{matrix} 0 \\ 0 \end{matrix}} \tag{6.35}$$

and consolidate

$$\boxed{\begin{matrix} I(E-mc^2) & -\vec{\sigma}\cdot\mathring{\vec{p}}c \\ -\vec{\sigma}\cdot\mathring{\vec{p}}c & I(E+mc^2) \end{matrix}}\ \boxed{\begin{matrix} \psi_L \\ \psi_S \end{matrix}} = \boxed{\begin{matrix} 0 \\ 0 \end{matrix}} \tag{6.36}$$

to get the two complex equations

$$\left(E - mc^2\right)\psi_L - \vec{\sigma}\cdot\mathring{\vec{p}}c\,\psi_S = 0, \tag{6.37a}$$

$$\left(E + mc^2\right)\psi_S - \vec{\sigma}\cdot\mathring{\vec{p}}c\,\psi_L = 0. \tag{6.37b}$$

In the non-relativistic limit,

$$E + mc^2 = E - mc^2 + 2mc^2 = E_N + 2mc^2 \sim 2mc^2, \tag{6.38}$$

where the non-relativistic energy that appears in the Schrödinger equation $E_N \ll mc^2$. Assume the state

$$\psi_L \sim N e^{i\vec{p}\cdot\vec{r}/\hbar} = N e^{im\vec{v}\cdot\vec{r}/\hbar} \tag{6.39}$$

has a typical momentum

$$\mathring{\vec{p}}\psi_L = -i\hbar\vec{\nabla}\psi_L \sim m\vec{v}\,\psi_L. \tag{6.40}$$

Hence, the bottom Eq. 6.37b implies the small component

$$\psi_S = \frac{\vec{\sigma}\cdot\mathring{\vec{p}}c}{E+mc^2}\psi_L \sim \frac{mvc}{2mc^2}\psi_L = \frac{1}{2}\left(\frac{v}{c}\right)\psi_L \ll \psi_L \tag{6.41}$$

is much smaller than the large component, and so the top Eq. 6.37a implies

$$E_N\psi_L = \left(E - mc^2\right)\psi_L = \vec{\sigma}\cdot\mathring{\vec{p}}c\,\psi_S = \frac{\left(\vec{\sigma}\cdot\mathring{\vec{p}}c\right)\left(\vec{\sigma}\cdot\mathring{\vec{p}}c\right)}{E+mc^2}\psi_L \sim \frac{\left(\vec{\sigma}\cdot\mathring{\vec{p}}\right)^2}{2m}\psi_L. \tag{6.42}$$

In 1927, Wolfgang Pauli [6] first formulated the corresponding time-dependent equation

$$i\hbar\partial_t\Psi = \frac{\left(\vec{\sigma}\cdot\mathring{\vec{p}}\right)^2}{2m}\Psi \tag{6.43}$$

as an ad hoc explanation for electron spin and magnetic moment.

6.2.2 Two Famous Identities

For constant vectors $\vec{u}$ and $\vec{v}$, the Eq. 6.16 Pauli algebra implies the algebraic identity

$$
\begin{aligned}
(\vec{\sigma}\cdot\vec{u})(\vec{\sigma}\cdot\vec{v}) &= (u_x\sigma_x + u_y\sigma_y + u_z\sigma_z)(v_x\sigma_x + v_y\sigma_y + v_z\sigma_z) \\
&= u_x v_x \sigma_x^2 + u_x v_y \sigma_x \sigma_y + u_x v_z \sigma_x \sigma_z \\
&\quad + u_y v_x \sigma_y \sigma_x + u_y v_y \sigma_y^2 + u_y v_z \sigma_y \sigma_z \\
&\quad + u_z v_x \sigma_z \sigma_x + u_z v_y \sigma_z \sigma_y + u_z v_z \sigma_z^2 \\
&= I u_x v_x + \boldsymbol{i} u_x v_y \sigma_z - \boldsymbol{i} u_x v_z \sigma_y \\
&\quad - \boldsymbol{i} u_y v_x \sigma_z + I u_y v_y + \boldsymbol{i} u_y v_z \sigma_x \\
&\quad + \boldsymbol{i} u_z v_x \sigma_y - \boldsymbol{i} u_z v_y \sigma_x + I u_z v_z \\
&= I\,(u_x v_x + u_y v_y + u_z v_z) \\
&\quad + \boldsymbol{i}\big(\,(u_y v_z - u_z v_y)\,\sigma_x \\
&\qquad + (u_z v_x - u_x v_z)\,\sigma_y \\
&\qquad + (u_x v_y - u_y v_x)\,\sigma_z\big) \\
&= I\vec{u}\cdot\vec{v} + \boldsymbol{i}\vec{u}\times\vec{v}\cdot\vec{\sigma}.
\end{aligned} \tag{6.44}
$$

For mechanical momentum $\mathring{\vec{p}} = \mathring{\vec{P}} - q\vec{\mathcal{A}}$ and for any wave function $\psi[\vec{r}]$, careful application of a vector derivative product rule implies the differential identity

$$
\begin{aligned}
\mathring{\vec{p}}\times\mathring{\vec{p}}\,\psi &= (-\boldsymbol{i}\hbar\vec{\nabla} - q\vec{\mathcal{A}})\times(-\boldsymbol{i}\hbar\vec{\nabla} - q\vec{\mathcal{A}})\psi \\
&= (-\boldsymbol{i}\hbar\vec{\nabla} - q\vec{\mathcal{A}})\times(-\boldsymbol{i}\hbar\vec{\nabla}\psi - q\vec{\mathcal{A}}\psi) \\
&= -\vec{\nabla}\times\vec{\nabla}\psi + \boldsymbol{i}\hbar q\vec{\nabla}\times\left(\vec{\mathcal{A}}\psi\right) + \boldsymbol{i}\hbar q\vec{\mathcal{A}}\times\vec{\nabla}\psi + \hbar^2 q^2\vec{\mathcal{A}}\times\vec{\mathcal{A}}\psi \\
&= \vec{0} + \boldsymbol{i}\hbar q\left(\left(\vec{\nabla}\times\vec{\mathcal{A}}\right)\psi + \left(\vec{\nabla}\psi\right)\times\vec{\mathcal{A}}\right) + \boldsymbol{i}\hbar q\vec{\mathcal{A}}\times\vec{\nabla}\psi + \vec{0} \\
&= +\boldsymbol{i}\hbar q\left(\psi\vec{\nabla}\times\vec{\mathcal{A}} - \vec{\mathcal{A}}\times\vec{\nabla}\psi\right) + \boldsymbol{i}\hbar q\vec{\mathcal{A}}\times\vec{\nabla}\psi \\
&= \boldsymbol{i}\hbar q\vec{\nabla}\times\vec{\mathcal{A}}\psi \\
&= \boldsymbol{i}\hbar q\vec{\mathcal{B}}\psi
\end{aligned} \tag{6.45}
$$

and so

$$\mathring{\vec{p}}\times\mathring{\vec{p}} = \boldsymbol{i}\hbar q\vec{\mathcal{B}}. \tag{6.46}$$

6.2.3 Electron Magnetic Moment

The Section 6.2.2 identities imply

$$\left(\vec{\sigma}\cdot\mathring{\vec{p}}\right)^2 = \mathring{\vec{p}}\cdot\mathring{\vec{p}} + i\vec{\sigma}\cdot\mathring{\vec{p}}\times\mathring{\vec{p}} = \mathring{\vec{p}}^2 - q\hbar\vec{\sigma}\cdot\vec{\mathcal{B}} \tag{6.47}$$

so the time-independent "large" Eq. 6.42 becomes

$$\begin{aligned} E_N\psi_L &= \frac{\mathring{\vec{p}}^2 - q\hbar\vec{\sigma}\cdot\vec{\mathcal{B}}}{2m}\psi_L \\ &= \frac{\mathring{\vec{p}}^2}{2m}\psi_L - \frac{q\hbar}{2m}\vec{\sigma}\cdot\vec{\mathcal{B}}\psi_L \\ &= \frac{\mathring{\vec{p}}^2}{2m}\psi_L + U\psi_L, \end{aligned} \tag{6.48}$$

with the corresponding time-dependent equation

$$i\hbar\partial_t\Psi_L = \frac{\mathring{\vec{p}}^2}{2m}\Psi_L + U\Psi_L, \tag{6.49}$$

where the potential energy

$$U = -\vec{\mu}\cdot\vec{\mathcal{B}} = -2\frac{q}{2m}\frac{\hbar}{2}\vec{\sigma}\cdot\vec{\mathcal{B}}, \tag{6.50}$$

and the magnetic moment

$$\vec{\mu} = g\frac{q}{2m}\vec{S}, \tag{6.51}$$

and the spin angular momentum

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma}, \tag{6.52}$$

and

$$g = 2. \tag{6.53}$$

Problems

1. Verify all of the Eq. 6.16 Pauli identities by explicit matrix multiplication.

2. Verify all of the Eq. 6.45 vector derivative expansion of $\vec{\nabla} \times \left(\vec{\mathcal{A}}\psi\right)$ by working in rectangular coordinates. (Hint: Without loss of generality, rotate the coordinates so that the vector points along the x-axis.)

3. Using the block matrices

$$\{\gamma_t, \gamma_x, \gamma_y, \gamma_z\} = \left\{ \boxed{\begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix}}, \boxed{\begin{matrix} 0 & \sigma_x \\ -\sigma_x & 0 \end{matrix}}, \boxed{\begin{matrix} 0 & \sigma_y \\ -\sigma_y & 0 \end{matrix}}, \boxed{\begin{matrix} 0 & \sigma_z \\ -\sigma_z & 0 \end{matrix}} \right\} \tag{6.54}$$

and the **Feynman slash notation**

$$\not{a} = \gamma_t a_t + \gamma_x a_x + \gamma_y a_y + \gamma_z a_z, \tag{6.55}$$

write the Dirac Eq. 6.19 in natural $\hbar = c = 1$ units as in the Fig. 1.1 memorial marker,

$$\boldsymbol{i}\not{\partial}\Psi = m\Psi. \tag{6.56}$$

Chapter 7

Afterword

In 1948, in a short letter to the Physical Review at the dawn of **quantum electrodynamics** (QED), Julian Schwinger [8] claimed that, actually, the electron magnetic dipole moment

$$g = 2\left(1 + \frac{\alpha}{2\pi} + \cdots\right) = 2.0023\ldots, \tag{7.1}$$

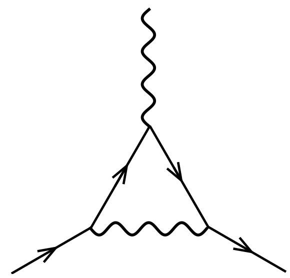
where the **fine structure constant**

$$\alpha = \left(\frac{q_e}{q_P}\right)^2 = \frac{q_e^2}{4\pi\epsilon_0\hbar c} \approx 0.0072973 \approx \frac{1}{137} \ll 1 \tag{7.2}$$

and the **Planck charge** $q_P \approx 1.8755$ aC, which was subsequently verified experimentally. Schwinger's triumph is commemorated in his Fig. 7.1 tombstone.

Figure 7.1: Julian Schwinger's tombstone includes his first-order correction to $g = 2$. (Creative Commons credit: Jacob Bourjaily, 2013.)

In 1949, while “bringing QED to the masses”, Richard Feynman [3] represented the electron **anomalous magnetic moment** by the diagram



Appendix A

Used Math

Briefly review the mathematics used in this text

A.1 Calculus in Higher Dimensions

A.1.1 Partial Derivatives

In one-dimension, the function

$$f[x] = 3x^2 + 1 \tag{A.1}$$

has the **derivative**

$$\frac{df}{dx} = 6x + 0 = 6x. \tag{A.2}$$

In two dimensions, the function

$$f[x, y] = 3xy^2 + 2x + 3y + 2 \tag{A.3}$$

has the **partial derivatives**

$$\frac{\partial f}{\partial x} = 3y^2 + 2 + 0 + 0 = 3y^2 + 2, \tag{A.4a}$$

$$\frac{\partial f}{\partial y} = 6xy + 0 + 3 + 2 = 6xy + 5, \tag{A.4b}$$

which are just like ordinary derivatives, but with other variables held constant.

A.1.2 Vector Derivatives

In three dimensions, the **gradient** of a scalar function

$$\begin{aligned}\vec{\nabla} f &= \hat{x}\frac{\partial f}{\partial x} + \hat{y}\frac{\partial f}{\partial y} + \hat{z}\frac{\partial f}{\partial z} \\ &= \hat{x}\,\partial_x f + \hat{y}\,\partial_y f + \hat{z}\,\partial_z f \end{aligned} \tag{A.5}$$

is the direction and rate of fastest increase of the function, where $\hat{x}$ is a unit vector in the direction of increasing x, and so on. In analogy with the Greek letters δ and $\mathit{\Delta}$, pronounce the pseudo-letters ∂ and ∇ "del" and "big del" (and enter them as "partial" and "nabla" in LaTeX).

The two complementary kinds of vector products imply two complementary kinds of vector derivatives. In rectangular coordinates, the **divergence**

$$\begin{aligned}
\vec{\nabla} \cdot \vec{v} &= (\hat{x}\,\partial_x + \hat{y}\,\partial_y + \hat{z}\,\partial_z) \cdot (\hat{x}\,v_x + \hat{y}\,v_y + \hat{z}\,v_z) \\
&= \hat{x} \cdot \hat{x}\,\partial_x v_x + \hat{x} \cdot \hat{y}\,\partial_x v_y + \hat{x} \cdot \hat{z}\,\partial_x v_z \\
&\quad + \hat{y} \cdot \hat{x}\,\partial_y v_x + \hat{y} \cdot \hat{y}\,\partial_y v_y + \hat{y} \cdot \hat{z}\,\partial_y v_z \\
&\quad + \hat{z} \cdot \hat{x}\,\partial_z v_x + \hat{z} \cdot \hat{y}\,\partial_z v_y + \hat{z} \cdot \hat{z}\,\partial_z v_z \\
&= \partial_x v_x + 0 + 0 \\
&\quad + 0 + \partial_y v_y + 0 \\
&\quad + 0 + 0 + \partial_z v_z \\
&= \partial_x v_x + \partial_y v_y + \partial_z v_z
\end{aligned} \tag{A.6}$$

is a scalar function, and the **curl**

$$\begin{aligned}
\vec{\nabla} \times \vec{v} &= (\hat{x}\,\partial_x + \hat{y}\,\partial_y + \hat{z}\,\partial_z) \times (\hat{x}\,v_x + \hat{y}\,v_y + \hat{z}\,v_z) \\
&= \hat{x} \times \hat{x}\,\partial_x v_x + \hat{x} \times \hat{y}\,\partial_x v_y + \hat{x} \times \hat{z}\,\partial_x v_z \\
&\quad + \hat{y} \times \hat{x}\,\partial_y v_x + \hat{y} \times \hat{y}\,\partial_y v_y + \hat{y} \times \hat{z}\,\partial_y v_z \\
&\quad + \hat{z} \times \hat{x}\,\partial_z v_x + \hat{z} \times \hat{y}\,\partial_z v_y + \hat{z} \times \hat{z}\,\partial_z v_z \\
&= +\vec{0} + \hat{z}\,\partial_x v_y - \hat{y}\,\partial_x v_z \\
&\quad - \hat{z}\,\partial_y v_x + \vec{0} + \hat{x}\,\partial_y v_z \\
&\quad + \hat{y}\,\partial_z v_x - \hat{x}\,\partial_z v_y + \vec{0} \\
&= \hat{x}\,(\partial_y v_z - \partial_z v_y) + \hat{y}\,(\partial_z v_x - \partial_x v_z) + \hat{z}\,(\partial_x v_y - \partial_y v_x)
\end{aligned} \tag{A.7}$$

is a vector function. Be especially careful not to confuse the gradient and the divergence.

A.1.3 Integrals of Derivatives

In one-dimensional calculus, the integral of the derivative of a function *is* the function. In three-dimensional calculus, three different derivatives imply three analogous results, all special cases of **Stokes'** theorem. The integral of the gradient of a scalar function along a path

$$\int_\ell \vec{\nabla} f \cdot d\ell = f\Big|_{\partial\ell} = f_2 - f_1 \tag{A.8}$$

is the difference of the function at the path's boundary points, where the unadorned ∂ is the **boundary operator**. The integral of the curl of a vector function over an area

$$\iint_a \vec{\nabla} \times \vec{v} \cdot d\vec{a} = \oint_{\ell=\partial a} \vec{v} \cdot d\ell \tag{A.9}$$

is the closed path integral of the function over the area's boundary. The integral of the divergence of a vector function over a volume

$$\iiint_V \vec{\nabla} \cdot \vec{v}\, dV = \oiint_{a=\partial V} \vec{v} \cdot d\vec{a} \tag{A.10}$$

is the closed surface integral of the function over the volume's boundary.

A.1.4 Product Rules

In one-dimensional calculus, the **product rule** for differentiation is

$$\frac{d}{dx}(fg) = \frac{df}{dx}g + f\frac{dg}{dx}. \tag{A.11}$$

In three-dimensional calculus, different kinds of products and derivatives resulting in six different product rules exist. Derive these by carefully employing the mixed "box" product

$$\vec{A} \cdot \left(\vec{B} \times \vec{C}\right) = \vec{C} \cdot \left(\vec{A} \times \vec{B}\right) = \vec{B} \cdot \left(\vec{C} \times \vec{A}\right), \tag{A.12}$$

the vector double product

$$\vec{A} \times \left(\vec{B} \times \vec{C}\right) = \vec{B}\left(\vec{A} \cdot \vec{C}\right) - \vec{C}\left(\vec{A} \cdot \vec{B}\right) \tag{A.13}$$

or

$$\vec{B}\left(\vec{A} \cdot \vec{C}\right) = \vec{A} \times \left(\vec{B} \times \vec{C}\right) + \vec{C}\left(\vec{A} \cdot \vec{B}\right), \tag{A.14}$$

and the linearity of the derivative

$$\vec{\nabla} = \vec{\nabla}_A + \vec{\nabla}_B. \tag{A.15}$$

The **gradient** of the product of two scalar fields is

$$\vec{\nabla}\left(fg\right) = \left(\vec{\nabla} f\right) g + f\left(\vec{\nabla} g\right). \tag{A.16}$$

The gradient of the scalar product of two vector fields is

$$\begin{aligned}
\vec{\nabla}\left(\vec{A} \cdot \vec{B}\right) &= \vec{\nabla}_A \left(\vec{A} \cdot \vec{B}\right) + \vec{\nabla}_B \left(\vec{A} \cdot \vec{B}\right) \\
&= \vec{\nabla}_A \left(\vec{A} \cdot \vec{B}\right) + \vec{\nabla}_B \left(\vec{B} \cdot \vec{A}\right) \\
&= \vec{B} \times \left(\vec{\nabla}_A \times \vec{A}\right) + \left(\vec{B} \cdot \vec{\nabla}_A\right) \vec{A} + \vec{A} \times \left(\vec{\nabla}_B \times \vec{B}\right) + \left(\vec{A} \cdot \vec{\nabla}_B\right) \vec{B} \\
&= \vec{B} \times \left(\vec{\nabla} \times \vec{A}\right) + \left(\vec{B} \cdot \vec{\nabla}\right) \vec{A} + \vec{A} \times \left(\vec{\nabla} \times \vec{B}\right) + \left(\vec{A} \cdot \vec{\nabla}\right) \vec{B}
\end{aligned} \tag{A.17}$$

The **divergence** of the product of a scalar field and a vector field is

$$\vec{\nabla} \cdot \left(f\vec{A} \right) = \left(\vec{\nabla} f \right) \cdot \vec{A} + f \left(\vec{\nabla} \cdot \vec{A} \right). \tag{A.18}$$

The divergence of the vector product of two vector fields is

$$\begin{aligned} \vec{\nabla} \cdot \left(\vec{A} \times \vec{B} \right) &= \vec{\nabla}_A \cdot \left(\vec{A} \times \vec{B} \right) + \vec{\nabla}_B \cdot \left(\vec{A} \times \vec{B} \right) \\ &= \vec{\nabla}_A \cdot \left(\vec{A} \times \vec{B} \right) - \vec{\nabla}_B \cdot \left(\vec{B} \times \vec{A} \right) \\ &= \vec{B} \cdot \left(\vec{\nabla}_A \times \vec{A} \right) - \vec{A} \cdot \left(\vec{\nabla}_B \times \vec{B} \right) \\ &= \vec{B} \cdot \left(\vec{\nabla} \times \vec{A} \right) - \vec{A} \cdot \left(\vec{\nabla} \times \vec{B} \right) \end{aligned} \tag{A.19}$$

The **curl** of the product of a scalar field and a vector field is

$$\vec{\nabla} \times \left(f\vec{A} \right) = \left(\vec{\nabla} f \right) \times \vec{A} + f \left(\vec{\nabla} \times \vec{A} \right). \tag{A.20}$$

The curl of the vector product of two vector fields is

$$\begin{aligned} \vec{\nabla} \times \left(\vec{A} \times \vec{B} \right) &= \vec{\nabla}_A \times \left(\vec{A} \times \vec{B} \right) + \vec{\nabla}_B \times \left(\vec{A} \times \vec{B} \right) \\ &= \vec{\nabla}_A \times \left(\vec{A} \times \vec{B} \right) + \vec{\nabla}_B \times \left(\vec{A} \times \vec{B} \right) \\ &= \left(\vec{B} \cdot \vec{\nabla}_A \right) \vec{A} - \vec{B} \left(\vec{\nabla}_A \cdot \vec{A} \right) + \vec{A} \left(\vec{\nabla}_B \cdot \vec{B} \right) - \left(\vec{A} \cdot \vec{\nabla}_B \right) \vec{B} \\ &= \left(\vec{B} \cdot \vec{\nabla} \right) \vec{A} - \vec{B} \left(\vec{\nabla} \cdot \vec{A} \right) + \vec{A} \left(\vec{\nabla} \cdot \vec{B} \right) - \left(\vec{A} \cdot \vec{\nabla} \right) \vec{B} \end{aligned} \tag{A.21}$$

A.2 Complex Numbers

Complex numbers complete real numbers. The algebraic equation $x^2 + 1 = 0$ has no real solutions. *Imagine* that it has solutions $\pm \boldsymbol{i}$. By successive multiplication, the imaginary number $\boldsymbol{i}$ satisfies $\boldsymbol{i}^2 = -1$, $\boldsymbol{i}^3 = -\boldsymbol{i}$, $\boldsymbol{i}^4 = +1$, $\boldsymbol{i}^5 = +\boldsymbol{i}$, $\boldsymbol{i}^6 = -1$, and so on in a 4-cycle. Hence the absolutely convergent **Taylor expansions** of common functions dramatically reorganize when evaluated at imaginary numbers. For example, the exponential

$$\begin{aligned} e^{\boldsymbol{i}x} &= 1 + \boldsymbol{i}x + \frac{(\boldsymbol{i}x)^2}{2!} + \frac{(\boldsymbol{i}x)^3}{3!} + \frac{(\boldsymbol{i}x)^4}{4!} + \frac{(\boldsymbol{i}x)^5}{5!} + \frac{(\boldsymbol{i}x)^6}{6!} + \frac{(\boldsymbol{i}x)^7}{7!} + \frac{(\boldsymbol{i}x)^8}{8!} + \cdots \\ &= 1 + \boldsymbol{i}x - \frac{x^2}{2!} - \boldsymbol{i}\frac{x^3}{3!} + \frac{x^4}{4!} + \boldsymbol{i}\frac{x^5}{5!} - \frac{x^6}{6!} - \boldsymbol{i}\frac{x^7}{7!} + \frac{x^8}{8!} + \cdots \\ &= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots \right) + \boldsymbol{i} \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \right) \\ &= \cos x + \boldsymbol{i} \sin x, \end{aligned} \tag{A.22}$$

becomes a linear superposition of sinusoids known as **Euler's identity**. The choice $x = \pi$ generates the famously beautiful special case

$$e^{\boldsymbol{i}\pi} + 1 = 0, \tag{A.23}$$

which relates the five most important mathematical constants $e, \boldsymbol{i}, \pi, 1, 0$ in a simple formula.

A general complex number is the linear combination

$$z = z_R + z_I \boldsymbol{i} = x + \boldsymbol{i} y = \{x, y\}, \tag{A.24}$$

where the real and imaginary components $x = z_R$ and $y = z_I$ are real numbers. The sum of two complex numbers

$$z + \lambda z' = x + \boldsymbol{i} y + \lambda(x' + \boldsymbol{i} y') = x + \lambda x' + \boldsymbol{i}(y + \lambda y') \tag{A.25}$$

and the product of two complex numbers

$$zz' = (x + \boldsymbol{i} y)(x' + \boldsymbol{i} y') = xx' - yy' + \boldsymbol{i}(yx' + xy') \tag{A.26}$$

are also complex numbers. A complex number's **conjugate**

$$z^* = \bar{z} = x - \boldsymbol{i} y \tag{A.27}$$

negates the imaginary part, so the **norm**

$$|z| = \sqrt{zz^*} = \sqrt{x^2 + y^2} \tag{A.28}$$

is the square root of the product with the conjugate.

A.3 Matrices

Matrices, arrays, or tableaux of numbers have been used to solve math problems for thousands of years. Consider the linear transformation

$$x' = ax + by, \tag{A.29a}$$

$$y' = cx + dy. \tag{A.29b}$$

In **box notation**, collect the variables in **column matrices**

$$\vec{r} = \boxed{\begin{matrix} x \\ y \end{matrix}}\,, \tag{A.30a}$$

$$\vec{r}\,' = \boxed{\begin{matrix} x' \\ y' \end{matrix}}\,, \tag{A.30b}$$

and collect the coefficients in the **square matrix**

$$M = \boxed{\begin{matrix} a & b \\ c & d \end{matrix}}\,, \tag{A.31}$$

and form the equivalent **matrix equation**

$$\begin{array}{|c|}\hline x' \\ y' \\ \hline\end{array} = \begin{array}{|cc|}\hline a & b \\ c & d \\ \hline\end{array} \begin{array}{|c|}\hline x \\ y \\ \hline\end{array}, \tag{A.32}$$

where the color guides the eye in checking the matrix multiplication, and where rows are dot-producted with columns. In **bracket notation**, write

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \tag{A.33}$$

and represent matrix equations symbolically as

$$\vec{r}\,' = M\,\vec{r}. \tag{A.34}$$

The **identity matrix**

$$I = \begin{array}{|cc|}\hline 1 & 0 \\ 0 & 1 \\ \hline\end{array} \neq \begin{array}{|cc|}\hline 1 & 1 \\ 1 & 1 \\ \hline\end{array} \tag{A.35}$$

has 1s on the primary $\searrow$ diagonal and 0s on the secondary $\swarrow$ diagonal.

Concatenate a second linear transformation to the first to get

$$\begin{aligned} x'' &= a'x' + b'y' = a'(ax+by) + b'(cx+dy) = (a'a+b'c)x + (a'b+b'd)y, \\ y'' &= c'x' + d'y' = c'(ax+by) + d'(cx+dy) = (c'a+d'c)x + (c'b+d'd)y, \end{aligned} \tag{A.36}$$

or in matrix notation

$$\begin{aligned} \begin{array}{|c|}\hline x'' \\ y'' \\ \hline\end{array} &= \begin{array}{|cc|}\hline a' & b' \\ c' & d' \\ \hline\end{array} \begin{array}{|c|}\hline x' \\ y' \\ \hline\end{array} \\ &= \begin{array}{|cc|}\hline a' & b' \\ c' & d' \\ \hline\end{array} \begin{array}{|cc|}\hline a & b \\ c & d \\ \hline\end{array} \begin{array}{|c|}\hline x \\ y \\ \hline\end{array} \\ &= \begin{array}{|cc|}\hline a'a+b'c & a'b+b'd \\ c'a+d'c & c'b+d'd \\ \hline\end{array} \begin{array}{|c|}\hline x \\ y \\ \hline\end{array}. \end{aligned} \tag{A.37}$$

Thus the product of two matrices

$$\begin{array}{|cc|}\hline a & b \\ c & d \\ \hline\end{array} \begin{array}{|cc|}\hline a' & b' \\ c' & d' \\ \hline\end{array} = \begin{array}{|cc|}\hline aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \\ \hline\end{array}, \tag{A.38}$$

and like vectors (or complex numbers), the sum of two matrices

$$\boxed{\begin{matrix} a & b \\ c & d \end{matrix}} + \lambda \boxed{\begin{matrix} a' & b' \\ c' & d' \end{matrix}} = \boxed{\begin{matrix} a + \lambda a' & b + \lambda b' \\ c + \lambda c' & d + \lambda d' \end{matrix}}, \tag{A.39}$$

and similarly for higher dimensions.

Appendix B

Coordinate Systems

Multiple coordinate systems are useful to solve problems of different symmetries, including rectangular, spherical, and cylindrical.

B.1 Curvilinear Coordinates

Consider a general curvilinear coordinate system $\{u_1, u_2, u_3\}$ whose axes are orthogonal at point. An infinitesimally small cube with edges parallel to the local curvilinear coordinate directions has edges of lengths $h_1 du_1$, $h_2 du_2$, and $h_2 du_2$, as in Fig. B.1.

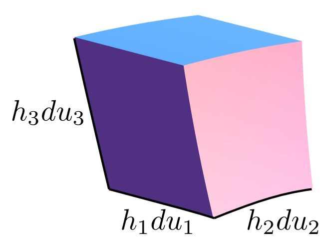


Figure B.1: Generic coordinate system $\{u_1, u_2, u_3\}$ and infinitesimal volume element of size $h_1 du_1$ by $h_2 du_2$ by $h_3 du_3$.

The square of the distance across opposite corners of the cube is

$$ds^2 = (h_1 du_1)^2 + (h_2 du_2)^2 + (h_3 du_3)^2 = h_1^2 du_1^2 + h_2^2 du_2^2 + h_3^2 du_3^2. \tag{B.1}$$

The volume of the cube is

$$dV = (h_1 du_1)(h_2 du_2)(h_3 du_3) = h_1 h_2 h_3 \, du_1 du_2 du_3. \tag{B.2}$$

A component of the gradient of a scalar field $S[\vec{r}]$ is the change of the scalar field along one edge of the infinitesimal cube divided by the length of that edge. Hence,

$$\vec{\nabla} S = \hat{u}_1 \frac{1}{h_1} \frac{\partial S}{\partial u_1} + \hat{u}_2 \frac{1}{h_2} \frac{\partial S}{\partial u_2} + \hat{u}_3 \frac{1}{h_3} \frac{\partial S}{\partial u_3}. \tag{B.3}$$

The divergence of a vector field $\vec{v}[\vec{r}]$ is the flux of the vector field through the faces of the infinitesimal cube divided by the volume of the cube. Hence,

$$\vec{\nabla} \cdot \vec{v} = \frac{1}{h_1 h_2 h_3} \left(\frac{\partial}{\partial u_1} (h_2 h_3 \, v_1) + \frac{\partial}{\partial u_2} (h_3 h_1 \, v_2) + \frac{\partial}{\partial u_3} (h_1 h_2 \, v_3) \right). \tag{B.4}$$

The Laplacian of a vector field is the divergence of the gradient, so

$$\nabla^2 S = \frac{1}{h_1 h_2 h_3} \left(\frac{\partial}{\partial u_1} \left(\frac{h_2 h_3}{h_1} \frac{\partial S}{\partial u_1} \right) + \frac{\partial}{\partial u_2} \left(\frac{h_3 h_1}{h_2} \frac{\partial S}{\partial u_2} \right) + \frac{\partial}{\partial u_3} \left(\frac{h_1 h_2}{h_3} \frac{\partial S}{\partial u_3} \right) \right). \tag{B.5}$$

A component of the curl of a vector field is the circulation of the vector field around a face of the the infinitesimal cube divided by the area of that face. Hence,

$$\begin{aligned} \vec{\nabla} \times \vec{v} = \hat{u}_1 & \frac{1}{h_2 h_3} \left(\frac{\partial}{\partial u_2} (h_3 v_3) - \frac{\partial}{\partial u_3} (h_2 v_2) \right) + \\ \hat{u}_2 & \frac{1}{h_3 h_1} \left(\frac{\partial}{\partial u_3} (h_1 v_1) - \frac{\partial}{\partial u_1} (h_3 v_3) \right) + \\ \hat{u}_3 & \frac{1}{h_1 h_2} \left(\frac{\partial}{\partial u_1} (h_2 v_2) - \frac{\partial}{\partial u_2} (h_1 v_1) \right). \end{aligned} \tag{B.6}$$

B.2 Polar Spherical Coordinates

Define spherical coordinates $\{u_1, u_2, u_3\} = \{r, \theta, \phi\}$ by

$$\begin{aligned} x &= r \sin\theta \cos\phi, \\ y &= r \sin\theta \sin\phi, \\ z &= r \cos\theta, \end{aligned} \tag{B.7}$$

where θ is the co-latitude and ϕ is the longitude, as in Fig. B.2. By inspection, the scale factors

$$\begin{aligned} h_1 &= 1, \\ h_2 &= r, \\ h_3 &= r \sin\theta. \end{aligned} \tag{B.8}$$

Hence, the diagonal square distance

$$ds^2 = dr^2 + (r\, d\theta)^2 + (r \sin\theta \, d\phi)^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2, \tag{B.9}$$

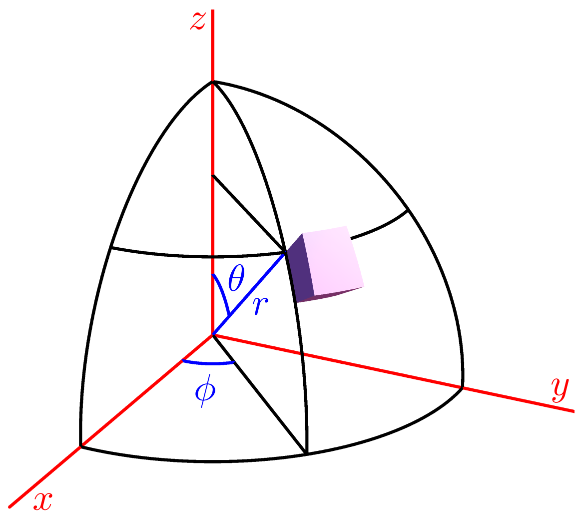


Figure B.2: Polar spherical coordinate system $\{r, \theta, \phi\}$ and infinitesimal volume element of size dr by $r\,d\theta$ by $r\sin\theta\,d\phi$.

and the elemental volume

$$dV = (dr)(r\,d\theta)(r\sin\theta\,d\phi) = r^2\sin\theta\,dr\,d\theta\,d\phi. \tag{B.10}$$

The spherical gradient

$$\vec{\nabla}S = \hat{r}\frac{\partial S}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial S}{\partial \theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial S}{\partial \phi}. \tag{B.11}$$

The spherical divergence

$$\vec{\nabla}\cdot\vec{v} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2\,v_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\,v_\theta) + \frac{1}{r\sin\theta}\frac{\partial v_\phi}{\partial\phi}. \tag{B.12}$$

The spherical Laplacian

$$\nabla^2 S = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial S}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial S}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial}{\partial\phi}\left(\frac{\partial S}{\partial\phi}\right). \tag{B.13}$$

The spherical curl

$$\begin{aligned}\vec{\nabla}\times\vec{v} = {} & \hat{r}\frac{1}{r\sin\theta}\left(\frac{\partial}{\partial\theta}(\sin\theta\,v_\phi) - \frac{\partial v_\theta}{\partial\phi}\right) + \\ & \hat{\theta}\frac{1}{r}\left(\frac{1}{\sin\theta}\frac{\partial v_r}{\partial\phi} - \frac{\partial}{\partial r}(r v_\phi)\right) + \\ & \hat{\phi}\frac{1}{r}\left(\frac{\partial}{\partial r}(r v_\theta) - \frac{\partial v_r}{\partial\theta}\right).\end{aligned} \tag{B.14}$$

In the $\theta = \pi/2$ equatorial plane, polar spherical coordinates become polar coordinates $\{r, \phi\}$.

B.3 Cylindrical Coordinates

Define cylindrical coordinates $\{u_1, u_2, u_3\} = \{s, \phi, z\}$ by

$$
\begin{aligned}
x &= s\cos\phi, \\
y &= s\sin\phi, \\
z &= z,
\end{aligned} \tag{B.15}
$$

where $s = r_\perp$ is the perpendicular distance form the axis and ϕ is the longitude, as in Fig. B.3. By inspection, the scale factors

$$
\begin{aligned}
h_1 &= 1, \\
h_2 &= s, \\
h_3 &= 1.
\end{aligned} \tag{B.16}
$$

Hence, the diagonal square distance

$$ds^2 = dr^2 + (s\,d\phi)^2 + dz^2 = dr^2 + s^2 d\phi^2 + dz^2 \tag{B.17}$$

and the elemental volume

$$dV = (ds)(s\,d\phi)(dz) = s\,ds d\phi\,dz. \tag{B.18}$$

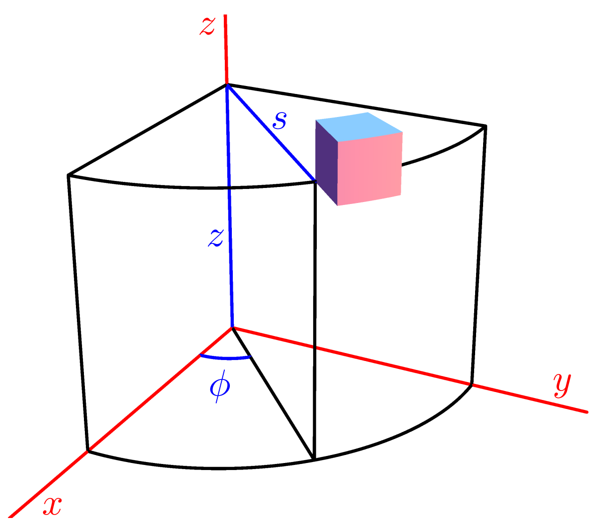


Figure B.3: Cylindrical coordinate system $\{s, \phi, z\}$ and infinitesimal volume element of size ds by $s\,d\phi$ by dz.

The cylindrical gradient

$$\vec{\nabla} S = \hat{s}\frac{\partial S}{\partial s} + \hat{\phi}\frac{1}{s}\frac{\partial S}{\partial \theta} + \hat{z}\frac{\partial S}{\partial z}. \tag{B.19}$$

The cylindrical divergence

$$\vec{\nabla} \cdot \vec{v} = \frac{1}{s}\frac{\partial}{\partial s}(s\, v_s) + \frac{1}{s}\frac{\partial v_\theta}{\partial \phi} + \frac{\partial v_z}{\partial \phi}. \tag{B.20}$$

The cylindrical Laplacian

$$\nabla^2 S = \frac{1}{s}\frac{\partial}{\partial s}\left(s\frac{\partial S}{\partial s}\right) + \frac{1}{s^2}\frac{\partial}{\partial \phi}\left(\frac{\partial S}{\partial \phi}\right) + \frac{\partial}{\partial z}\left(\frac{\partial S}{\partial z}\right). \tag{B.21}$$

The cylindrical curl

$$\begin{aligned} \vec{\nabla} \times \vec{v} = \hat{s}\left(\frac{1}{s}\frac{\partial v_z}{\partial \phi} - \frac{\partial v_\phi}{\partial z}\right) + \\ \hat{\phi}\left(\frac{\partial v_s}{\partial z} - \frac{\partial v_z}{\partial s}\right) + \\ \hat{z}\,\frac{1}{s}\left(\frac{\partial}{\partial s}(s v_\phi) - \frac{\partial v_s}{\partial \phi}\right). \end{aligned} \tag{B.22}$$

Appendix C

Notation

Table C.1 summarizes the symbols of this text. Some symbols are more universal than others.

Standard mathematics notation suffers from a serious ambiguity involving parentheses. In particular, parentheses can be used to denote multiplication, as in $a(b+c) = ab + ac$ and $f(g) = fg$, or they can be used to denote functions evaluated at arguments, as in $f(t)$ and $g(b+c)$. It can be a struggle to determine the intended meaning from context.

To avoid ambiguity, this text always uses round parentheses $(\bullet)$ to group for multiplication and square brackets $[\bullet]$ to list function arguments. Thus, $a(b) = ab$ denotes the product of two factors a and b, while $f[x]$ denotes a function f evaluated at an argument x. *Mathematica* employs the same convention.

Advanced physics texts often write integrals

$$\mathcal{I} = \int_a^b f[x]\,dx = \int_a^b dx\, f[x], \tag{C.1}$$

with the integration variable indicator dx near the integral sign $\int$, like summations

$$\mathcal{S} = \sum_{n=a}^{b} f_n, \tag{C.2}$$

with the summation variable n near the summation symbol $\sum$.

Table C.1: Symbols used in this text.

Quantity	Symbol	Alternates
vector	$\vec{v}$	$\mathbf{v}, \overline{v}, \underline{v}, \overrightarrow{v}, \lvert v\rangle$
unit vector	$\hat{v} = \vec{v}/v = \vec{v}/\lvert\vec{v}\rvert$	$\mathbf{u}, \overline{u}, \underline{u}, \vec{e}_v$
matrix symbol	M	$\mathbf{M}, \overline{\overline{M}}, \underline{\underline{M}}$
matrix	$\boxed{\begin{matrix} a & b \\ c & d \end{matrix}}$	$\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{pmatrix} a & b \\ c & d \end{pmatrix}$
functions	$f[x], \vec{\mathcal{A}}[\vec{r}, t]$	$f(x), \vec{\mathcal{A}}(\vec{r}, t)$
complex numbers	$z = x + \boldsymbol{i}y = \{x, y\}$	$x + iy$
derivatives	$\dot{x}, dx/dt, \partial f/\partial y, \partial_y f$	$x'(t), f_y$
(differential) operators	$\mathring{E} = +\boldsymbol{i}\hbar\partial_t,\ \mathring{\vec{p}} = -\boldsymbol{i}\hbar\vec{\nabla}$	$\hat{E}, E_{\mathrm{op}}$
Pauli (spin) matrices	$\sigma_x, \sigma_y, \sigma_z$	$\sigma_1, \sigma_2, \sigma_3$
position	$\vec{r} = \{x, y, z\}$	$\vec{x} = \{x_1, x_2, x_3\}$
fields	$\vec{\mathcal{E}}, \vec{\mathcal{B}}$	$\mathbf{E}, \mathbf{B}, \mathbf{D}, \mathbf{H}$
potentials	$\mathcal{V}, \vec{\mathcal{A}}$	$\varphi, \mathbf{A}$
sources	q, Q, i, I	e
densities	$\rho = dq/dV, \vec{J} = d\vec{I}/da$	$\mathbf{J}$
dipole moments	$\vec{\mu}_{\mathcal{E}}, \vec{\mu}_{\mathcal{B}}$	$\vec{p}, \vec{m}, \mathbf{p}, \mathbf{m}$
electron charge	$q_e < 0$	$-e < 0$

Bibliography

[1] Paul Adrien Maurice Dirac. The quantum theory of the electron. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 117(778):610–624, 1928.

[2] Albert Einstein. Zur Elektrodynamik bewegter Körper [On the Electrodynamics of Moving Bodies]. *Annalen der Physik*, 322(10):891–921, January 1905.

[3] Richard P. Feynman. Space-time approach to quantum electrodynamics. *Phys. Rev.*, 76:769–789, Sep 1949.

[4] Ben Yu-Kuang Hu. Introducing electromagnetic field momentum. *European Journal of Physics*, 33(4):873–881, May 2012.

[5] N. David Mermin. Relativity without light. *American Journal of Physics*, 52(2):119–124, 1984.

[6] Wolfgang Pauli. Zur quantenmechanik des magnetischen elektrons [On the quantum mechanics of the magnetic electron]. *Zeitschrift fur Physik*, 43(9-10):601–623, September 1927.

[7] Erwin Schrödinger. Quantisierung als eigenwertproblem [Quantization as an Eigenvalue Problem]. *Annalen der Physik*, 384(4):361–376, 1926.

[8] Julian Schwinger. On quantum-electrodynamics and the magnetic moment of the electron. *Phys. Rev.*, 73:416–417, Feb 1948.