Companion Notes to Peskin & Schroeder's QFT

焯焯侠*

January 7, 2026

Contents

*School of Physics, Nankai University. Major: Theoretical Physics. Email: 1687591283@qq.com

1 Introduction

These notes serve as a commentary on Peskin & Schroeder's textbook [1], supplementing the original text by providing the missing intermediate calculation steps. They are designed to be accessible to any graduate student with a foundation in classical mechanics and quantum mechanics.

The notes also include the author's personal perspectives. Given the author's background in mathematics, more rigorous mathematical proofs will be integrated into future versions. The exercise solutions provided in each section reference A Complete Solution to Problems in 'An Introduction to Quantum Field Theory' by Zhong-Zhi Xianyu (Harvard University).

The primary references for these notes are:

1. An Introduction to Quantum Field Theory by Peskin and Schroeder [1].
2. Modern Quantum Mechanics by Sakurai [2].
3. A Complete Solution to Problems in "An Introduction to Quantum Field Theory" by Zhong-Zhi Xianyu [3].
4. Modern Electrodynamics by Zangwill [4].

2 General Electromagnetic Fields

2.1 Maxwell's Equations

We begin by formulating Maxwell's equations in a vacuum. This fundamental set comprises four partial differential equations: Gauss's law for electricity, Gauss's law for magnetism, Faraday's law of induction, and the Ampère−Maxwell law. We first establish the fundamental quantities that act as sources for the electromagnetic field. For detailed derivations, please refer to Chapter 2 of Zangwill's Modern Electrodynamics [4, Chapter 2].

In the macroscopic limit, we treat charge not as discrete points but as a continuous volume charge density, defined as:

$$\rho(\mathbf{r}, t) = \frac{dq}{dV},$$

which represents the average electric charge per unit volume.

To describe the dynamic transport of this charge, we introduce the current density vector $\mathbf{j}$. For a charge distribution moving with a velocity field $\mathbf{v}$, this is defined as:

$$\mathbf{j} = \rho\mathbf{v},$$

representing the charge flux through a unit area per unit time.

Finally, the physical principle of charge conservation dictates that these quantities are intrinsically coupled: any temporal change in the amount of charge within a volume must be accounted for by the flow of current across its boundary. This conservation law is mathematically codified in the continuity equation:

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0 \tag{2.1}$$

By the expirmental results, we define equation:

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int d^3r' \rho\left(\mathbf{r}'\right) \frac{\mathbf{r} - \mathbf{r}'}{\left|\mathbf{r} - \mathbf{r}'\right|^3}. \tag{2.2}$$

to be the electric field for any choice of charge density $\rho(\mathbf{r})$. Given (2.2), we can compute the divergence of $\mathbf{E}$:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad \text{and} \quad \nabla \times \mathbf{E} = 0. \tag{2.3}$$

The first equation in (2.3) is **Gauss's law** in differential form. The second equation states is only valid for electrostatics, where charges are stationary.

The magnetic field $\mathbf{B}$ is defined by the Biot-Savart in 1820, and we define the magnetic field produced by any time-independent current density $\mathbf{j}(\mathbf{r})$ as:

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int d^3r' \frac{\mathbf{j}\left(\mathbf{r}'\right) \times \left(\mathbf{r} - \mathbf{r}'\right)}{\left|\mathbf{r} - \mathbf{r}'\right|^3}. \tag{2.4}$$

In 1851, William Thomson (later Lord Kelvin) used an equivalence between current loops and permanent magnets due to Ampère to show that the magnetic field produced by both types of sources satisfies

$$\nabla \cdot \mathbf{B} = 0 \quad \text{and} \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{j}. \tag{2.5}$$

We confirm that (2.4) is consistent with both equations in (2.5) as long as the current density satisfies the steady-current condition: $\nabla \cdot \mathbf{j} = 0$. The first equation in (2.5) states is called **Gauss's law for magnetism**, and the second equation is known as **Ampère's law**.

By Faraday's observation and Stokes' theorem, we yields the differential form of **Faraday's law**:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.6}$$

The displacement current $\mathbf{j}_D = \epsilon_0 \partial \mathbf{E}/\partial t$ is Maxwell's transcendent contribution to the theory of electromagnetism. We insert $\mathbf{j}_D$ into the second equation of (2.5) to obtain the modified **Ampère-Maxwell law**:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}. \tag{2.7}$$

where $c = 1/\sqrt{\mu_0 \epsilon_0}$ is the speed of light in vacuum.

Taken (2.3),(2.5),(2.6) and (2.7) together, we have the complete set of **Maxwell's equations in vacuum**:

$$\nabla \cdot \mathbf{B} = 0 \qquad \nabla \cdot \mathbf{E} = \rho/\epsilon_0 \tag{2.8}$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \qquad \nabla \times \mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{j} \tag{2.9}$$

2.2 Scalar Potential and Vector Potential

To simplify the description of the electric and magnetic fields, we introduce the scalar potential and the vector potential, respectively.

By using the Helmholtz theorem, we can provide the explicit formula of $\mathbf{E}(\mathbf{r})$ [4, Chapter 3]:

$$\mathbf{E}(\mathbf{r}) = -\nabla \frac{1}{4\pi\epsilon_0} \int d^3r' \frac{\rho\left(\mathbf{r}'\right)}{\left|\mathbf{r} - \mathbf{r}'\right|}. \tag{2.10}$$

The electric field integral(2.2) is difficult to evaluate for all but the simplest choices of $\rho(\mathbf{r})$. However, a glance back at (2.10) shows that we can define a function called the electrostatic scalar potential,

$$\varphi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int d^3r' \frac{\rho\left(\mathbf{r}'\right)}{\left|\mathbf{r} - \mathbf{r}'\right|},$$

and write the electric field in the form

$$\mathbf{E}(\mathbf{r}) = -\nabla\varphi(\mathbf{r}). \tag{2.11}$$

Similarly, we can get an explicit formula of $\mathbf{B}(\mathbf{r})$ by Helmholtz theorem. [4, Chapter 10]:

$$\mathbf{B}(\mathbf{r}) = \nabla \times \frac{\mu_0}{4\pi} \int d^3r' \frac{\mathbf{j}\left(\mathbf{r}'\right)}{|\mathbf{r}-\mathbf{r}'|} \tag{2.12}$$

The magnetic scalar potential formula $\mathbf{B} = -\nabla\psi$ is not valid at points in space where $\mathbf{j}(\mathbf{r}) \neq 0$. A more general approach to $\mathbf{B}(\mathbf{r})$ exploits the zero-divergence condition $\nabla \cdot \mathbf{B} = 0$ to infer that a vector potential $\mathbf{A}(\mathbf{r})$ exists such that

$$\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A}(\mathbf{r}). \tag{2.13}$$

We can easily appreciated this fact already from our application of the Helmholtz theorem to get(2.12), where

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int d^3r' \frac{\mathbf{j}\left(\mathbf{r}'\right)}{|\mathbf{r}-\mathbf{r}'|}.$$

Thus, in general, we can express the electrostatics and magnetostatics in terms of the scalar(2.11) and vector potentials(2.13).

2.3 Electromagnetic Potentials

The scalar potential $\varphi(\mathbf{r})$ and vector potential $\mathbf{A}(\mathbf{r})$ played prominent simplifying roles in electrostatics and magnetostatics. Their time-dependent counterparts do the same for time-varying fields. The starting point, as always, is the Maxwell equations in vacuum (2.8) and (2.9).

We consider the time-varying generalization of (2.13):

$$\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t). \tag{2.14}$$

Further progress comes from inserting (2.14) into Faraday's law on the left side of (2.8). This gives

$$0 = \nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right)$$

Recall now that $\nabla \times \nabla f = 0$ is an identity for any $f(\mathbf{r}, t)$. We can set $\varphi\left(\mathbf{r}, t\right)$ in such a way that:

$$\mathbf{E}(\mathbf{r}, t) = -\nabla\varphi(\mathbf{r}, t) - \frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t}. \tag{2.15}$$

Thus,the potentials reduce the number of functions to be determined from six (the scalar components of $\mathbf{E}$ and $\mathbf{B}$) to four (the scalar components of φ and the vector components of $\mathbf{A}$).

2.3.1 Gauge Invariance

Equations of motion for the potentials follow by substituting (2.15) and (2.14) into the inhomogeneous Maxwell equations, the equations on the right sides of (2.8) and (2.9) where the charge density and current density appear. Making use of $\nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A}$, we find

$$\nabla^2\varphi + \frac{\partial}{\partial t}(\nabla \cdot \mathbf{A}) = -\rho/\epsilon_0 \tag{2.16}$$

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla\left(\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial\varphi}{\partial t}\right) = -\mu_0\mathbf{j} \tag{2.17}$$

Like their static counterparts, $\varphi(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$ are not uniquely defined. This has no observable consequences because the non-uniqueness does not affect the electric and magnetic fields that enter the Coulomb-Lorentz law. We therefore let $\Lambda(\mathbf{r},t)$ be an arbitrary gauge function of space and time and define a new vector potential $\mathbf{A}'$ and a new scalar potential φ' using

$$\mathbf{A}' = \mathbf{A} + \nabla\Lambda, \tag{2.18}$$

and

$$\varphi' = \varphi - \frac{\partial\Lambda}{\partial t}. \tag{2.19}$$

We can verify that the fields $\mathbf{E}$ and $\mathbf{B}$, calculated using (2.14) and (2.15), remain invariant under (2.19) and (2.18).

Thus, we need to choose a gauge function $\Lambda(\mathbf{r},t)$ so that (2.16) and (2.17) become simple and easy to solve when written in the primed variables. Two common choices are the Coulomb gauge and the Lorenz gauge, which are defined by the condition

$$\nabla\cdot\mathbf{A} = 0 \qquad \text{(Coulomb gauge)} \tag{2.20}$$

$$\nabla\cdot\mathbf{A} + \frac{1}{c^2}\frac{\partial\varphi}{\partial t} = 0 \qquad \text{(Lorenz gauge)} \tag{2.21}$$

2.3.2 The Coulomb Gauge and the Lorenz Gauge

The **Coulomb gauge** choice (2.20) reduces (2.16) and (2.17) to

$$\nabla^2\varphi_C = -\rho/\epsilon_0$$

and

$$\nabla^2\mathbf{A}_C - \frac{1}{c^2}\frac{\partial^2\mathbf{A}_C}{\partial t^2} = -\mu_0\mathbf{j} + \frac{1}{c^2}\nabla\frac{\partial\varphi_C}{\partial t}.$$

The scalar potential obeys Poisson's equation. Therefore, if we specify that $\varphi_C(\mathbf{r},t)\to 0$ as $|\mathbf{r}|\to\infty$, we know from electrostatics that

$$\varphi_C(\mathbf{r},t) = \frac{1}{4\pi\epsilon_0}\int d^3r'\,\frac{\rho\,(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|} \tag{2.22}$$

This shows that at least a part of the electric field (2.15) is the familiar, instantaneous Coulomb electric field.

The **Lorenz gauge** choice (2.21) uncouples (2.16) and (2.17) to give

$$\nabla^2\varphi_L - \frac{1}{c^2}\frac{\partial^2\varphi_L}{\partial t^2} = -\rho/\epsilon_0 \tag{2.23}$$

and

$$\nabla^2\mathbf{A}_L - \frac{1}{c^2}\frac{\partial^2\mathbf{A}_L}{\partial t^2} = -\mu_0\mathbf{j}. \tag{2.24}$$

We obeserve that the charge density ρ determines φ_L in exactly the same way that the Cartesian components of the current density $\mathbf{j}$ determine the Cartesian components of $\mathbf{A}_L$. This characteristic makes the Lorenz gauge very popular for problems where the Coulomb potential (2.22) does not simplify the physics.

It is crucial to distinguish the Lorenz gauge named after the Danish physicist**Ludvig Lorenz** who formulated this gauge condition from the works of **H.A. Lorentz** (the Dutch physicist famous for the Lorentz force and transformations in special relativity), as well as the meteorologist **Edward Lorenz**, who pioneered chaos theory and the butterfly effect.

2.4 Conversation of Energy

We now turn to considering the work done by electromagnetic [4, Chapter 15]. The magnetic Lorentz force does no work , so all the work is done by the electric Coulomb force. Specifically, the rate at which $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{B}(\mathbf{r}, t)$ do mechanical work on a collection of particles with charge density $\rho(\mathbf{r}, t)$ and current density $\mathbf{j}(\mathbf{r}, t) = \rho(\mathbf{r}, t)\boldsymbol{v}(\mathbf{r}, t)$ confined to a volume V is

$$\frac{dW_{\text{mech}}}{dt} = \int_V d^3r(\rho\mathbf{E} + \mathbf{j} \times \mathbf{B}) \cdot \boldsymbol{v} = \int_V d^3r\mathbf{j} \cdot \mathbf{E}. \tag{2.25}$$

We try to rewrite this in the form of a conservation law. The first step eliminates the current density $\mathbf{j}$ on the far right side of (2.25) using the Ampère-Maxwell equation. This gives

$$\frac{dW_{\text{mech}}}{dt} = \int_V d^3r \left[\frac{1}{\mu_0}\nabla \times \mathbf{B} - \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}\right] \cdot \mathbf{E}$$

Next, using Faraday's law and specific identity, we have

$$\nabla \cdot (\mathbf{E} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{B}) = -\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} - \mathbf{E} \cdot (\nabla \times \mathbf{B}).$$

Combining the result of this substitution with the original equation (2.25) gives the desired conversation of energy law in a form known as Poynting's theorem:

$$\int_V d^3r \frac{\partial}{\partial t}\frac{1}{2}\epsilon_0 \left[\mathbf{E} \cdot \mathbf{E} + c^2\mathbf{B} \cdot \mathbf{B}\right] = -\int_V d^3r\mathbf{j} \cdot \mathbf{E} - \int_V d^3r\nabla \cdot \mathbf{S}.$$

where we have defined the Poynting vector

$$\mathbf{S} = \frac{1}{\mu_0}\mathbf{E} \times \mathbf{B}.$$

We therefore define the total electromagnetic energy as:

$$U_{\text{EM}} = \frac{1}{2}\epsilon_0 \int_V d^3r \left[\mathbf{E} \cdot \mathbf{E} + c^2\mathbf{B} \cdot \mathbf{B}\right]. \tag{2.26}$$

This definition is physically plausible because it reduces to the sum of the electrostatic total energy U_E [4, Section 3.6] and the magnetostatic total energy U_B [4, Section 12.6] in the static limit.

A spatially local statement of energy conservation follows from (2.25) if we use (2.26) to define an electromagnetic energy density,

$$u_{\text{EM}} = \frac{1}{2}\epsilon_0 \left(\mathbf{E} \cdot \mathbf{E} + c^2\mathbf{B} \cdot \mathbf{B}\right).$$

We will rewrite this in terms of the electromagnetic tensor $F^{\mu\nu}F_{\mu\nu}$ later.

3 Special Relativity

Special relativity is the theory of how different observers, moving at constant velocity with respect to one another, report their experience of the same physical event [4, Chapter 22].

Our discussion begins with the physical postulates of relativity, the Lorentz transformation, and some of the simpler consequences of the Lorentz transformation. We treat the kinematics and dynamics of point particles quite briefly and do not discuss spin at all. A central topic is the transformation laws for electromagnetic quantities like charge density, current density, the electromagnetic potentials, and the electromagnetic fields. Using these, we revisit the physics of moving point charges and plane electromagnetic waves. We then introduce the concept of the Lorentz tensor and derive manifestly covariant representations for the Maxwell equations and for the conservation laws of electrodynamics.

3.1 Galileo's and Einstein's Relativity

Before special relativity was formulated, the fundamental laws of physics were understood to obey Galileo' s principle of relativity.Galileo's relativity principle states that the laws of bodily motion are the same in all **inertial frames**, where we define an inertial frame as one in which a body not subject to any forces moves with constant velocity.

Consider, for example, the inertial frames K and K' shown in Figure (1).

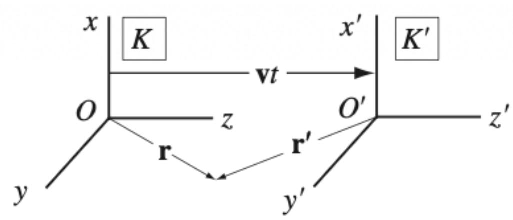

Figure 1: Two inertial frames K and K' moving at a constant relative velocity $\mathbf{v}$ along their common z-axis. At time $t = t' = 0$, the origins of the two frames coincide.

Special relativity appeared at a time when scientists were struggling to understand the absence of Galilean relative motion effects in light-propagation experiments. Einstein resolved the conceptual issues associated with the electrodynamics of moving bodies by rejecting the universal validity of Newton's laws and embracing the universal validity of Maxwell's laws. His famous and highly readable 1905 paper on the subject frames the solution using two postulates:

1. **The Principle of Relativity**: The laws of physics are the same in all inertial frames.
2. **The Constancy of the Speed of Light**: The speed of light in vacuum has the same value c in all inertial frames, independent of the motion of the source or observer.

Postulate 1 is a generalization of Galileo's relativity principle to include Maxwell's laws of electrodynamics.Postulate 2 explained the failure to detect relative motion between light and the aether by the simple expedient of making the aether superfluous.

We will show below that Einstein's two postulates (and the tacit assumptions that empty space is isotropic and spatially homogeneous) are sufficient to construct the entire edifice of special relativity. Part of the program is to discover the transformation laws that preserve the forms of the wave equation and the Maxwell equations, and to discover the dynamical law of motion that replaces Newton's laws. An equally important part of the program is to discover the physical consequences of the postulates. We begin with the most important of the physical consequences.

3.2 Lorentz Transformation

The different perception of time by different inertial observers leads us to treat space and time on an equal footing and to locate events in a venue called space-time. The most general transformation law between two inertial frames K and K' in space-time is

$$x' = x'(x, y, z, t) \quad y' = y'(x, y, z, t) \quad z' = z'(x, y, z, t) \quad t' = t'(x, y, z, t).$$

Because **space are homogeneou**, the infinitesimal displacement

$$dx' = \frac{\partial x'}{\partial x}dx + \frac{\partial x'}{\partial y}dy + \frac{\partial x'}{\partial z}dz + \frac{\partial x'}{\partial t}dt = 0. \tag{3.1}$$

cannot be an explicit function of (x, y, z, t), which means the partial derivatives in (3.1) are constants. We therefore let r_μ (with $\mu = 1, 2, 3, 4$) stand for x, y, z, ct and write the transformation law to this point in the form

$$r'_\mu = L_{\mu\nu} r_\nu + a_\mu. \tag{3.2}$$

3.2.1 Boosting the Standard Configuration

The unkown parameters of (3.2) can be droped to a handful from 20 if we assume that

i the coordinate axes in frame K are aligned with their counterparts in frame K';

ii the origins of the two frames coincide when $t = t' = 0$;

iii the velocity vector which "boosts" frame K to frame K' is $\mathbf{v} = v\hat{\mathbf{z}}$.

This returns us to the "standard configuration" of Figure(1), where the most general transformation law consistent with the rotational invariance of isotropic free space is

$$x' = Cx \quad y' = Cy \quad z' = Az + Bt \quad t' = Dz + Et. \tag{3.3}$$

We can derive that $B = -vA$, beacuse the standard configuration requires $z' = 0$ to coincide with $z = vt$. $C = 1$, because C cannot depend on the direction of motion of one frame with respect to the other.We therefore have to determine the three unknown parameters A, D, and E in (3.3), where (3.3) becomes

$$x' = x \quad y' = y \quad z' = A(z - vt) \quad t' = Dz + Et. \tag{3.4}$$

At this point, we need to employ the **method of undetermined coefficients** to determine the specific expressions for these three unknowns. We consider a point source of light that emits a spherical wave at $t = 0$ from the origin of K, which reaches the point (x, y, z) at time t and speed c such that

$$x^2 + y^2 + z^2 = c^2 t^2. \tag{3.5}$$

According to **postulate** 2, the constancy of the speed of light, we have the same formula at the frame K':

$$x'^2 + y'^2 + z'^2 = c^2 t'^2. \tag{3.6}$$

We substitute (3.4) into (3.6) and compare the result with (3.5) to obtain three equations for the three unknowns A, D, and E:

$$A^2 - c^2 D^2 = 1$$
$$E^2 - \frac{v^2}{c^2} A^2 = 1$$
$$vA^2 + c^2 DE = 0$$

We can solve these equations to find

$$A = E = \gamma \equiv \frac{1}{\sqrt{1 - v^2/c^2}} \quad \text{and} \quad D = -\frac{v\gamma}{c^2}.$$

where $A = E > 0$, beacuse we have to recover the Galilean transformation in the limit $v \ll c$.

We therefore have the Lorentz transformation between the coordinates of the two inertial frames K and K' in standard configuration:

$$x' = x \quad y' = y \quad z' = \gamma(z - \beta ct), \quad ct' = \gamma(ct - \beta z). \tag{3.7}$$

and

$$x = x' \quad y = y' \quad z = \gamma\left(z' + \beta ct'\right) \quad ct = \gamma\left(ct' + \beta z'\right). \tag{3.8}$$

where dimensionless velocity is $\beta = v/c$ and the Lorentz factor is $\gamma = (1 - \beta^2)^{-1/2}$.

We thus see that, by this extension, no material particle or object at rest in an inertial frame can be accelerated to the speed of light.

3.2.2 The Invariant Interval

We now turn to finding realativistic invariant quantity takes the same numerical value in every inertial frame. Let two events be denoted by coordinates (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) in frame K, with their separations defined as $\Delta x = x_1 - x_2$, and so forth.We have the same definition of $\Delta x' = x_1' - x_2'$,and so forth in frame K', which we have not chosen $\mathbf{v}$ between two frames. Thus, for,

$$\Delta z' = \gamma(\Delta z - \beta c\Delta t) \quad c\Delta t' = \gamma(c\Delta t - \beta \Delta z)$$

and

$$\Delta z = \gamma\left(\Delta z' + \beta c\Delta t'\right) \quad c\Delta t = \gamma\left(c\Delta t' + \beta \Delta z'\right),$$

we define square of the interval between these two events in frame K as:

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - (c\Delta t)^2 \tag{3.9}$$

and in frame K' as:

$$(\Delta s')^2 = (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2 - (c\Delta t')^2$$

It can be shown that $(\Delta s)^2 = (\Delta s')^2$,which means this quantity like c is also a Lorentz invariance.

The interval between two events play an inportant role in special relativity and quantum field theory, because it makes **causality**. The $(\Delta s)^2$ is called space-like separation if $(\Delta s)^2 > 0$, time-like separation if $(\Delta s)^2 < 0$, and light-like or null separation if $(\Delta s)^2 = 0$.

A pair of events with null separation can be connected by a signal traveling at the speed of light.

A pair of events with a time-like separation can always be made to occur at a single point in space $\Delta z' = 0$.

To achieve this, based on the Lorentz transformation formula $\Delta z' = \gamma(\Delta z - v\Delta t)$, the velocity of the inertial reference frame along the z-direction can be set to

$$v = \frac{\Delta z}{\Delta t} < c \qquad \text{(because of time-like separation)}.$$

This demonstrates that the two events are connected, thereby preserving the principle of causality.

For space-like separation, attempting to use a Lorentz transformation to place these two events at the same spatial point would require

$$v = \frac{\Delta z}{\Delta t} > c \qquad \text{(because of space-like separation).}$$

which is prohibited. Therefore, these events cannot be considered the same event; that is, no causal relationship exists between them.

The Light Cone and CausalityThe light cone acts as the fundamental geometric boundary of causality in spacetime, defined by the paths of light rays ($v = c$) emanating from a single event,which is shown as:

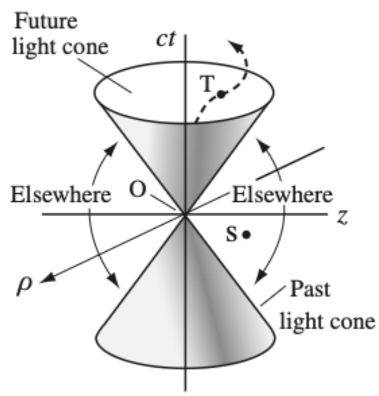

Figure 2: The light cone structure in spacetime. The vertical axis represents time, while the horizontal axes represent spatial dimensions. The cone's surface corresponds to light-like intervals, separating time-like and space-like regions.

It divides spacetime into two distinct regimes: the interior time-like region, where events are causally accessible via subluminal speeds ($v < c$), and the exterior space-like region.

3.2.3 Boosting a General Configuration

The proper time is an invariant measure of the motion of a particle along its trajectory in space-time. Considering the interval between two points on the world line which lie infinitesimally close to one another, we have:

$$(ds)^2 = (d\mathbf{r})^2 - (cdt)^2 = -(cdt)^2 \left[1 - \frac{u^2(t)}{c^2}\right],$$

where non-uniform velocity is $\mathbf{u}(t) = d\mathbf{r}/dt$.

We therefore can define an another invariant quantity named proper time $d\tau$ in frame K as:

$$d\tau = \sqrt{-\frac{(ds)^2}{c^2}} = \sqrt{1 - \frac{u^2(t)}{c^2}}dt = \frac{dt}{\gamma(u)} \tag{3.10}$$

We close with two points. First, the last equality in (3.10) generalizes the meaning of γ in (3.7) so **the argument u in $\gamma(u)$ can be a particle speed rather than merely the speed of a Lorentz boost from one inertial frame to another**. Second, the definition of $d\tau$ tells us that the "proper time" is the time measured by a clock in its own rest frame.

We have the most general Lorentz transformation between two inertial frames K and K' moving with a relative velocity $\mathbf{v}$:

$$\begin{aligned} \mathbf{r}'_\perp &= \mathbf{r}_\perp & \mathbf{r}_\perp &= \mathbf{r}'_\perp \\ \mathbf{r}'_\parallel &= \gamma\left(\mathbf{r}_\parallel - \boldsymbol{\beta} ct\right) & \mathbf{r}_\parallel &= \gamma\left(\mathbf{r}'_\parallel + \boldsymbol{\beta} ct'\right) \\ ct' &= \gamma\left(ct - \boldsymbol{\beta}\cdot\mathbf{r}_\parallel\right) & ct &= \gamma\left(ct' + \boldsymbol{\beta}\cdot\mathbf{r}'_\parallel\right). \end{aligned} \tag{3.11}$$

where we decompose $\mathbf{r}$ into its components $\mathbf{r}_\parallel$and $\mathbf{r}_\perp$ which lie parallel and perpendicular to $\boldsymbol{\beta} = \mathbf{v}/c$.

3.3 Four-Vectors

We denote a four-vector in Minkowski space by

$$\vec{a} = (a_1, a_2, a_3, a_4).$$

where we require the scalar of two four-vectors $\vec{a}$ and $\vec{b}$ be invariant to rotations, and Lorentz boosts from one inertial frame to another. That is,

$$\vec{a}\cdot\vec{b} = a_\mu b_\mu = a'_\mu b'_\mu = \vec{a'}\cdot\vec{b'}.$$

The prototype of a four-vector in special relativity is the space-time coordinate,

$$\vec{r} = (x, y, z, ict) = (\mathbf{r}, ict). \tag{3.12}$$

The fourth component of (3.12) is a pure imaginary number. This choice ensures that (3.9) produces the appropriate minus sign.

It was precisely the assumed invariance of these quantities which led us to the standard-configuration Lorentz transformation (3.7) and its inverse (3.8). The matrix representations for these transformations are

$$r'_\mu = \left[\frac{\partial r'_\mu}{\partial r_\nu}\right] r_\nu = L_{\mu\nu} r_\nu \quad \text{and} \quad r_\mu = \left[\frac{\partial r_\mu}{\partial r'_\nu}\right] r'_\nu = L^{-1}_{\mu\nu} r'_\nu,$$

where

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & 0 & -i\beta\gamma & \gamma \end{bmatrix} \quad \mathbf{L}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & -i\beta\gamma \\ 0 & 0 & i\beta\gamma & \gamma \end{bmatrix}. \tag{3.13}$$

where $\mathbf{L}^{\mathrm{T}} = \mathbf{L}^{-1}$, and $L_{\mu\lambda} L_{\nu\lambda} = \delta_{\mu\nu}$. The determinant of these transformation matrices is one:

$$|\mathbf{L}| = \left|\mathbf{L}^{-1}\right| = 1.$$

By definition, the components of an arbitrary four-vector $\vec{a}$ transform exactly like $\vec{r}$. Hence, a_1, a_2, and a_3 are often called the "space components" of $\vec{a}$, and a_4 is called the "time component" of $\vec{a}$. More precisely, $\vec{a}$ is a four-vector if

Referring back to (3.11), $a'_\mu = L_{\mu\nu} a_\nu$ is equivalent to

$$\begin{aligned} \mathbf{a}'_\perp &= \mathbf{a}_\perp & \mathbf{a}_\perp &= \mathbf{a}'_\perp \\ \mathbf{a}'_\parallel &= \gamma\left(\mathbf{a}_\parallel + i\boldsymbol{\beta} a_4\right) & \mathbf{a}_\parallel &= \gamma\left(\mathbf{a}'_\parallel - i\boldsymbol{\beta} a'_4\right) \\ a'_4 &= \gamma\left(a_4 - i\boldsymbol{\beta}\cdot\mathbf{a}_\parallel\right) & a_4 &= \gamma\left(a'_4 + i\boldsymbol{\beta}\cdot\mathbf{a}'_\parallel\right). \end{aligned} \tag{3.14}$$

3.3.1 The Four-Velocity

We now turn to introducing the concept of the **four-velocity**. The four-velocity $\vec{U}$ of a particle is another prototype four-vector.We define the four-velocity as:

$$\vec{U} = \frac{d\vec{r}}{d\tau} = \gamma(u)\frac{d}{dt}(\mathbf{r}, ict) = \gamma(u)\left(\frac{d\mathbf{r}}{dt}, ic\right) = \gamma(u)(\mathbf{u}, ic) \equiv (\mathbf{U}, U_4)\,. \tag{3.15}$$

It can be verified that its inner product is a Lorentz invariant scalar:

$$\vec{U}\cdot\vec{U} = \mathbf{U}\cdot\mathbf{U} + U_4^2 = \frac{\mathbf{u}\cdot\mathbf{u} - c^2}{1-u^2/c^2} = -c^2. \tag{3.16}$$

The four-velocity in this inertial reference frame K is related to the four-velocity in another inertial reference frame K' by a Lorentz transformation, the details of which will not be elaborated upon here.

3.3.2 The Four-momentum and Energy

The four-momentum $\vec{p}$ plays a central role in relativistic particle dynamics. Given the four-velocity in (3.15), we define $\vec{p}$ using a scalar $\mathcal{E}$ and a three-vector $\mathbf{p}$:

$$\vec{p} = m\vec{U} = m\,(\mathbf{U}, U_4) = (\mathbf{p}, i\mathcal{E}/c). \tag{3.17}$$

where the **rest mass** m is a Lorentz invariant scalar, and

$$\mathcal{E}/c = -imU_4 = \gamma(u)mc \quad \text{and} \quad \mathbf{p} = m\mathbf{U} = \gamma(u)m\mathbf{u}. \tag{3.18}$$

The meaning of $\mathcal{E}$ becomes clear when we Taylor expand $\gamma(u)mc$ in (3.18) for $u \ll c$ to get

$$\mathcal{E} = \frac{mc^2}{\sqrt{1-u^2/c^2}} = mc^2 + \frac{1}{2}mu^2 + \frac{3}{8}m\frac{u^4}{c^2} + \cdots\,. \tag{3.19}$$

The second term on the far right side of (3.19) is the familiar low-velocity kinetic energy.

The first term is a constant which may sensibly be called the **rest energy** and the total energy is $\mathcal{E}$.

The meaning of $\mathbf{p}$ in (3.18) emerges similarly from a Taylor expansion of $\gamma(u)m\mathbf{u}$ for $u \ll c$. The result,

$$\mathbf{p} = \frac{m\mathbf{u}}{\sqrt{1-u^2/c^2}} = m\mathbf{u}\left[1 + \frac{1}{2}\frac{u^2}{c^2} + \frac{3}{8}\frac{u^4}{c^4} + \cdots\right],$$

shows that $\mathbf{p}$ reduces to the ordinary Newtonian linear momentum, $m\mathbf{u}$, when the particle speed is very small compared to the speed of light.Using (3.16), the Lorentz invariant length of the energy-momentum four-vector (3.18) is

$$\vec{p}\cdot\vec{p} = m^2\vec{U}\cdot\vec{U} = -m^2c^2. \tag{3.20}$$

and

$$\vec{p}\cdot\vec{p} = \mathbf{p}\cdot\mathbf{p} - \frac{\mathcal{E}^2}{c^2} = p^2 - \frac{\mathcal{E}^2}{c^2}. \tag{3.21}$$

Combining (3.20) and (3.21), we obtain the important relativistic relation between energy, momentum, and rest mass:

$$\mathcal{E} = \sqrt{p^2c^2 + m^2c^4}. \tag{3.22}$$

Eliminating $\gamma(u)$ from the left and right sides of (3.18) relates the three-velocity to the threemomentum and the total energy:

$$\mathbf{u} = \frac{c^2\mathbf{p}}{\mathcal{E}}. \tag{3.23}$$

Experiment shows that (3.22) and (3.23) remain valid for **zero-mass particles**. The $m = 0$ limits of these formulae are

$$\mathcal{E} = cp \quad \text{and} \quad \mathbf{u} = c\frac{\mathbf{p}}{p}.$$

The energy-momentum four-vector (3.17) in this case is $\vec{p} = (\mathbf{p}, ip)$. This shows that there is no frame of reference where $\mathbf{p} = 0$ unless $\mathcal{E} = 0$, in which case the "particle" does not exist. Photons and neutrinos have **no rest frame**.

3.4 Electromagnetic Quantities

We now make some praperations for covariant electrodynamics by showing how various electromagnetic quantities transform between inertial frames [4, Section 22.6].

3.4.1 The Continuity Equation

Combining with (3.11) and (3.14),we are fully justified in defining the **four-vector operator**, which is also called the four-divergence, as

$$\vec{\nabla} = \left(\nabla, \frac{\partial}{\partial(ict)}\right). \tag{3.24}$$

Its components transform between inertial frames according to the same rules as those for an arbitrary four-vector $\vec{a}$. The inner product of this four-vector operator is

$$\vec{\nabla} \cdot \vec{\nabla} = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}. \tag{3.25}$$

Einstein's **postulate** 1 states that the laws of electromagnetism are valid in every inertial frame. We can rewirte the continuity equation (2.1) by the product of $\vec{j}$ and four-divergence as:

$$\vec{\nabla} \cdot \vec{j} = 0.$$

where $\vec{j}$ is defined as

$$\vec{j} = (\mathbf{j}, ic\rho). \tag{3.26}$$

We can verify that $\vec{j} = (\mathbf{j}, ic\rho)$ is a four-vector, as its inner product is:

$$\vec{j} \cdot \vec{j} = -c^2\rho_0^2$$

where ρ_0 is the charge density in the rest frame.

3.4.2 Lorenz Gauge Potentials

The gauge freedom enjoyed by the scalar potential $\varphi(\mathbf{r}, t)$ and the vector potential $\mathbf{A}(\mathbf{r}, t)$ imply that these quantities possess no intrinsic transformation properties when we change inertial frames. However, they acquire quite specific transformation properties if we choose a gauge constraint that is preserved by a Lorentz transformation. Using (2.1) as a model, the Lorenz gauge condition,

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \varphi}{\partial t} = 0, \tag{3.27}$$

has this property if we can define the four-vector

$$\vec{A} = (\mathbf{A}, i\varphi/c). \tag{3.28}$$

and use (3.24) to write (3.27) as an invariant four-divergence:

$$\vec{\nabla} \cdot \vec{A} = 0. \tag{3.29}$$

To confirm that (3.28) is indeed a four-vector, we recall that the electromagnetic potentials in the Lorenz gauge satisfy the inhomogeneous wave equations (2.23) and (2.24):

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\varphi = -\rho/\epsilon_0 \quad \text{and} \quad \left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\mathbf{A} = -\mu_0\mathbf{j}.$$

We therefore have the combined equation

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right](\mathbf{A}, i\varphi/c) = -\mu_0(\mathbf{j}, ic\rho). \tag{3.30}$$

Combining (3.30) with the Lorentz invariance of the wave operator (3.25) and the four-current character of (3.26) shows that the transformation properties on the left and right sides of (3.30) will not be the same unless (3.28) is indeed a four-vector.

3.5 Covariant Electrodynamics

We now turn to constructing a new respresentation of the Maxwell equations called manifestly covariant form which makes covariance more obvious.

3.5.1 Lorentz Tensors

A Lorentz tensor of rank 0 is what we have previously called a Lorentz scalar: a one-component quantity which is invariant to a change of inertial frame:

$$c' = c.$$

A Lorentz tensor of rank 1 is what we have previously called a four-vector: an object whose four components transform according to the Lorentz transformation matrix (3.13):

$$a'_\mu = L_{\mu v} a_v$$

A Lorentz tensor of rank 2 is an object whose 16 components transform according to the rule

$$s'_{\mu v} = L_{\mu\alpha} L_{v\beta} s_{\alpha\beta}.$$

Lorentz tensors of higher rank are defined similarly.

Consider a rank 1 tensor b_μ and a rank 2 tensor $W_{\mu\nu}$. We have the contraction theorem states that multiplying the two together and contracting (summing over) the common index μ produces an object d_ν which is a rank 1 tensor:

$$b_\mu W_{\mu\nu} = d_\nu$$

We thus can rewrite equations (2.1), (3.29) and (3.30) as :

$$\partial_\mu j_\mu = 0, \quad \partial_\mu A_\mu = 0, \quad \text{and} \quad \partial_\mu \partial_\mu A_v = j_v.$$

3.5.2 The Maxwell Equations

The path to writing the Maxwell equations in manifestly covariant form begins with the observation that the four-gradient $\partial_\mu = (\partial_1, \partial_2, \partial_3, \partial_4) = (\nabla, \partial/\partial(ict))$ and the four-potential $A_\mu = (A_1, A_2, A_3, A_4) = (\mathbf{A}, i\varphi/c)$ are sufficient to write out the two equations in(2.14) and (2.15):

$$\mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A}.$$

We therefore define one type of **electromagnetic field tensor** $F_{\mu\nu}$ as

$$F_{\mu\nu} = \partial_\mu A_v - \partial_v A_\mu.$$

Its matrix form is

$$\mathbf{F} = \begin{bmatrix} 0 & B_z & -B_y & -iE_x/c \\ -B_z & 0 & B_x & -iE_y/c \\ B_y & -B_x & 0 & -iE_z/c \\ iE_x/c & iE_y/c & iE_z/c & 0 \end{bmatrix}.$$

where the Spatial components $(i, j = 1, 2, 3)$ is corresponding to the magnetic field, e.g., $F_{12} = \partial_1 A_2 - \partial_2 A_1 = B_z$, and the time-space components ($i = 1, 2, 3$ and 4) is corresponding to the electric field, e.g., $F_{14} = \partial_1 A_4 - \partial_4 A_1 = -iE_x/c$.

The electromagnetic field-strength tensor $F_{\mu\nu}$ has only 6 (rather than 16) independent components because it is asymmetric ($F_{\mu\nu} = -F_{\nu\mu}$) and the diagonal elements ($\mu = \nu$) are zero.

We can derive another noteworthy is the Lorentz invariant scalar,

$$F_{\alpha\beta} F_{\alpha\beta} = 2\left(\mathbf{B} \cdot \mathbf{B} - \mathbf{E} \cdot \mathbf{E}/c^2\right)$$

3.5.3 Lagrangian of Free Electromagnetism Field

Consider a free electromagnetic field with no charged particles present. We define the free-field Lagrangian:

$$L_f = \frac{1}{2}\epsilon_0 \int d^3 r \left(\mathbf{E} \cdot \mathbf{E} - c^2 \mathbf{B} \cdot \mathbf{B}\right) = \frac{1}{4}\epsilon_0 \int d^3 r F_{\mu\nu} F_{\mu\nu} = \int d^3 r \mathcal{L}_f \tag{3.31}$$

where this Lagrangian density can be constructed via the inverse Legendre transformation of the Hamiltonian (2.26)

$$U_f = \frac{1}{2}\epsilon_0 \int_V d^3 r \left[\mathbf{E} \cdot \mathbf{E} + c^2 \mathbf{B} \cdot \mathbf{B}\right] = \int_V d^3 r \mathcal{H}_f.$$

For specific details, please refer to Professor Chen Tong's lecture notes [5].

The rightmost member of (3.31) defines the free-field Lagrangian density $\mathcal{L}_f$ and **a deductive approach to this quantity aims to establish the particular form shown in the middle member of** (3.31) **from first principles.** For guidance, we use the fact that Lagrange's equations derived from this Lagrangian should produce the Maxwell equations for a free electromagnetic field.

We will derive the Euler-Lagrange equations for the electromagnetic field using the metric formalism later.

3.6 Minkowski Metric and Some Notations

3.6.1 Units

We will work in "God-given" units, where

$$\hbar = c = \epsilon_0 = \mu_0 = 1.$$

In this system,

$$[\text{ length }] = [\text{ time }] = [\text{ energy }]^{-1} = [\text{ mass }]^{-1}$$

The mass (m) of a particle is therefore equal to its rest energy $\left(mc^2\right)$, and also to its inverse Compton wavelength $(mc/\hbar)$. For example,

$$m_{\text{electron}} = 9.109 \times 10^{-28}\ \text{g} = 0.511\text{MeV} = \left(3.862 \times 10^{-11}\ \text{cm}\right)^{-1}.$$

3.6.2 Minkowski Metric and Relativity Notations

Since we will be combining Special Relativity with Quantum Mechanics, a process that involves the frequent use of four-vectors and Lorentz invariant inner products, we will adopt an another notation system.

Our conventions for relativity follow Jackson (1975), Bjorken and Drell (1964, 1965), and nearly all recent field theory texts. We use the **Minkowski metric** tensor

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

with Greek indices running over $0, 1, 2, 3$ or t, x, y, z. Roman indicesi, j, etc.-denote only the three spatial components. Repeated indices are summed in all cases. Four-vectors, like ordinary numbers, are denoted by light italic type; three-vectors are denoted by boldface type; unit three-vectors are denoted by a light italic label with a hat over it.

We will redefine our notation following the conventions in Peskin & Schroeder [1]. The original 4-vector (3.12):

$$\vec{r} = (x, y, z, ict) = (\mathbf{r}, ict).$$

is redefined as:

$$x^\mu = \left(x^0, \mathbf{x}\right), \quad x_\mu = g_{\mu\nu} x^\nu = \left(x^0, -\mathbf{x}\right);$$

which represent the **contravariant** and **covariant** vectors, respectively. Consequently, the original inner product

$$\vec{r} \cdot \vec{r} = x^2 + y^2 + z^2 - (ct)^2, \tag{3.32}$$

becomes

$$x^2 = x^\mu x_\mu = g_{\mu\nu} x^\mu x^\nu = x^0 x^0 - \mathbf{x} \cdot \mathbf{x}. \tag{3.33}$$

where $x^0 = t$. Although (3.32)and Eq. (3.33) differ by a minus sign, the property of (3.33) as a Lorentz scalar remains unchanged. This notation **avoids the appearance of the imaginary unit** i.

For the original 4-momentum (3.17),

$$\vec{p} = m\vec{U} = m\left(\mathbf{U}, U_4\right) = (\mathbf{p}, i\mathcal{E}/c).$$

we redefine it as:

$$p^\mu = \left(p^0, \mathbf{p}\right), \quad p_\mu = g_{\mu\nu} p^\nu = \left(p^0, -\mathbf{p}\right);$$

where $p^0 = \mathcal{E} = E$. Thus, (3.22) becomes

$$p^0 = \sqrt{|\mathbf{p}|^2 + m^2}.$$

The scalar product of two vectors $\vec{p}$ and $\vec{x}$ is

$$p \cdot x = g_{\mu\nu} p^\mu x^\nu = p^0 x^0 - \mathbf{p} \cdot \mathbf{x}.$$

A massive particle has

$$p^2 = p^\mu p_\mu = E^2 - |\mathbf{p}|^2 = m^2.$$

For the original four-gradient operator (3.24):

$$\vec{\nabla} = \left(\nabla, \frac{\partial}{\partial(ict)}\right),$$

we redefine it as:

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial x^0}, \nabla\right), \quad \partial^\mu = g^{\mu\nu}\partial_\nu = \left(\frac{\partial}{\partial t}, -\nabla\right).$$

We define the totally antisymmetric tensor $\epsilon^{\mu\nu\rho\sigma}$ so that

$$\epsilon^{0123} = +1$$

where $\epsilon_{0123} = -1$ and $\epsilon^{1230} = -1$.

3.6.3 Quantum Mechanics

We will often work with the Schrödinger wavefunctions of single quantum mechanical particles. We represent the energy and momentum operators acting on such wavefunctions following the usual conventions:

$$E = i\frac{\partial}{\partial x^0}, \quad \mathbf{p} = -i\nabla.$$

These equations can be combined into

$$p^\mu = i\partial^\mu;$$

raising the index on ∂^μ conveniently accounts for the minus sign. The plane wave $e^{-ik\cdot x}$ has momentum k^μ, since

$$i\partial^\mu\left(e^{-ik\cdot x}\right) = k^\mu e^{-ik\cdot x}.$$

Thus far, we have established the foundational material for QFT, specifically Electromagnetism and Special Relativity. The subsequent sections will proceed with the material from Peskin & Schroeder.

References

[1] Michael E Peskin and Daniel V Schroeder. *An Introduction to Quantum Field Theory*. Addison-Wesley, Reading, MA, 1995.

[2] Jun John Sakurai and Jim Napolitano. *Modern Quantum Mechanics*. Cambridge University Press, Cambridge, 2nd edition, 2017.

[3] Zhong-Zhi Xianyu. A complete solution to problems in "an introduction to quantum field theory". Lecture notes, Harvard University, 2015. Unpublished solution manual.

[4] Andrew Zangwill. *Modern Electrodynamics*. Cambridge University Press, Cambridge, 2013.

[5] 陈童. 经典场论新讲*:* 现代导引. 2023.