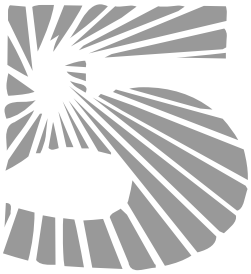


# Gauge symmetries and interactions

It is well known that the overall phase of a wave-function in quantum mechanics is not measurable. On the other hand, the so-called Aharonov–Bohm effect [☞ e.g., the texts [407, 471, 480, 472, 29, 324]] is based on the interference of two wave-functions and measures the *relative* phase, which proves that it is not possible to circumnavigate the complex nature of wave-functions. This then shows: (**1**) phases of wave-functions *are* physically relevant variables, and (**2**) any change in the *overall* (common) phase in a wave-function of the whole system must be a symmetry. This and the next chapters focus on this symmetry, and the corresponding conserved charge guaranteed to exist by Noether's theorem.

Moreover, this phase should be variable *locally*: in one way in one spacetime point, in another way in another spacetime point. It turns out that this seemingly simple (gauge) principle is actually the foundation of the contemporary understanding of all fundamental interactions [☞ [31] for the most complete review to date]. These five chapters (5–9) are dedicated to the application of this gauge idea, from technically simple examples towards more complex and realistic applications, and not following the history of its development but using the benefit of hindsight and the lessons of that history. For a flippant introduction of this idea, see also Refs. [33, 275, 269].

## 5.1 The non-relativistic $U(1)$ example

Start with the well-known non-relativistic quantum-mechanical description of a particle under the influence of a potential $V(\vec{r}\,)$, the wave-function of which is determined by the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\Psi(\vec{r},t) = \Big[-\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{r},t)\Big]\Psi(\vec{r},t), \tag{5.1}$$

and by the boundary conditions. In part, the boundary conditions follow from the shape of the potential and the chosen energy $E$ of the system, and are in part specified by choice. For example, in directions/regions where $V(\vec{r},t) > E$ as $r \to \infty$, we require $\lim_{r\to\infty}\Psi(\vec{r},t) = 0$; we also require that both $\int_V \mathrm{d}^3\vec{r}\ |\Psi(\vec{r},t)|^2$ and $\int_V \mathrm{d}^3\vec{r}\ \Psi^*(\vec{r},t)H\Psi(\vec{r},t)$ integrals are finite for every choice of the volume $V$. In the direction $\hat{\mathrm{e}}_k$ where $r \to \infty$ is not obstructed by a boundary condition, we may require that

$$\Psi(\vec{r},t) \sim \exp\Big\{+i\int \mathrm{d}(\hat{\mathrm{e}}_k\cdot\vec{r})\sqrt{2m[E-V(\vec{r},t)]}/\hbar\Big\}, \qquad r\to\infty. \tag{5.2}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

Such a particle may freely "escape to infinity" in the direction $\hat{\mathrm{e}}_k$, along which the kinetic energy remains positive, $\lim_{r\to\infty}\left(E - V(\vec{r}\,)\right) > 0$.

It is very well known that in this formalism the complex wave-function $\Psi(\vec{r},t)$ in its entirety does not correspond to any measurable quantity, but that $|\Psi(\vec{r},t)|^2$ *is* a physically measurable probability density of finding the particle in an infinitesimal volume $\mathrm{d}^3\vec{r}$ at the point $\vec{r}$ in space and $t$ in time. It follows that the *phase* of the complex function $\Psi(\vec{r},t)$ is not measurable,[1] so that no transformation

$$\Psi(\vec{r},t) \quad\to\quad e^{i\varphi}\,\Psi(\vec{r},t) \tag{5.3}$$

can have any physical (measurable) consequence. The transformation (5.3) is a symmetry of the Schrödinger equation (5.1) if and only if the phase $\varphi$ is a constant. In other words, the transformation (5.3) is a symmetry of the Schrödinger equation of the physical system described by the equation if and only if the identical transformation is applied to all points in space and each moment of time. Such a symmetry transformation is called *global*. Its existence is the necessary and sufficient condition for the application of Noether's theorem, and – therefore – for the existence of a corresponding conserved charge.

However, there should exist no physical obstacle for a transformation such as (5.3) to be performed with the phase $\varphi$ in one point of space and at one moment in time, and a completely different phase in another point of space and at another moment in time. Indeed, the choice of the wave-function phase should be a completely arbitrary choice of an unmeasurable degree of freedom, with no measurable consequence. In other words, the transformation (5.3) would have to be an exact symmetry of the physical system even if the phase $\varphi$ is an arbitrary function of $\mathrm{x} = (c\,t,\vec{r}\,)$. Such transformations and symmetries are called *local*.

**Digression 5.1** A rather formal justification for the transformation (5.3) to be a symmetry of the system is provided by noting that the formulation (5.3) in fact unnecessarily relies on the coordinate representation of the abstract state $|\Psi(t)\rangle$. Furthermore, it is known that only *pure* quantum states may be represented by a state vector $|\Psi(t)\rangle$, while a general state must be represented by a real, convex, normalized linear combination

$$\boldsymbol{\rho} = \sum_n r_n\,|n\rangle\langle n|, \quad \text{such that} \quad r_n \in \mathbb{R},\ \ 0 \leqslant r_n \leqslant 1,\ \ \sum_n r_n = 1. \tag{5.4a}$$

This is called the *state operator* [29], a.k.a. the *density matrix/operator* [471, 472, 360, for example]. Equivalently, $\boldsymbol{\rho}^\dagger = \boldsymbol{\rho}$, $\mathrm{Tr}[\boldsymbol{\rho}] = 1$ and $\langle u|\boldsymbol{\rho}|u\rangle \geqslant 0$ for every $|u\rangle$. A state operator (5.4a) represents a *pure* state if there exists a $|\Psi\rangle = \sum_n c_n|n\rangle$ such that $\boldsymbol{\rho} = |\Psi\rangle\langle\Psi|$; otherwise, $\boldsymbol{\rho}$ represents a *mixed* state.

The phase transformation (5.3) of the state vectors $|n\rangle$, written as $|n\rangle \to e^{i\boldsymbol{\varphi}}|n\rangle$, leaves the state operator $\boldsymbol{\rho}$ invariant:

$$\boldsymbol{\rho} \to \sum_n r_n\Big(e^{i\boldsymbol{\varphi}}|n\rangle\Big)\Big(\langle n|e^{-i\boldsymbol{\varphi}^\dagger}\Big) = \sum_n r_n|n\rangle\langle n| = \boldsymbol{\rho}, \qquad \text{if and only if } \ \big[\,e^{i\boldsymbol{\varphi}},\,|n\rangle\langle n|\,\big] = 0 \quad \text{and} \quad \boldsymbol{\varphi}^\dagger = \boldsymbol{\varphi}. \tag{5.4b}$$

[1] Here, we have in mind only the overall phase. In the transformation of the linear combination $\Psi = \Psi_1 + \Psi_2 \to e^{i\varphi_1}\Psi_1 + e^{i\varphi_2}\Psi_2$, the phase $(\varphi_1+\varphi_2)$ is the unmeasurable overall phase, while the *relative* phase $(\varphi_1-\varphi_2)$ is measurable by means of *interference*. This overwhelmingly reminds us of the fact that the absolute values of coordinates (and the phase is indeed a kind of coordinate) are not measurable quantities, while coordinate *differences* – i.e., distances – are.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

In turn, the information about the change of this (or any other) choice cannot be transported instantly, and there will have to exist some physical mechanism for transporting this information from point to point in space and time.

It is not hard to verify that the transformation (5.3) with $\varphi = \varphi(\vec{r},t)$ is *not* a symmetry of the Schrödinger equation (5.1):

$$i\hbar\frac{\partial}{\partial t}\Psi = \Big[-\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{r},t)\Big]\Psi, \tag{5.5}$$

transformation (5.3), with $\varphi = \varphi(\vec{r},t)$

$$i\hbar\frac{\partial}{\partial t}\Big(e^{i\varphi}\Psi\Big) = \Big[-\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{r},t)\Big]\Big(e^{i\varphi}\Psi\Big),$$

$$i\hbar\, e^{i\varphi}\Big(i\frac{\partial\varphi}{\partial t}\Big)\Psi + i\hbar\, e^{i\varphi}\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\vec{\nabla}\cdot\Big(e^{i\varphi}(i\vec{\nabla}\varphi)\Psi + e^{i\varphi}\vec{\nabla}\Psi\Big) + V(\vec{r},t)e^{i\varphi}\Psi,$$

$$\begin{aligned} i\hbar\, e^{i\varphi}\Big(i\frac{\partial\varphi}{\partial t}\Big)\Psi + i\hbar\, e^{i\varphi}\frac{\partial\Psi}{\partial t} \\ = -\frac{\hbar^2}{2m}\Big(e^{i\varphi}(i\vec{\nabla}\varphi)^2\Psi + e^{i\varphi}(i\vec{\nabla}^2\varphi)\Psi + 2e^{i\varphi}(i\vec{\nabla}\varphi)\cdot(\vec{\nabla}\Psi) + e^{i\varphi}\vec{\nabla}^2\Psi\Big) + V(\vec{r},t)e^{i\varphi}\Psi, \end{aligned}$$

so, using the original equation (5.1) and upon dividing by $\Psi(\vec{r},t)$, we obtain

$$\frac{\partial\varphi}{\partial t} = \frac{\hbar}{2m}\Big(i(\vec{\nabla}^2\varphi) + 2i(\vec{\nabla}\varphi)\cdot(\vec{\nabla}\ln(\Psi)) - (\vec{\nabla}\varphi)^2\Big). \tag{5.6}$$

This result is *absolutely unacceptable*! Not only did the (*unmeasurable!*) phase $\varphi(\vec{r},t)$ turn out not to be an arbitrarily selectable function of space and time, but it would have to satisfy a differential equation (5.6) that depends on the particular state of the system represented by the wave-function $\Psi(\vec{r},t)$!

The resolution of this seeming paradox can only lie in changing the Schrödinger equation, but in a way that does not ruin any of the many confirmed results obtained from this equation. Evidently, this is a very demanding request.

Following the computation (5.5)–(5.6) closely, one notices that the ultimate – and *absolutely unacceptable* – result stems from the fact that *derivatives* of the "new" wave-function $e^{i\varphi}\Psi(\vec{r},t)$ differ from the $e^{i\varphi}$-multiples of the derivative of the "old" wave-function $\Psi(\vec{r},t)$. With this hint, introduce a new kind of derivative:

$$\frac{\partial}{\partial t} \to D_t := \frac{\partial}{\partial t} + X, \qquad\qquad \vec{\nabla} \to \vec{D} := \vec{\nabla} + \vec{Y}, \tag{5.7a}$$

where the quantities $X$ and $\vec{Y}$ will be determined so that these newfangled $D$-derivatives satisfy the relations

$$D_t'\Psi' = D_t'(e^{i\varphi}\Psi) = e^{i\varphi}(D_t\Psi), \qquad \vec{D}'\Psi' = \vec{D}'(e^{i\varphi}\Psi) = e^{i\varphi}(\vec{D}\Psi). \tag{5.7b}$$

By writing $\Psi = e^{-i\varphi}\Psi'$, these requirements show that

$$(D_t'\cdots) = e^{i\varphi}(D_t e^{-i\varphi}\cdots), \qquad (\vec{D}'\cdots) = e^{i\varphi}(\vec{D}e^{-i\varphi}\cdots), \tag{5.7c}$$

where $D_t', \vec{D}'$ denotes these new derivatives *after* the $\Psi \to e^{i\varphi}\Psi$ transformation. In turn, with these newfangled derivatives, the Schrödinger equation becomes

$$i\hbar D_t\Psi = \Big[-\frac{\hbar^2}{2m}\vec{D}^2 + V(\vec{r})\Big]\Psi, \qquad \text{or} \qquad \Big[i\hbar D_t + \frac{\hbar^2}{2m}\vec{D}^2 - V(\vec{r})\Big]\Psi = 0, \tag{5.8}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

and changes under the transformation (5.3) as

$$0 = \Big[i\hbar D'_t + \frac{\hbar^2}{2m}\vec{D}'\cdot\vec{D}' - V(\vec{r}\,)\Big](e^{i\varphi}\Psi) = \Big[i\hbar e^{i\varphi}D_t + \frac{\hbar^2}{2m}\vec{D}'\cdot e^{i\varphi}\vec{D} - e^{i\varphi}V(\vec{r}\,)\Big]\Psi$$
$$= e^{i\varphi}\Big[i\hbar D_t + \frac{\hbar^2}{2m}\vec{D}\cdot\vec{D} - V(\vec{r}\,)\Big]\Psi, \tag{5.9}$$

which is satisfied precisely when equation (5.8) is. Thus, with these newfangled derivatives $D_t$ and $\vec{D}$, which themselves change via the transformation (5.3), the new Schrödinger equation (5.8) remains invariant.

***Comment 5.1*** *It is not at all unreasonable that the procedure for computing a rate of change (the derivative operator) needed adjustment. Recall that the total derivative* $\frac{d}{dt}f\big(t,g(t)\big) = \big[\frac{\partial}{\partial t} + \frac{\partial g}{\partial t}\frac{\partial}{\partial g}\big]f$ *may be viewed as the partial derivative* $\frac{\partial}{\partial t}$ *corrected for the fact that the function* $f$ *also depends on* $t$ *implicitly, via its dependence on* $g(t)$*. By the same token, complex wave-functions also depend on the spacetime coordinates implicitly, via their dependence on the choice of a spacetime variable phase.*

It remains to examine the nature of these newfangled derivatives (5.7), as well as the differences between the new Schrödinger equation (5.8) and the old one (5.1). The newfangled derivatives satisfy (5.7c)

$$\Big[\Big(\frac{\partial}{\partial t} + X'\Big)\cdots\Big] = e^{i\varphi}\Big[\Big(\frac{\partial}{\partial t} + X\Big)e^{-i\varphi}\cdots\Big],$$
$$\big[(\vec{\nabla} + \vec{Y}')\cdots\big] = e^{i\varphi}\big[(\vec{\nabla} + \vec{Y})e^{-i\varphi}\cdots\big]; \tag{5.10}$$

which yields

$$X' = X - i\frac{\partial\varphi}{\partial t} \qquad \text{and} \qquad \vec{Y}' = \vec{Y} - i(\vec{\nabla}\varphi). \tag{5.11}$$

The relations (5.11) ought to be familiar to all Students who have successfully completed a standard electrodynamics course! With the definitions

$$\Phi := \frac{\hbar}{iq}X, \qquad \vec{A} := \frac{i\hbar}{q}\vec{Y}, \qquad \lambda := \frac{\hbar}{q}\varphi, \tag{5.12}$$

the definitions (5.7a) become

$$D_t := \frac{\partial}{\partial t} + i\frac{q}{\hbar}\Phi, \qquad \vec{D} := \vec{\nabla} - i\frac{q}{\hbar}\vec{A}, \tag{5.13}$$

and are called the *covariant derivatives*. Combining, we have

$$\Phi \to \Phi' = \Phi - \frac{\partial\lambda}{\partial t}, \qquad \vec{A} \to \vec{A}' = \vec{A} + (\vec{\nabla}\lambda), \tag{5.14a}$$
$$\Psi(\vec{r},t) \to \Psi'(\vec{r},t) = e^{iq\lambda(\vec{r},t)/\hbar}\,\Psi(\vec{r},t). \tag{5.14b}$$

The first two relations are the standard gauge transformations of the vector and the scalar electromagnetic potentials. The third relation is the corresponding gauge transformation of the wave-function $\Psi(\vec{r},t)$ of a particle with the electric charge $q$, which is evidently a translation of the phase of this function.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

***Comment 5.2*** *The action of the gauge transformation (5.14b) implies that the complex conjugate wave-function $\Psi^*(\vec{r},t)$ represents a particle with the charge that is opposite to the particle represented by $\Psi(\vec{r},t)$: $q(\Psi^*) = -q(\Psi)$.*

The first two transformation equations (5.14a) clearly imply that the effect of $\lambda(\vec{r},t)$ is indistinguishable from that of $\lambda(\vec{r},t) + \lambda_0$, where $\lambda_0 = const.$, and the single-valuedness of $\Psi(\vec{r},t)$ then implies that $\lambda_0$ must be an integral multiple of $2\pi\cdot\frac{\hbar}{q_0}$, where $q_0$ then must be a minimal, unit electric charge. That is, the transformation function $\lambda(\vec{r},t)$ takes $2\pi\cdot\frac{\hbar}{q_0}$-periodic distinct values, i.e., along a circle of radius $\frac{\hbar}{q_0}$. In turn, the exponential $U_\lambda := e^{iq\lambda(\vec{r},t)/\hbar}$ is unitary: $(U_\lambda)^\dagger = (U_\lambda)^{-1}$, and such $\lambda$-parametrized exponentials form the *gauge group*, called $U(1)$ [☞ Appendix A, and especially A.2].

More precisely, note that the transformation function, $\lambda = \lambda(\vec{r},t)$, remains an unrestricted, arbitrary function of space and time[2] – true to the original insight and definition as discussed above. The combined transformation (5.14) is then the true and complete *local* symmetry: a *continuous family of $U(1)$ gauge groups of symmetries*, one independent $U(1)$ symmetry in every point of space and time!

Owing to the identity $\vec{\nabla}\times(\vec{\nabla} f) \equiv 0$ valid for any scalar function $f$, it follows that $(\vec{\nabla}\times\vec{A})$ is invariant with respect to the transformations (5.14). Similarly, since the transformation of $\vec{\nabla}\Phi$ is precisely opposite of the transformation of $\frac{\partial}{\partial t}\vec{A}$, the sum $(\vec{\nabla}\Phi + \frac{\partial}{\partial t}\vec{A})$ is also invariant. These expressions are, of course, familiar:

$$\vec{B} := \vec{\nabla}\times\vec{A} \qquad \text{and} \qquad \vec{E} := -\left(\vec{\nabla}\Phi + \frac{\partial\vec{A}}{\partial t}\right) \tag{5.15}$$

are the magnetic and the electric fields, expressed in terms of the electromagnetic potentials. The ability to define gauge-*invariant* fields $\vec{B}$ and $\vec{E}$ will be shown to be an exceptional consequence of the abelian (commutative) nature of the $U(1)$ gauge transformation (5.14).

---

**Digression 5.2** The term "gauge transformation" for the relations (5.14) is a historical atavism: It is a derivative of the literally translated German original coinage by Hermann Weyl, *Eichinvarianz*, by which he denoted the invariance with respect to transformations (5.14) [564]. Weyl noticed that Einstein's general theory of relativity is invariant with respect to complex rescalings. His original idea that the imaginary part of the rescaling function $\varphi(\vec{r},t)$ in the transformation (5.14b) may unite gravity with electromagnetism turned out unphysical. Such a rescaling symmetry would permit fixing a length unit in Nature, for which Weyl used the German verb *eichen*, meaning to gauge, to calibrate. The word *gauge* and its derivatives that are used in the English literature, *jauge* in French, βαθμίδας in Greek, *mérték* in Hungarian, (simply imported) *gauge* in Italian, *калибро́вочная* in Russian, *de gauge* in Spanish, etc. are all literal translations of the German verb *eichen*.

Soon, Vladimir A. Fok (first, according to Professor Okun [394], in 1926), Hermann Weyl, Erwin Schrödinger and Fritz London noticed that quantum mechanics, as governed by the Schrödinger equation, has a symmetry with respect to the combined transformations (5.14) using a real function $\varphi(\vec{r},t)$.[3] This was derived here as a

---

[2] Well, yes: $\lambda(\vec{r},t)$ clearly must be differentiable, at least once with respect to both $t$ and $\vec{r}$ for the equations (5.14a) to be well defined; see however also Section 5.2.3.

[3] Woit recounts [577, pp. 61–62] that Schrödinger hinted at this in a 1922 paper, but was chidingly reminded of this neglected "tidbit" in December of 1926 by the young London; see also the account in Ref. [119].

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

transformation stemming from the innate property of wave-functions that their phases are not measurable.

*Fundamental physics* is indubitably quantum. Equations (5.1)–(5.14) and their logic then indicate the fundamental nature of this principle to be that of a ***local symmetry***, emphasizing that a spacetime variable parameter $\lambda(\vec{r},t)$ in the transformations (5.14) parametrizes a spacetime continuum of $U(1)$ symmetries. ***Local symmetry*** is then used as a conceptually correct alternative for the historically well-entrenched term *gauge symmetry* or the descriptive but rarely used modifier *phase symmetry*.

***Comment 5.3*** *Note that the transformation (5.14) may be understood as a spacetime-dependent translation of sorts in the (abstract, target) space of values of the functions defined over the spacetime; a translation of the electromagnetic potentials and a* ***phase-shift*** *of the wave-function:*

$$\textit{Eq. (5.14)} \;\Rightarrow\; \mathrm{Arg}\left[\Psi(\vec{r},t)\right] \to \mathrm{Arg}\left[\Psi(\vec{r},t)\right] + \varphi(\vec{r},t), \quad \mathrm{Arg}[z] := \tfrac{1}{2i}\ln\left(\tfrac{z}{z^*}\right). \tag{5.16}$$

*The electromagnetic potentials and the phase of the wave-function are all physically* ***unmeasurable*** *variables, the existence of which is however necessary for the consistency of the model. Lorentz symmetry requires the gauge potentials to be 4-vectors, although only two polarizations (components) have a physical meaning; the complex-analytic structure of the Schrödinger and Dirac equations requires the use of complex wave-functions, although the (overall) phase is not physically measurable.*

With the definitions (5.12), the Schrödinger equation (5.8) becomes

$$i\hbar\Big[\frac{\partial}{\partial t} + \frac{iQ}{\hbar}\Phi\Big]\Psi = \Big[-\frac{\hbar^2}{2m}\Big(\vec{\nabla} + \frac{Q}{i\hbar}\vec{A}\Big)^2 + V(\vec{r},t)\Big]\Psi. \tag{5.17}$$

That is,

$$i\hbar\frac{\partial}{\partial t}\Psi(\vec{r},t) = H_{EM}\,\Psi(\vec{r},t), \tag{5.18}$$

where

$$H_{EM} = \frac{1}{2m}\Big(\frac{\hbar}{i}\vec{\nabla} - Q\,\vec{A}(\vec{r},t)\Big)^2 + \Big[V(\vec{r},t) + Q\,\Phi(\vec{r},t)\Big] \tag{5.19}$$

is the Hamiltonian for a particle of mass $m$ and electric charge $q := Q(\Psi)$. The dynamics of this particle is affected by the interaction with the potential $V(\vec{r},t)$, as well as the electromagnetic potentials $\vec{A}(\vec{r},t)$ and $\Phi(\vec{r},t)$.

***Conclusion 5.1*** *The transformation (5.14) with (5.16) is the fundamental assertion that we are at liberty to* ***arbitrarily*** *change the quantities that were introduced in the (mathematical) model of the physical system for its consistency, but which* ***on principle*** *represent no physically measurable quantity.*

It is worth noticing that the quantum description of the interaction of a charged particle with the electromagnetic field is inherently described in terms of the electromagnetic potentials $\vec{A}, \Phi$ and not in terms of the electric and magnetic field, $\vec{E}, \vec{B}$. Moreover, the Hamiltonian (5.19) cannot, in the general case, be expressed locally (without integration) as an interaction of a charged particle with the $\vec{E}$- and $\vec{B}$-fields.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

The following facts will be shown to be consequences of the abelian (commutative) nature of the $U(1)$ symmetry group: (**1**) the Maxwell equations[4] (5.72) as well as the corresponding Lagrangian and Hamiltonian for the electromagnetic field *can* be expressed exclusively in terms of the electric and the magnetic field, and (**2**) electromagnetic potentials *can* be fully eliminated from the equations of motion, the Lagrangian and the Hamiltonian except if there is matter that interacts with the electromagnetic field.

Indeed, the transformations (5.14) are parametrized by one function, $\lambda(\vec{r},t)$, which defines the local unitary *operator*

$$\varphi(\vec{r},t) \quad \mapsto \quad U_\varphi := \exp\left\{i\varphi(\vec{r},t)\, Q\right\} \tag{5.20}$$

as in equation (A.37), where the operator $Q$ may be regarded:

1. from the mathematical vantage point, as the generator of the $U(1)$ symmetry,
2. from the physical vantage point, the electric charge operator. The electric charge of a particle is then the eigenvalue and the wave-function of the particle the eigenfunction of the operator $Q$.

At every point of spacetime $\mathrm{x} = (ct,\vec{r}\,)$ separately, the (continuously many) operators $U_\varphi$ defined by equation (5.20) form an abelian (commutative) group, denoted $U(1)$. Since the function in the exponent manifestly satisfies $\varphi \simeq \varphi + 2\pi$, this group is sometimes identified with the circle, $S^1$. To repeat: Since $\varphi = \varphi(\vec{r},t)$ gives an independent "angle"-transformation at every point in space and time, we have a 4-dimensional *continuum* of $U(1)$ symmetry groups.

> ***Comment 5.4*** *The full space of "coordinates" in electrodynamics is therefore of the form* $(\textit{spacetime} \times S^1)$ *– a 5-dimensional topological space, equipped with a particular geometry; this was clear as early as in 1914 to Gunnar Nordstrøm [☞ Digression 11.5 on p. 414].*

### 5.1.1 Exercises for Section 5.1

✎ ***5.1.1*** *Fill in the details of the computation (5.7)–(5.14).*

✎ ***5.1.2*** *From the definitions (5.15), derive Gauss's law for the magnetic field and Faraday's law of induction. (This proves that the equations (5.72b) are* ***consequences*** *of Maxwell's definitions (5.15).)*

✎ ***5.1.3*** *Show that the gauge-invariant scalar functions of $\epsilon_0$, $\mu_0$, $\vec{E}$ and $\vec{B}$ with the dimensions of (volume) energy density and which are analytic functions of the components of the vectors $\vec{E}$ and $\vec{B}$ must be of the form*

$$c_1\big(\epsilon_0\,\vec{E}^2\big) + c_2\big(\tfrac{1}{\mu_0}\,\vec{B}^2\big) + c_3\big(\sqrt{\tfrac{\epsilon_0}{\mu_0}}\,\vec{E}\cdot\vec{B}\big). \tag{5.21}$$

*The results in Table C.4 on p. 527, should be useful.*

✎ ***5.1.4*** *Determine the constants $c_1, c_2, c_3, c_4, c_5$ so that*

$$\int dt\, d^3\vec{r}\,\Big\{\, c_1\big(\epsilon_0\,\vec{E}^2\big) + c_2\big(\tfrac{1}{\mu_0}\,\vec{B}^2\big) + c_3\big(\sqrt{\tfrac{\epsilon_0}{\mu_0}}\,\vec{E}\cdot\vec{B}\big) + c_4\,\rho\,\Phi + c_5\,\vec{j}\cdot\vec{A}\,\Big\} \tag{5.22}$$

[4] James Clerk Maxwell described electrodynamics, originally in 1873, as a system of equations which would today be written as $\vec{E} := -\vec{\nabla}\Phi - \frac{\partial\vec{A}}{\partial t}$ and $\vec{B} := \vec{\nabla}\times\vec{A}$, and then $\vec{\nabla}\cdot(\epsilon_0\vec{E}) = \rho$ and $\vec{\nabla}\times(\vec{B}/\mu_0) - \frac{\partial(\epsilon_0\vec{E})}{\partial t} = \vec{j}$. By the Maxwell equations (5.72) today, one understands the *consequences* of the first two of these equations together with the latter two, expressed exclusively in terms of the electric and the magnetic field, where the electromagnetic potentials, $\vec{A}$ and $\Phi$, are eliminated, and where there are neither (monopole) magnetic charges nor magnetic currents: $\rho_m = 0 = \vec{j}_m$.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

*is the Hamilton action the variation of which by $\Phi$ and $\vec{A}$, using the relations (5.15), produces Gauss and Ampère's law (5.72a).*

## 5.2 Electrodynamics with leptons

By quantum electrodynamics one understands the relativistic theory that describes the interaction of photons and electrically charged particles. Unlike leptons, quarks and hadrons also interact via the much stronger strong nuclear interaction, so the analysis of their interactions is considerably more complicated. This section is limited to the electromagnetic interactions of leptons, and the next one turns to the electromagnetic interactions of the hadrons.

It follows from the relations (A.43d)–(A.43f) that the components of the radius-vector, and then also any other vector quantity, span a spin-1 representation of the rotation group. One thus says that the photon (represented by the vector potential,[5] $\vec{A}$) has spin 1. On the other hand, it is well known that electrically charged particles such as the electron and the quarks, which make up all tangible matter, have spin $\frac{1}{2}$.

Thus, we must first establish the relativistic generalization of the Schrödinger equation for particles of spin $\frac{1}{2}$ and 1, as well as the argument from the previous section, which specifies the interaction between them.

### 5.2.1 Relativistic spinors and the Dirac equation

The Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi(\vec{r},t) = H\Psi(\vec{r},t) \qquad \Leftrightarrow \qquad \Psi(\vec{r},t) = e^{-i\hbar^{-1}\int_{t_0}^{t} \mathrm{d}t'\, H(t')}\,\Psi(\vec{r},t'), \quad t > t_0 \tag{5.23}$$

is simply the statement that the Hamiltonian generates the time evolution of the wave-function $\Psi(\vec{r},t)$. In non-relativistic physics (here, without electromagnetic potentials),

$$i\hbar\frac{\partial}{\partial t} = H = \frac{1}{2m}\left(\frac{\hbar}{i}\vec{\nabla}\right)^2 + V(\vec{r},t) \qquad \Leftrightarrow \qquad E = \frac{\vec{p}^{\,2}}{2m} + V(\vec{r},t), \tag{5.24}$$

the combination of which with equation (5.23) is the diffusion equation: of second order in spatial derivatives, but first order in the time derivative. This also implies the "quantization correspondence" (in the coordinate representation)

$$\vec{p} \;\leftrightarrow\; \vec{p} = \frac{\hbar}{i}\vec{\nabla}, \qquad \text{and} \qquad E \;\leftrightarrow\; H = i\hbar\frac{\partial}{\partial t}. \tag{5.25}$$

Instead of the non-relativistic relation (5.24), the relativistic version of the Schrödinger equation would have to correspond to the relativistic relation (3.37), and using the correspondences (5.25) we obtain

$$\begin{aligned} \vec{p}^{\,2}c^2 + m^2c^4 = E^2 \qquad &\leftrightarrow \qquad \left[c^2\left(\frac{\hbar}{i}\vec{\nabla}\right)^2 + m^2c^4\right]\Psi(\vec{r},t) = \left(i\hbar\frac{\partial}{\partial t}\right)^2\Psi(\vec{r},t), \\ &\Rightarrow \quad \left[\Box + \left(\frac{mc}{\hbar}\right)^2\right]\Psi(\vec{r},t) = 0. \end{aligned} \tag{5.26}$$

This is the so-called Klein–Gordon equation, where

$$\Box := \left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2\right] \tag{5.27}$$

is called the d'Alembertian or the wave operator.

[5] It will soon be shown that, as a consequence of the $U(1)$ gauge symmetry, the four functions $\Phi, \vec{A}$ represent only two physical degrees of freedom, which may be identified with two components of the vector $\vec{A}$ that are perpendicular to the direction of the photon motion.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

**Digression 5.3** Ironically, Schrödinger seems to have known [243] about the equation (5.26) before publishing the equation that soon acquired his name, but discarded it in the belief that the double-valuedness of the solution (3.37), $E = \pm c\sqrt{\vec{p}^2 + m^2c^2}$, precludes a probabilistic interpretation $|\Psi(\vec{r},t)|^2$. Wolfgang Pauli and Victor Weisskopf proved in 1934 that the essential obstacle to this interpretation of the quantity $|\Psi(\vec{r},t)|^2$ in relativistic physics is the fact that relativistic physics *must* contain the possibility of creating and annihilating particles, as permitted by conservation of energy, linear and angular momentum, charge, etc. This implies that the number of particles in relativistic physics is not a conserved quantity, and contradicts the elementary consequence of the Schrödinger equation:

$$(5.1) \quad \Rightarrow \quad \frac{\partial \varrho}{\partial t} = -\vec{\nabla}\cdot\vec{\mathcal{J}} + \frac{2}{\hbar}\,\Im m\left(V(\vec{r})\right)\varrho, \tag{5.28a}$$

$$\varrho(\vec{r},t) := |\Psi(\vec{r},t)|^2, \qquad \vec{\mathcal{J}}(\vec{r},t) := \frac{\hbar}{m}\,\Im m\left[\Psi^*(\vec{r},t)\vec{\nabla}\Psi(\vec{r},t)\right]. \tag{5.28b}$$

This shows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathscr{V}} \mathrm{d}^3\vec{r}\,\varrho(\vec{r},t) = -\oint_{\partial\mathscr{V}} \mathrm{d}^2\vec{\sigma}\cdot\vec{\mathcal{J}} \;+\; \frac{2}{\hbar}\int_{\mathscr{V}} \mathrm{d}^3\vec{r}\,\Im m\left(V(\vec{r})\right)\varrho(\vec{r},t). \tag{5.28c}$$

The probability of finding the particle (represented by $\Psi$) within the volume $\mathscr{V}$ changes only by the probability flowing through $\partial\mathscr{V}$ (the boundary of the volume $\mathscr{V}$) – if and only if the potential $V(\vec{r},t)$ is a real function where $\varrho(\vec{r},t)$ is nonzero. The number of particles is then also conserved, and this is indeed the case in standard quantum mechanics.

Motivated by the fact that the non-relativistic Schrödinger equation is of first order in time derivatives, while the Klein–Gordon equation is of the second order, Paul Dirac found a way to factorize the Klein–Gordon equation and so obtain a differential equation that is of first order both in spatial and in time derivatives. Indeed, in the rest-frame of the particle, $\vec{p} = 0$, so that the relativistic relation (3.37) reduces to

$$E^2 - m^2c^4 = 0 \qquad \Rightarrow \qquad (E + mc^2)(E - mc^2) = 0, \tag{5.29}$$

which is the desired factorization. With $\vec{p} \neq 0$, the desired factorization of the equivalent equation (3.36) is of the form

$$\begin{aligned} \mathrm{p}^2 - m^2c^2 = 0 \qquad \Rightarrow \qquad 0 &= (\boldsymbol{\beta}^\mu p_\mu + mc)(\boldsymbol{\gamma}^\nu p_\nu - mc), \\ &= \boldsymbol{\beta}^\mu\boldsymbol{\gamma}^\nu\, p_\mu p_\nu + mc(\boldsymbol{\gamma}^\mu - \boldsymbol{\beta}^\mu)p_\mu - m^2c^2. \end{aligned} \tag{5.30}$$

As the original equation $\mathrm{p}^2 - m^2c^2 = 0$ has no linear terms in the 4-momentum p, it must be that $\boldsymbol{\beta}^\mu = \boldsymbol{\gamma}^\mu$. Equating the quadratic terms one then obtains that

$$\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\, p_\mu p_\nu = \mathrm{p}^2 \equiv \eta^{\mu\nu}\, p_\mu p_\nu. \tag{5.31}$$

Since $p_\mu p_\nu = p_\nu p_\mu$, we in fact have the conditions

$$\left\{\,\boldsymbol{\gamma}^\mu\,,\,\boldsymbol{\gamma}^\nu\,\right\} = 2\eta^{\mu\nu}, \tag{5.32}$$

where $[\eta^{\mu\nu}] = \mathrm{diag}(1,-1,-1,-1)$ is the matrix-inverse of the metric tensor (3.19) of empty spacetime. This yields

$$\mathrm{p}^2 - m^2c^2 \;=\; 0 \;=\; (\boldsymbol{\gamma}^\mu p_\mu - mc)(\boldsymbol{\gamma}^\mu p_\mu + mc). \tag{5.33}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

Using the relativistic combination of the correspondences (5.25), this produces the Dirac equation:

$$p_\mu \to \frac{\hbar}{i}\partial_\mu \qquad \Rightarrow \qquad \left[i\hbar\boldsymbol{\gamma}^\mu\partial_\mu - mc\right]\Psi(\mathrm{x}) = 0, \tag{5.34}$$

where the standard abbreviation [☞ Digression 3.6 on p. 93]

$$\partial_\mu := \frac{\partial}{\partial x^\mu}, \qquad \to (-\tfrac{1}{c}\partial_t, \vec{\nabla}), \tag{5.35}$$

was introduced. The choice of the second of the two factors in equation (5.33) for defining the Dirac equation is an arbitrary, but standard choice.

The question remains, what sort of objects the $\boldsymbol{\gamma}^\mu$ are so as to satisfy the relations (5.32).

#### The Dirac spinor

Relations of the type (5.32) define so-called Clifford algebras. Their abstract structure, properties and representation theory had been established by mathematicians William Kingdon Clifford and Hermann Grassmann back in the second half of the nineteenth century. However, in the first half of the twentieth century, this was unknown among physicists, and Dirac independently found the smallest matrix realization of the $\boldsymbol{\gamma}^\mu$ objects, which today we call Dirac matrices; relation (5.32) then implicitly contains the unit $4\times4$ matrix in the right-hand side. There exist several "standard" choices of Dirac matrices; here we follow the traditional sources [64, 63] and use the so-called Dirac basis:

$$\boldsymbol{\gamma}^0 = \begin{bmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & -\mathbb{1} \end{bmatrix}, \qquad \boldsymbol{\gamma}^i = \begin{bmatrix} \mathbb{0} & \boldsymbol{\sigma}^i \\ -\boldsymbol{\sigma}^i & \mathbb{0} \end{bmatrix}, \quad i = 1,2,3. \tag{5.36}$$

To satisfy the relations (5.32), $\boldsymbol{\gamma}^\mu$ cannot be "ordinary" numbers but *can* be matrices. This implies that the operator that acts upon $\Psi(\mathrm{x})$ in the Dirac equation (5.34) also has to be a $4\times4$ matrix,[6] so $\Psi(\mathrm{x})$ must be a column-matrix with four components!

Recall that the solutions of the Schrödinger equation, e.g., for the hydrogen atom (4.8), yield $\Psi(\vec{r},t)$ as an expansion over spherical harmonics, $Y_\ell^m(\theta,\phi)$, which correspond to components of the "spin-$\ell$" representation[7] of the $SO(3) \overset{1-2}{\approx} SU(2)$ rotation group [☞ Table A.2 on p. 469]. For example, the hydrogen atom states with $\ell = 1$ and $m = \pm1, 0$ span the 3-vector representation of the rotation group, where it is also easy to define the Cartesian basis:

$$\begin{array}{ll} (\Psi_n)_x & := \frac{1}{2}\left(\Psi_{n,1,+1} + \Psi_{n,1,-1}\right), \\ (\Psi_n)_y & := \frac{1}{2i}\left(\Psi_{n,1,+1} - \Psi_{n,1,-1}\right), \end{array} \qquad (\Psi_n)_z := \Psi_{n,1,0}. \tag{5.37}$$

The elements of the $(2\ell+1)$-dimensional vector space $\{\Psi_{n,\ell,m},\ \text{for } |m| \leqslant \ell,\ \triangle m \in \mathbb{Z}\}$ may just as easily be represented as $(2\ell+1)$-component column-matrices.

However, the 4-component nature of the solutions to the Dirac equation represents an *additional* degree of freedom, a relativistic generalization of the "spin" factor that we used in Section 4.4.2, such as in the factorization (4.123), for example. Even for $\ell = 0$, the Dirac equation has four linearly independent solutions. In the simple case when $\vec{p} = 0$, the Dirac equation reduces to

$$\left[\frac{i\hbar}{c}\boldsymbol{\gamma}^0\frac{\partial}{\partial t} - mc\mathbb{1}\right]\Psi = 0. \tag{5.38}$$

[6] Since the $\boldsymbol{\gamma}^\mu$'s are $4\times4$ matrices, the Dirac equation should, pedantically, be written as $[i\hbar\boldsymbol{\gamma}^\mu p_\mu - mc\mathbb{1}]\Psi = 0$.

[7] When it denotes a rotation group representation, the term "spin-$j$" is simply short for "the total angular momentum where the eigenvalue of the quadratic operator $J^2$ equals $j(j+1)$," regardless of the physical original and composition of this total angular momentum.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

The solutions in Dirac's basis of $\boldsymbol{\gamma}$-matrices are

$$\Psi_A(t) = e^{-i(mc^2/\hbar)t}\begin{bmatrix}\Psi_1(0)\\ \Psi_2(0)\end{bmatrix}, \qquad \text{and} \qquad \Psi_B(t) = e^{+i(mc^2/\hbar)t}\begin{bmatrix}\Psi_3(0)\\ \Psi_4(0)\end{bmatrix}, \tag{5.39}$$

where $\Psi_B(t)$ represents the solutions with negative energy; i.e., *anti*-solutions with positive energy that move backwards in time, according to the Stückelberg–Feynman interpretation that is by now the standard understanding: $\Psi_B(t) \to \overline{\Psi}_B(-t)$ [☞ definition (5.49)].

Using the redefinition of solutions (wave-functions for particles) with negative energy as anti-solutions (wave-functions for antiparticles) with positive energy, the standard solutions (following the conventions of Ref. [243]) are

$$u^{\uparrow} \propto \begin{bmatrix}1\\ 0\\ \frac{p_z c}{E+mc^2}\\ \frac{(p_x+ip_y)c}{E+mc^2}\end{bmatrix}, \qquad u^{\downarrow} \propto \begin{bmatrix}0\\ 1\\ \frac{(p_x-ip_y)c}{E+mc^2}\\ \frac{p_z c}{E+mc^2}\end{bmatrix}, \tag{5.40}$$

$$v^{\downarrow} \propto \begin{bmatrix}\mathbb{O} & \mathbb{1}\\ \mathbb{1} & \mathbb{O}\end{bmatrix} u^{\uparrow} \propto \begin{bmatrix}\frac{p_z c}{E+mc^2}\\ \frac{(p_x+ip_y)c}{E+mc^2}\\ 1\\ 0\end{bmatrix}, \qquad v^{\uparrow} \propto \begin{bmatrix}\mathbb{O} & \mathbb{1}\\ \mathbb{1} & \mathbb{O}\end{bmatrix} u^{\downarrow} \propto \begin{bmatrix}\frac{(p_x-ip_y)c}{E+mc^2}\\ \frac{p_z c}{E+mc^2}\\ 0\\ 1\end{bmatrix}, \tag{5.41}$$

(5.42)

where $E = +\sqrt{\vec{p}^2c^2 + m^2c^4}$ always, and the solutions with negative energy are

$$u^{\uparrow}_{-}(E,\vec{p}) = -v^{\downarrow}(-E,-\vec{p}) \qquad \text{and} \qquad u^{\downarrow}_{-}(E,\vec{p}) = v^{\uparrow}(-E,-\vec{p}). \tag{5.43}$$

Note that $u^{\uparrow}, u^{\downarrow}, u^{\uparrow}_{-}, u^{\downarrow}_{-}$ are four linearly independent solutions to the Dirac equation (5.34), whereas $v^{\uparrow}, v^{\downarrow}$ satisfy the Dirac equation with $p_\mu \to -p_\mu$ – which precisely holds for the complementary factor in equation (5.33). The solutions to the Dirac equation may then be written as

$$\Psi(\mathrm{x}) = \sum_{s=\uparrow,\downarrow}\Big[N_u\, e^{-(i/\hbar)\mathrm{x}\cdot\mathrm{p}}\, u^s(\mathrm{p}) + N_v\, e^{-(i/\hbar)\mathrm{x}\cdot\mathrm{p}}\, v^s(\mathrm{p})\Big], \tag{5.44}$$

which represents the "plane wave" of a spin-$\frac{1}{2}$ particle, free of the influence of any potential. This $\Psi(\mathrm{x})$, however, is not a 4-vector in the 4-dimensional spacetime, but the so-called Dirac spinor, which we will see transforms with respect to Lorentz transformations, in an intrinsic fashion, distinct from 4-vectors.

#### Lorentz transformations of the Dirac spinor

From relation (A.121c), we see that the antisymmetrized products of two Dirac gamma matrices, $\boldsymbol{\gamma}^{\mu\nu} := \frac{i}{4}[\boldsymbol{\gamma}^{\mu}, \boldsymbol{\gamma}^{\nu}]$, close a Lie algebra:

$$\big[\,\boldsymbol{\gamma}^{\mu\nu}\,,\,\boldsymbol{\gamma}^{\rho\sigma}\,\big] = \eta^{\mu\rho}\boldsymbol{\gamma}^{\nu\sigma} - \eta^{\mu\sigma}\boldsymbol{\gamma}^{\nu\rho} + \eta^{\nu\sigma}\boldsymbol{\gamma}^{\mu\rho} - \eta^{\nu\rho}\boldsymbol{\gamma}^{\mu\sigma}. \tag{5.45}$$

It is not hard to verify that the definitions $J_j := \frac{1}{2i}\varepsilon_{jkl}\boldsymbol{\gamma}^{kl}$ and $K_j := i\boldsymbol{\gamma}^{0j}$ result in the commutation relations (5.45) written as

$$\big[\,J_j\,,\,J_k\,\big] = i\varepsilon_{jk}{}^{m}J_m, \qquad \big[\,J_j\,,\,K_k\,\big] = i\varepsilon_{jk}{}^{m}K_m, \qquad \big[\,K_j\,,\,K_k\,\big] = -i\varepsilon_{jk}{}^{m}J_m. \tag{5.46}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

While the $J_j$ elements generate $SO(3)$ rotations, the $K_j$ elements generate Lorentz boosts. The elements of the Lorentz group – in the representation that acts upon 4-component Dirac spinors – are obtained as exponential functions of the linear combinations of these six generators:

$$g(\vec{\varphi},\vec{\beta}) := \exp\left\{-i(\varphi^j J_j + \beta^j K_j)\right\} = \exp\left\{\beta_j \boldsymbol{\gamma}^{0j} - \varepsilon_{jkm}\varphi^j \boldsymbol{\gamma}^{km}\right\} = \exp\left\{\lambda_{\mu\nu}\boldsymbol{\gamma}^{\mu\nu}\right\}. \tag{5.47}$$

These may be shown to actually form a double covering of the $SO(1,3)$ group, denoted $Spin(1,3)$: to each non-identity element of the $SO(1,3)$ group there correspond precisely two elements of the $Spin(1,3)$ group. For example, the 360°-rotations of $SO(1,3)$ correspond to the $\pm\mathbb{1}$ elements of $Spin(1,3)$, and only the $SO(1,3)$ 720°-rotation corresponds to the unique element $\mathbb{1} \in Spin(1,3)$.

Let us just cite here [64] that the Lorentz boost in the $x^1$-direction causes the transformation

$$\Psi(\mathrm{x}) \to \left[\sqrt{\tfrac{1}{2}(\gamma+1)}\mathbb{1} - \sqrt{\tfrac{1}{2}(\gamma-1)}\boldsymbol{\gamma}^{01}\right]\Psi(\mathrm{x}), \tag{5.48}$$

where $\gamma$ with no index denotes the familiar relativistic factor $\gamma := \frac{1}{\sqrt{1-v^2/c^2}}$. It is then easy to verify that $\Psi^\dagger\Psi$ is not Lorentz-invariant, but that $\Psi^\dagger\boldsymbol{\gamma}^0\Psi$ is. One thus defines

$$\overline{\Psi} := \Psi^\dagger\boldsymbol{\gamma}^0 \tag{5.49}$$

as the Dirac-conjugate of the Dirac spinor $\Psi$, and note that $\overline{\Psi}\Psi$ is Lorentz-invariant.

Using the results from Appendix A.6.1, the following bilinear[8] functions may be constructed from a Dirac spinor and its Dirac-conjugate spinor:

| Expression | Lorentz representation | Number of independent components |
|---|---|---|
| $\overline{\Psi}\Psi$ | = scalar, | 1 |
| $\overline{\Psi}\boldsymbol{\gamma}^\mu\Psi$ | = 4-vector, | 4 |
| $\overline{\Psi}\boldsymbol{\gamma}^{\mu\nu}\Psi$ | = antisymmetric rank-2 tensor, | 6 |
| $\overline{\Psi}\boldsymbol{\gamma}^\mu\widehat{\boldsymbol{\gamma}}\Psi$ | = axial (i.e., pseudo-) 4-vector, | 4 |
| $\overline{\Psi}\widehat{\boldsymbol{\gamma}}\Psi$ | = pseudo-scalar, | 1 |

(5.50)

Since every complex 4×4 matrix may be written as a complex linear combination of 16 matrices (A.124) [580], the 16 functions (5.50) also form a complete system of bilinear functions of the Dirac spinor, $\Psi$. It is important to note that in the functions (5.50), the $\boldsymbol{\gamma}$-matrices do not transform with respect to the Lorentz transformations, but $\overline{\Psi}$ and $\Psi$ do, and in fact just so that each bilinear product as a whole transforms in the indicated fashion. For example, $\overline{\Psi}\boldsymbol{\gamma}^\mu\Psi$ really transforms, as a whole, as the components of any other contravariant 4-vector.

***Comment 5.5*** *The careful Reader may have questioned the identification of the matrices $J_j$ and $K_j$ as the rotation and Lorentz boosts. The list (5.50) gives unambiguous confirmation, in the form of the correct Lorentz transformations of the listed bilinear expressions.*

The notation (5.50) is standard, and supposes that one consistently uses that the $\boldsymbol{\gamma}^\mu, \boldsymbol{\gamma}^{\mu\nu}$ and $\widehat{\boldsymbol{\gamma}}$ are all 4×4 matrices, $\overline{\Psi}$ is a 4-component row-matrix, and $\Psi$ a 4-component column-matrix. Instead, one may also use the index notation, so the $A$th element of the column-matrix $\Psi$ is written $\Psi^A$, the $A$th element of the row-matrix $\overline{\Psi}$ is $\overline{\Psi}_A$, and similarly for the $\boldsymbol{\gamma}$-matrices, so that the expressions (5.50) become

$$\overline{\Psi}_A\Psi^A, \quad \overline{\Psi}_A\left(\gamma^\mu\right)^A{}_B\Psi^B, \quad \overline{\Psi}_A\left(\gamma^{\mu\nu}\right)^A{}_B\Psi^B, \quad \overline{\Psi}_A\left(\gamma^\mu\right)^A{}_B\left(\hat{\gamma}\right)^B{}_C\Psi^C, \quad \overline{\Psi}_A\left(\hat{\gamma}\right)^A{}_B\Psi^B. \tag{5.51}$$

[8] It is understood that "bilinear" here means "*anti*-linear + linear in $\overline{\Psi}$ and $\Psi$, respectively."

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

Similarly, instead of column-matrices (5.40)–(5.41), we may write[9] $u^{\uparrow A}$, etc., where, for example,

$$u^{\uparrow 1} = N, \quad u^{\uparrow 2} = 0, \quad u^{\uparrow 3} = N\frac{p_z c}{E + mc^2}, \quad u^{\uparrow 4} = N\frac{(p_x + ip_y)c}{E + mc^2}, \quad \text{etc.} \tag{5.52}$$

The normalizing factors in equation (5.44) are chosen so that

$$\overline{u^{\uparrow}}\, u^{\uparrow} = 2mc = \overline{u^{\downarrow}}\, u^{\downarrow} \qquad \text{and} \qquad \overline{v^{\uparrow}}\, v^{\uparrow} = -2mc = \overline{v^{\downarrow}}\, v^{\downarrow}. \tag{5.53}$$

The solutions of equations (5.40)–(5.41) are also complete, in the sense that[10]

$$\sum_{s=\uparrow,\downarrow} u^s\, \overline{u^s} = \not{p} + mc\mathbb{1} \qquad \text{and} \qquad \sum_{s=\uparrow,\downarrow} v^s\, \overline{v^s} = \not{p} - mc\mathbb{1}, \tag{5.54}$$

that is,

$$\sum_{s=\uparrow,\downarrow} u^{s,A}\, \overline{u^s}_B = (\gamma^\mu)^A{}_B p_\mu + mc\delta^A_B \qquad \text{and} \qquad \sum_{s=\uparrow,\downarrow} v^{s,A}\, \overline{v^s}_B = (\gamma^\mu)^A{}_B p_\mu - mc\delta^A_B. \tag{5.55}$$

The matrix (5.54) and the (explicit) index notation (5.55) may be used interchangeably, as needed and for the sake of compactness and clarity. Also, by the general Dirac spinor $\Psi$ one understands a general linear combination

$$\Psi \;:=\; \hat{\mathrm{e}}_A\, \Psi^A, \tag{5.56}$$

just as we write $\mathrm{x} = \hat{\mathrm{e}}_\mu\, x^\mu$ for a 4-vector. However, one must keep in mind that the $\hat{\mathrm{e}}_\mu$ are (Cartesian) unit vectors in the 4-dimensional spacetime in which we too move, whereas the $\hat{\mathrm{e}}_A$ are unit vectors in an abstract vector space of solutions to the Dirac equation.

#### Helicity, chirality and the Weyl equation

It is useful to note a very important difference between two seemingly similar properties of spin-$\frac{1}{2}$ particles: helicity and chirality. Much of the analysis here may be found in standard texts on particle physics and field theory as cited in the preface, but there is also a book dedicated to all matters of spin in particle physics [334]. The generalization of this analysis of course also exists for particles with arbitrary spin, subject however to the Weinberg–Witten theorem 6.1 on p. 249, as well as to higher-dimensional spacetime as needed in string theory.

Using the projectors (A.121b)

$$\boldsymbol{\gamma}_\pm := \tfrac{1}{2}[\mathbb{1} \pm \widehat{\boldsymbol{\gamma}}], \tag{5.57}$$

one defines in a fully Lorentz-invariant way:

$$\Psi_\pm := \boldsymbol{\gamma}_\pm \Psi, \qquad \text{so} \qquad \Psi_+ + \Psi_- = \Psi, \quad \boldsymbol{\gamma}_\pm \Psi_\pm = \Psi_\pm, \quad \boldsymbol{\gamma}_\pm \Psi_\mp = 0. \tag{5.58}$$

For $\Psi_+$ (also written as $\Psi_R$) one says that it has right-handed chirality, and $\Psi_-$ (also $\Psi_L$) has left-handed chirality. To this end, Weyl's basis (A.132) of Dirac matrices is particularly convenient. The complex 2-component projections $\Psi_\pm$ are Weyl spinors.

Independently of chirality, for particles with linear momentum $\vec{p}$ and spin $\vec{S}$, one defines the *helicity* operator, $h := \hat{p}\cdot\vec{S}/\hbar$, the eigenvalue of which is the helicity of the particle. With the mental (mnemonic and entirely fictitious!) image of the intrinsic angular momentum (spin) of the particle represented as the rotation of the particle itself, helicity may be represented as the "projection of the spin in the direction of motion." For example, a spin-$\frac{1}{2}$ particle may have helicity $+\frac{1}{2}$ or $-\frac{1}{2}$,

[9] Caution: the Dirac 4-spinors $u^\uparrow, u^\downarrow, v^\uparrow$ and $v^\downarrow$ are linearly independent and each has four components. Only a total of four of these components are linearly independent.

[10] Caution: the normalizations (5.53) and (5.54) differ from the standard quantum mechanical ones.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

depending on whether, respectively, it "spins" about the direction of motion in the right-hand sense or the left-hand sense.

Helicity is not defined in a Lorentz-invariant manner. Indeed, a particle with a nonzero mass always has a rest-frame wherein it does not move, and where $\vec{p} = 0$, so the eigenvalues of $h$ vanish. Also, it is always possible to pass such a particle, i.e., Lorentz-boost, into a coordinate system wherein the particle moves in the direction opposite to the original $\vec{p}$. Since this changes the sign of $\vec{p}$ but not of $\vec{S}$, the eigenvalues of $h$ also change their sign. It follows that helicity cannot be Lorentz-invariant for particles with nonzero mass.

For particles with no mass, helicity *is* Lorentz-invariant, and coincides with chirality.

— ❦ —

The solutions (5.40)–(5.43) of the Dirac equation (5.34) indicate that the upper and lower components of the Dirac spinor are not independent and it is not possible to separate them in a Lorentz-invariant way. The relations (A.121b) define the projectors $\boldsymbol{\gamma}_{\pm}$ that are Lorentz-invariant since the $\boldsymbol{\gamma}$-matrices do not change with respect to Lorentz transformations, which gives rise to the hope that the Dirac 4-component spinor may be separated into two 2-component spinors in a Lorentz-invariant way.

---

**Digression 5.4** One often finds a "quick" argument in the literature that $\boldsymbol{\gamma}$-matrices are Lorentz-invariant: supposedly, in the product $\boldsymbol{\gamma}^\mu p_\mu$, the Lorentz transformations act upon the physical quantity, the 4-momentum, and not on the $\boldsymbol{\gamma}$-matrices. This recalls the view that rotations of a vector $\vec{v} = \hat{\mathrm{e}}_i v^i$ act upon the basis elements $\hat{\mathrm{e}}_i$, not on the components, which are "only numerical values" in a given coordinate system. However, it is equally reasonable to adopt the vantage point where the *inverse* rotations act upon the components $v^i$, and not upon the basis elements $\hat{\mathrm{e}}_i$. Both applications of the transformations produce a net change in the physical quantity $\hat{\mathrm{e}}_i v^i$, which is regarded as the "active" transformation. By contrast, the "passive" transformation *simultaneously* rotates both the basis vectors $\hat{\mathrm{e}}_i$ as well as the components $v^i$ (in the inverse sense), so that the physical quantity $\vec{v}$ remains invariant.

However, this is not a case of active/passive action of the Lorentz transformations: The Dirac $\boldsymbol{\gamma}$-matrices indeed are components of a 4-vector, but those components are *matrices*, the rows of which are in the basis of the Dirac 4-component spinor $\Psi$, and the columns of which are in the basis of the Dirac-conjugated spinor $\overline{\Psi}$. The Lorentz transformations act upon all three bases, and those actions mutually cancel so that the $\boldsymbol{\gamma}$-matrices remain invariant. In other words, the product $\overline{\Psi}\boldsymbol{\gamma}^\mu\Psi\, p_\mu$ is evidently Lorentz-invariant: $\overline{\Psi}\boldsymbol{\gamma}^\mu\Psi$ is a contravariant 4-vector and $p_\mu$ a covariant one, so $\overline{\Psi}\boldsymbol{\gamma}^\mu\Psi\, p_\mu$ is the scalar product of a contravariant 4-vector and a covariant 4-vector. By adapting the index notation so as to also count the components of the Dirac spinor (5.51), we have

$$(\overline{\Psi}\boldsymbol{\gamma}^\mu\Psi)\, p_\mu = \big(\overline{\Psi}_A\, (\boldsymbol{\gamma}^\mu)^A{}_B\, \Psi^B\big)\, p_\mu, \tag{5.59a}$$

so that the numerical values $(\boldsymbol{\gamma}^\mu)^A{}_B$ for each fixed $\mu, A, B$ are simply the Clebsch–Gordan coefficients in the expansion of the product $\overline{\Psi} \times \Psi$ in a spacetime 4-vector basis. In turn, the coefficients $(\boldsymbol{\gamma}^\mu)^A{}_B$ also appear in the tri-linear Lorentz-*invariant* contraction of the basis vectors $\hat{\mathrm{e}}_A (\boldsymbol{\gamma}^\mu)^A{}_B\, \hat{\mathrm{e}}_\mu\, \hat{\mathrm{e}}^B$ [☞ Section A.6]. Lastly, rewriting the above equation as

$$(\overline{\Psi}\boldsymbol{\gamma}^\mu\Psi)\, p_\mu = (\boldsymbol{\gamma}^\mu)^A{}_B\, \big(\overline{\Psi}_A\, \Psi^B\, p_\mu\big) \tag{5.59b}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

re-interprets the matrices $\boldsymbol{\gamma}^\mu$ as the (in general) $\mathbb{C}$-valued projection of the direct product of Dirac-conjugate Dirac-spinors, Dirac-spinors and 4-momenta to Lorentz-invariant (in general) complex numbers: $\boldsymbol{\gamma} : \{\overline{\Psi}\} \times \{\Psi\} \times \{\mathrm{p}\} \to \mathbb{C}$.

However, it is not hard to show that helicity projections do not commute with the Dirac matrices:

$$[\boldsymbol{\gamma}_\pm, \boldsymbol{\gamma}^\mu] \neq 0 : \qquad [\mathbb{1}, \boldsymbol{\gamma}^\mu] = 0 = \{\widehat{\boldsymbol{\gamma}}, \boldsymbol{\gamma}^\mu\} \quad \Rightarrow \quad \boldsymbol{\gamma}_\pm \boldsymbol{\gamma}^\mu = \boldsymbol{\gamma}^\mu \boldsymbol{\gamma}_\mp. \tag{5.60}$$

Owing to this, an attempt to use the projections (5.58) on the Dirac equation yields

$$\begin{aligned} \boldsymbol{\gamma}_\pm [i\hbar \boldsymbol{\gamma}^\mu \partial_\mu - mc\mathbb{1}]\Psi &= [i\hbar \boldsymbol{\gamma}_\pm \boldsymbol{\gamma}^\mu \partial_\mu - mc\boldsymbol{\gamma}_\pm \mathbb{1}]\Psi = [i\hbar \boldsymbol{\gamma}^\mu \boldsymbol{\gamma}_\mp \partial_\mu - mc\mathbb{1}\boldsymbol{\gamma}_\pm]\Psi \\ &= i\hbar \boldsymbol{\gamma}^\mu (\partial_\mu \Psi_\mp) - mc\Psi_\pm, \end{aligned} \tag{5.61}$$

which is a system of differential equations that couples $\Psi_+$ and $\Psi_-$ precisely when $m \neq 0$. Conversely,

$$\boldsymbol{\gamma}^\mu \partial_\mu \Psi_\pm = 0 \qquad \Leftrightarrow \qquad m\Psi_\pm = 0. \tag{5.62}$$

***Conclusion 5.2 (Weyl)*** *The Dirac spinor $\Psi$ separates in a Lorentz-invariant way into the* ***right-handed*** *$\Psi_+ \equiv \Psi_R := \boldsymbol{\gamma}_+ \Psi$ and* ***left-handed*** *$\Psi_- \equiv \Psi_L := \boldsymbol{\gamma}_- \Psi$ 2-component Weyl spinor (the eigen-spinors of the $\widehat{\boldsymbol{\gamma}}$ matrix) precisely when the mass of the particle vanishes.*

These (Weyl) spinors satisfy the simpler differential equations, $\boldsymbol{\gamma}^\mu \partial_\mu \Psi_\pm = 0$. Indeed, the Dirac differential equation (5.34) is a system of four coupled differential equations for four components of the Dirac spinor $\Psi$. By contrast, $\boldsymbol{\gamma}^\mu \partial_\mu \Psi_\pm = 0$ is a system of two coupled differential equations for two components of the Weyl spinor $\Psi_+$ and separately for $\Psi_-$.

Hermann Weyl noticed and published the characteristics of this special case of the Dirac equation in 1929. Yet, when Pauli invented the neutrino so as to preserve the energy conservation law, he did not want to use Weyl's equations on the grounds that they permit violating the symmetry of parity.[11] To wit, the Lorentz-invariant separation of $\Psi_+$ and $\Psi_- \propto P(\Psi_+)$ permits an independent – and different – treatment of these two halves of the Dirac spinor of opposite chirality. This is quite ironic, since Pauli did correctly predict the mass of the neutrino to be either very teeny or vanishing, and even during his own life it became clear that Nature really treats the left-handed neutrino very differently from the right-handed one. Until the discovery of the see-saw mechanism [☞ Section 7.3.2], the Weyl equations provided a much better model for neutrinos, and describe the maximal parity violation as observed in Nature.

— ❦ —

The frequent confusion of helicity and chirality has been fostered by the fact that massless particles are a specially simple case both for chirality and for helicity, where these two different physical quantities coincide. On the other hand, the Lorentz invariance of chirality is of fundamental importance in the contemporary formulation of weak and electroweak interactions, while helicity is easier to measure. The Reader should strive to conceptually differentiate and carefully distinguish between these two inherently different quantities.

[11] Up to the experimental confirmations of parity violation in weak interactions [☞ Sections 2.4.2 and 4.2.1], Pauli had, just as many other renowned physicists of the time, ardently advocated against ideas that include parity violation; see, e.g., A. Salam's Nobel lecture [473].

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

**The Dirac Lagrangian density**

The construction of the Dirac Lagrangian density is straightforward, if we only require the variation of the Hamilton action with that Lagrangian density to produce the Dirac equation. First, note that $\Psi$ and $\overline{\Psi}$ may *formally* be treated as independent quantities. The Dirac equation (5.34) is then simply multiplied *from the left* by $\overline{\Psi}$ (and by $c$ for units) and we identify

$$\mathscr{L}_D = \beta\overline{\Psi}(\mathrm{x})\left[c\not{p} + mc^2\mathbb{1}\right]\Psi(\mathrm{x}) = -\beta\overline{\Psi}(\mathrm{x})\left[i\hbar c\boldsymbol{\gamma}^\mu\partial_\mu - mc^2\mathbb{1}\right]\Psi(\mathrm{x}), \tag{5.63}$$

where $\beta$ is an arbitrary overall sign, since the variation by $\overline{\Psi}$ yields a $\beta$-multiple of equation (5.34). Variation by $\Psi$ yields the Hermitian conjugate of equation (5.34), i.e., nothing new (and nothing unneeded).

---

**Digression 5.5** The Dirac spinor $\Psi$ is a 4-tuple of formally anticommutative variables. In the general case, if $\psi$ and $\chi$ are anticommutative and $f$ and $g$ are commutative variables, we have that

$$[f,g] = 0, \qquad [f,\psi] = [f,\chi] = 0 = [g,\psi] = [g,\chi], \qquad \text{but} \qquad \{\psi,\chi\} = 0; \tag{5.64a}$$

$$\left[\frac{\partial}{\partial f},\frac{\partial}{\partial g}\right] = 0, \qquad \left[\frac{\partial}{\partial f},\frac{\partial}{\partial \psi}\right] = \left[\frac{\partial}{\partial f},\frac{\partial}{\partial \chi}\right] = 0 = \left[\frac{\partial}{\partial g},\frac{\partial}{\partial \psi}\right] = \left[\frac{\partial}{\partial g},\frac{\partial}{\partial \chi}\right], \quad \text{but} \quad \left\{\frac{\partial}{\partial \psi},\frac{\partial}{\partial \chi}\right\} = 0. \tag{5.64b}$$

Also,

$$\frac{\partial}{\partial\psi}\chi = -\chi\frac{\partial}{\partial\psi} \qquad \text{and} \qquad \frac{\partial}{\partial\chi}\psi = -\psi\frac{\partial}{\partial\chi}, \tag{5.64c}$$

which the Student must keep in mind when deriving the equations of motion from Lagrangian densities that also contain fermionic (anticommutative) variables. It is convenient to define the *right-derivative*:

$$\psi\frac{\overleftarrow{\partial}}{\partial\psi} = 1, \quad (\psi\chi)\frac{\overleftarrow{\partial}}{\partial\psi} = -\left(\psi\frac{\overleftarrow{\partial}}{\partial\psi}\right)\chi = -\chi, \quad (\psi\chi)\frac{\overleftarrow{\partial}}{\partial\chi} = \psi\left(\chi\frac{\overleftarrow{\partial}}{\partial\chi}\right) = \psi, \quad \text{etc.}, \tag{5.64d}$$

and diligently apply derivatives *either* from right *or* from left.

---

The definition of the Lagrangian allows us to identify the components of $\Psi$ as the *canonical coordinates*, so we may also define the canonically conjugate momentum densities:

$$\pi_\Psi := \mathscr{L}_D\frac{\overleftarrow{\partial}}{\partial\dot{\Psi}} = \left(-\beta\overline{\Psi}[i\hbar c\boldsymbol{\gamma}^\mu\partial_\mu - mc^2\mathbb{1}]\Psi\right)\frac{\overleftarrow{\partial}}{\partial(c\partial_0\Psi)} = -i\beta\hbar\overline{\Psi}\boldsymbol{\gamma}^0 = -i\beta\hbar\Psi^\dagger, \tag{5.65}$$

where we applied the right-derivative [☞ Digression 5.5 on p. 180]. The Hamiltonian then becomes

$$\begin{aligned}\mathscr{H}_D &= \pi_\Psi\dot{\Psi} - \mathscr{L}_D = (-i\beta\hbar\Psi^\dagger)(\dot{\Psi}) + \beta\overline{\Psi}[i\hbar c\boldsymbol{\gamma}^\mu\partial_\mu - mc^2\mathbb{1}]\Psi \\ &= -\beta\Psi^\dagger H\Psi + \beta\overline{\Psi}[i\hbar c\boldsymbol{\gamma}^\mu\partial_\mu - mc^2\mathbb{1}]\Psi, \qquad H \equiv i\hbar\tfrac{\partial}{\partial t}.\end{aligned} \tag{5.66}$$

The sign $\beta$ in the computation (5.65) may now be determined as follows: For an on-shell Dirac fermion, i.e., one that satisfies the equations of motion (5.34), the second term in the expression (5.66) vanishes, and we obtain

$$\mathscr{H}_D\big|_{(5.34)} = -\beta\Psi^\dagger H\Psi, \qquad \text{where} \quad [\Psi] = \tfrac{1}{L^{3/2}}. \tag{5.67}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

We thus choose $\beta = -1$ for the total energy (the Hamiltonian) of the Hamilton–Jacobi canonical formalism on-shell and the expectation value of the quantum-mechanical operator $H$ to have the same sign. To sum up:

$$\mathscr{L}_D = -\overline{\Psi}(\mathrm{x})\left[c\not{p} + mc^2\mathbb{1}\right]\Psi(\mathrm{x}) = \overline{\Psi}(\mathrm{x})\left[i\hbar c\boldsymbol{\gamma}^\mu\partial_\mu - mc^2\mathbb{1}\right]\Psi(\mathrm{x}), \tag{5.68a}$$

$$\pi_\Psi = i\hbar\Psi^\dagger, \tag{5.68b}$$

$$\mathscr{H}_D = \pi_\Psi\dot{\Psi} - \mathscr{L}_D = \Psi^\dagger H\Psi - \overline{\Psi}[i\hbar c\boldsymbol{\gamma}^\mu\partial_\mu - mc^2\mathbb{1}]\Psi. \tag{5.68c}$$

Also, since the Dirac equation (5.34) may be written as

$$i\hbar\dot{\Psi} = H_D\Psi := \left[(i\hbar c\vec{\boldsymbol{\gamma}}\cdot\vec{\nabla} + mc^2)\boldsymbol{\gamma}^0\right]\Psi, \tag{5.69}$$

and $H_D$ is the on-shell Dirac Hamiltonian operator: $\mathscr{H}_D \overset{(5.34)}{=} \overline{\Psi}H_D\Psi$, as arranged in equation (5.67).

**Digression 5.6** The Dirac Lagrangian densities are often "antisymmetrized" using the identity

$$\int \mathrm{d}^4\mathrm{x}\,\overline{\Psi}\boldsymbol{\gamma}^\mu\partial_\mu\Psi = \tfrac{1}{2}\int \mathrm{d}^4\mathrm{x}\,\overline{\Psi}\boldsymbol{\gamma}^\mu\partial_\mu\Psi \;+\; \tfrac{1}{2}\int \mathrm{d}^4\mathrm{x}\left[\partial_\mu(\overline{\Psi}\boldsymbol{\gamma}^\mu\Psi) - (\partial_\mu\overline{\Psi})\boldsymbol{\gamma}^\mu\Psi\right] \tag{5.70a}$$

$$= \tfrac{1}{2}\int \mathrm{d}^4\mathrm{x}\left[\overline{\Psi}\boldsymbol{\gamma}^\mu\partial_\mu\Psi - (\partial_\mu\overline{\Psi})\boldsymbol{\gamma}^\mu\Psi\right] \;+\; \underbrace{\tfrac{1}{2}\oint_{\mathscr{V}} \mathrm{d}^3(\mathrm{x})_\mu\,(\overline{\Psi}\boldsymbol{\gamma}^\mu\Psi)}_{=0}, \tag{5.70b}$$

where the third, 3-dimensional integral is computed over the 3-dimensional boundary of spacetime, which is "at infinity." Physical fields are required to vanish there. We thus write

$$\overline{\Psi}\boldsymbol{\gamma}^\mu\partial_\mu\Psi \simeq \tfrac{1}{2}\left[\overline{\Psi}\boldsymbol{\gamma}^\mu\partial_\mu\Psi - (\partial_\mu\overline{\Psi})\boldsymbol{\gamma}^\mu\Psi\right] =: \tfrac{1}{2}\left(\overline{\Psi}\boldsymbol{\gamma}^\mu\overset{\leftrightarrow}{\partial}_\mu\Psi\right), \tag{5.70c}$$

where the middle expression defines the symbol $\overset{\leftrightarrow}{\partial}_\mu$. So antisymmetrized, we have that

$$\mathscr{L}_D \simeq -\overline{\Psi}(\mathrm{x})\left[\tfrac{1}{2}c\overset{\leftrightarrow}{\not{p}} + mc^2\mathbb{1}\right]\Psi(\mathrm{x}) = \overline{\Psi}(\mathrm{x})\left[\tfrac{i}{2}\hbar c\overset{\leftrightarrow}{\not{\partial}} - mc^2\mathbb{1}\right]\Psi(\mathrm{x}). \tag{5.70d}$$

Finally, the components of the canonically conjugate momentum density (5.68b) are constantly proportional to the Hermitian conjugates of the components of the Dirac spinor itself. Roughly speaking, one half of the Dirac (4-component) spinor are canonical coordinates of the system, the other half are conjugate momenta. The choice of which particular components are regarded as coordinates and which are momenta is, of course, arbitrary – up to the condition that the relations

$$\left\{\,\Psi\,,\,(i\hbar\Psi^\dagger)\,\right\} = i\hbar\mathbb{1} \qquad \Rightarrow \qquad \left\{\,\Psi\,,\,\Psi^\dagger\,\right\} = \mathbb{1} \tag{5.71}$$

produce the canonical *anticommutation* relations between the canonical momenta and the canonical coordinates. This arbitrariness is identical to that in classical physics.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

### 5.2.2 The $U(1)$ gauge symmetry and photons

Classical electrodynamics builds on the Maxwell equations,

$$\text{(Gauss)}\begin{cases} \vec{\nabla}\cdot\vec{E} = \dfrac{1}{4\pi\epsilon_0}4\pi\rho_e, \qquad \vec{\nabla}\times(c\vec{B}) - \dfrac{1}{c}\dfrac{\partial\vec{E}}{\partial t} = \dfrac{1}{4\pi\epsilon_0}\dfrac{4\pi}{c}\vec{j}_e, & \text{(Ampère)} \quad (5.72a) \\ \vec{\nabla}\cdot(c\vec{B}) = \dfrac{\mu_0}{4\pi}4\pi\rho_m, \qquad -\vec{\nabla}\times\vec{E} - \dfrac{1}{c}\dfrac{\partial(c\vec{B})}{\partial t} = \dfrac{\mu_0}{4\pi}\dfrac{4\pi}{c}\vec{j}_m, & \text{(Faraday)} \quad (5.72b) \end{cases}$$

that encompass the indicated laws, and where $c = \frac{1}{\sqrt{\epsilon_0\,\mu_0}}$ is the speed of propagation of light in vacuum. The densities of the magnetic (monopole!) charges, $\rho_m$, and currents, $\vec{j}_m$, are included for later discussion of electro-magnetic duality [☞ Section 11.4]. No experiment indicates their existence, so that the equations (5.72b) are cited in the literature almost exclusively with $\rho_m \to 0$ and $\vec{j}_m \to 0$. However, note that the units satisfy $[\rho_e/\epsilon_0] = [\mu_0\,\rho_m]$, as well as $[\vec{j}_e/\epsilon_0] = [\mu_0\,\vec{j}_m]$.

#### The relativistic description

For the purposes of a relativistic description of electrodynamics [☞ also Comment 8.1 on p. 294], we introduce[12]

$$A_\mu := (\Phi, -c\vec{A}), \quad \text{(gauge potential)} \qquad A^\mu := \eta^{\mu\nu}A_\nu = (\Phi, c\vec{A}); \quad (5.73a)$$

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu, \quad \begin{pmatrix}\text{antisymmetric}\\ \text{rank-2 tensor}\end{pmatrix} \qquad F^{\mu\nu} := \eta^{\mu\rho}F_{\rho\sigma}\eta^{\sigma\nu}; \quad (5.73b)$$

and identify

$$F_{00} = 0, \qquad F^{00} = 0, \quad (5.73c)$$

$$F_{0i} = \partial_0 A_i - \partial_i A_0 = \frac{1}{c}\frac{\partial(-cA_i)}{\partial t} - \frac{\partial\Phi}{\partial x^i} = E_i, \qquad F^{0i} = \eta^{00}F_{0j}\eta^{ji} = -E_i, \quad (5.73d)$$

$$F_{ij} = \partial_i A_j - \partial_j A_i = \frac{\partial(-cA_j)}{\partial x^i} - \frac{\partial(-cA_i)}{\partial x^j}$$
$$= c\Big(\frac{\partial A_i}{\partial x^j} - \frac{\partial A_j}{\partial x^i}\Big) = c\varepsilon_{ji}{}^k B_k \qquad = -c\varepsilon_{ij}{}^k B_k, \quad F^{ij} = \eta^{ik}F_{kl}\eta^{jl} = -c\varepsilon^{ijk}B_k \quad (5.73e)$$

and, of course, $F_{\mu\nu} = -F_{\nu\mu}$. In matrix form, we have

$$\big[\,F_{\mu\nu}\,\big] = \begin{bmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -cB_3 & cB_2 \\ -E_2 & cB_3 & 0 & -cB_1 \\ -E_3 & -cB_2 & cB_1 & 0 \end{bmatrix}, \quad \big[\,F^{\mu\nu}\,\big] = \begin{bmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -cB_3 & cB_2 \\ E_2 & cB_3 & 0 & -cB_1 \\ E_3 & -cB_2 & cB_1 & 0 \end{bmatrix}. \quad (5.74)$$

Since $F_{\mu\nu}$ are components of a rank-2 tensor, it follows that the Lorentz transformations act by [☞ Digression 3.5 on p. 91]

$$y^\mu = \mathrm{L}^\mu{}_\nu x^\nu \quad \Rightarrow \quad F_{\mu\nu}(\mathrm{y}) = \mathrm{L}^\rho{}_\mu\, F_{\rho\sigma}(\mathrm{x})\,\mathrm{L}^\sigma{}_\nu, \quad \text{i.e.} \quad \mathbf{F}(\mathrm{y}) = \mathbf{L}^T\,\mathbf{F}(\mathrm{x})\,\mathbf{L}. \quad (5.75)$$

The familiar Lagrangian [☞ also Exercises 5.1.3 and 5.1.4] for the electromagnetic field may thus be written as

$$\mathscr{L}_{EM} = -\tfrac{4\pi\epsilon_0}{4}F_{\mu\nu}\,F^{\mu\nu}. \quad (5.76)$$

[12] The negative relative sign in the definition of $A_\mu$ cancels the difference in signs in the definition (5.13), an additional factor of $c$ equates the units of $\Phi$ and $\vec{A}$, which stem from the difference between $D_t$ and $\vec{D}$.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

**Example 5.1** Let $\vec{E} = \hat{e}^2 E_2$ and $\vec{B} = 0$ be given in an inertial Cartesian coordinate system $S$, and let the inertial system $\widetilde{S}$ move with respect to $S$ with the constant speed $\hat{e}^1 v_1$. The relations (5.75) then yield

$$\widetilde{E}_2 = \gamma E_2, \qquad \text{and also} \qquad \widetilde{B}_3 = \gamma \frac{v_1}{c^2} E_2. \tag{5.77}$$

A field that in one inertial coordinate system looks like a "purely" electric field, can in another inertial system easily be represented by a *combination* of electric and magnetic fields. Notice, however, that the equation $\vec{E}\cdot\vec{B} = 0$ remains valid. Indeed, this is a Lorentz-invariant characteristic of the specified field [☞ relations (5.80a)].

The Maxwell equations (5.72a) may then also be written as

$$\partial_\mu F^{\mu\nu} = \frac{1}{4\pi\epsilon_0}\frac{4\pi}{c} j_e^\nu, \tag{5.78}$$

where $\mathrm{j}_e = (c\rho_e, \vec{j}_e)$ is the 4-vector of electric charge and current densities. Analogously, the Maxwell equations (5.72b) may be written also as

$$\tfrac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\partial_\mu F_{\nu\rho} = \frac{\mu_0}{4\pi}\frac{4\pi}{c} j_m^\sigma, \tag{5.79}$$

where $\mathrm{j}_m = (c\rho_m, \vec{j}_m)$ is the 4-vector of (monopole) magnetic charge and current densities.

**Digression 5.7** Direct substitution yields

$$\tfrac{1}{2}F_{\mu\nu}F^{\mu\nu} = \vec{E}^2 - c^2\vec{B}^2 \qquad \text{and} \qquad \tfrac{1}{4}\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = -c\vec{E}\cdot\vec{B}, \tag{5.80a}$$

which, using the transformations (5.75), shows that these two bilinear expressions in $\vec{E}$ and $\vec{B}$ are Lorentz-invariant. Evidently, these are the only linearly independent Lorentz-invariant bilinear expressions in $F_{\mu\nu}$ and $F^{\mu\nu}$, and so then also in $\vec{E}$ and $\vec{B}$. Since the Lagrangian density for electrodynamics must be a scalar (invariant) density and quadratic in electric and magnetic fields, we find that the Lagrangian density must be of the form

$$\mathscr{L}_{EM} = C_1\, F_{\mu\nu}F^{\mu\nu} + C_2\, \varepsilon^{\mu\nu\rho\sigma}\, F_{\mu\nu}F_{\rho\sigma}. \tag{5.80b}$$

The coefficients $C_1, C_2$ are chosen so that the variation of the Hamilton action, $\delta\int \mathrm{d}^4\mathrm{x}\, \mathscr{L}_{EM} = 0$, reproduces the Maxwell equations. The fact that this renders $C_2 = 0$ then poses the (unanswered ☹) question: Why is, in the possible "addition"

$$\mathscr{L}_{\vartheta,EM} = \vartheta\, \tfrac{4\pi\epsilon_0}{4}\varepsilon^{\mu\nu\rho\sigma}\, F_{\mu\nu}F_{\rho\sigma}, \tag{5.80c}$$

to the standard Lagrangian density (5.76) of the parameter $\vartheta = 0$, either identically or up to experimental error (i.e., $\vartheta \lll 1$)?

Direct substitution of $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ on the left-hand side of equation (5.79) yields

$$\tfrac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\partial_\mu\left(\partial_\nu A_\rho - \partial_\rho A_\nu\right) = \tfrac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\partial_\mu\partial_\nu A_\rho - \tfrac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\partial_\mu\partial_\rho A_\nu, \tag{5.81}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

where both terms vanish separately, since

$$\varepsilon^{\mu\nu\rho\sigma}\partial_\mu\partial_\nu = \underbrace{\varepsilon^{\mu\nu\rho\sigma}\partial_\nu\partial_\mu}_{\mu\leftrightarrow\nu} = \varepsilon^{\nu\mu\rho\sigma}\partial_\mu\partial_\nu = -\varepsilon^{\mu\nu\rho\sigma}\partial_\mu\partial_\nu. \tag{5.82}$$

That is,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad \overset{(5.79)}{\Longleftrightarrow} \quad 0 = \frac{\mu_0}{4\pi}\frac{4\pi}{c} j^\sigma_m. \tag{5.83}$$

The existence of magnetic charges and currents would then be an *obstruction* for equating $F_{\mu\nu}$ with $\partial_\mu A_\nu - \partial_\nu A_\mu$, i.e., the electric and the magnetic fields could not be expressed in terms of an *unambiguously specified* 4-vector potential (5.15) [☞ Section 5.2.3], and conversely: if $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ for an unambiguously specified 4-vector potential $A_\mu(\mathrm{x})$, then no monopole magnetic charge or current can exist. We thus have:

> ***Conclusion 5.3*** *Electric and magnetic charges and currents exist* ***simultaneously*** *if and only if there can be no* ***unambiguously*** *specified 4-vector potential $A_\mu(\mathrm{x})$ for which the electromagnetic field would be $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ [☞ Section 5.2.3].*

---

**Digression 5.8** Define a "differential 2-form" $\mathbf{F} := F_{\mu\nu}\,\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu$, where "$\wedge$" denotes the antisymmetric product of the differentials, as well as the operator $\mathbf{d} := \mathrm{d}x^\mu\partial_\mu$. Then

$$\mathbf{d}\wedge\mathbf{d} \equiv \tfrac{1}{2}(\partial_\mu\partial_\nu - \partial_\nu\partial_\mu)\,\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu \equiv 0. \tag{5.84a}$$

The Maxwell equations (5.72b), i.e., (5.79), are then equivalent to

$$\mathbf{d}\wedge\mathbf{F} = \mathbf{j}_m, \qquad \mathbf{j}_m := \frac{\mu_0}{4\pi}\frac{4\pi}{c} j^\sigma_m\,\varepsilon_{\mu\nu\rho\sigma}\,\mathrm{d}x^\mu\wedge\mathrm{d}x^\nu\wedge\mathrm{d}x^\rho, \tag{5.84b}$$

and the differential 3-form $\mathbf{j}_m$ is the *obstruction* for equating the differential 2-form $\mathbf{F}$ with $\mathbf{d}\wedge\mathbf{A}$, for any differential 1-form $\mathbf{A} = A_\mu\mathrm{d}x^\mu$. $\mathbf{F}$ is said to be a *nontrivial* (non-exact) 2-form.[13]

On the other hand, equations (5.72a), i.e., (5.78), may also be written in the form (5.84b). To this end, however, we need one more item of notation: in tensorial notation, any antisymmetric rank-$r$ tensor may be turned into an antisymmetric rank-$(4-r)$ tensor by contracting with $\varepsilon_{\mu\nu\rho\sigma}$ or $\varepsilon^{\mu\nu\rho\sigma}$. Thus, a 4-vector $j^\mu_m$ is "translated" into a rank-3 tensor $j^\mu_m \to (j^\mu_m\varepsilon_{\mu\nu\rho\sigma})$ and a 3-form $\mathbf{j}_m$. A double use of this operation yields $\frac{1}{2}\epsilon^{\mu\nu\rho\sigma}\partial_\nu(\frac{1}{2}\varepsilon_{\rho\sigma\alpha\beta}F^{\alpha\beta}) = \partial_\nu F^{\mu\nu}$. The corresponding operation with differential forms is the so-called "Hodge star," which turns an $r$-form into a $(4-r)$-form: $*\mathbf{A}$ is a 3-form, $*\mathbf{j}_m$ a 1-form, etc. The Maxwell equations (5.72a) and (5.78) are thus equivalent to

$$\mathbf{d}\wedge *\,\mathbf{F} = \mathbf{j}_e, \qquad \mathbf{j}_e := \frac{1}{4\pi\epsilon_0}\frac{4\pi}{c} j^\mu_e\,\varepsilon_{\mu\nu\rho\sigma}\mathrm{d}x^\nu\wedge\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma. \tag{5.84c}$$

Equations (5.84b) and (5.84c) respectively provide a compact form of the Maxwell equations:

$$\mathbf{d}\wedge\mathbf{F} = \mathbf{j}_m \quad \text{and} \quad \mathbf{d}\wedge *\mathbf{F} = \mathbf{j}_e. \tag{5.84d}$$

Since $\mathbf{d}\wedge\mathbf{d} \equiv 0$, $\mathbf{d}\wedge\mathbf{d}\wedge(*\mathbf{F}) = \mathbf{d}\wedge\mathbf{j}_e$ produces $\mathbf{d}\wedge\mathbf{j}_e = 0$, which is the well-known continuity equation (2.66), the integral of which yields the electric charge conservation law [☞ also Section 6.1.2]. Similarly, $\mathbf{d}\wedge\mathbf{d}\wedge\mathbf{F} = \mathbf{d}\wedge\mathbf{j}_m$ implies $\mathbf{d}\wedge\mathbf{j}_m = 0$, the continuity equation, and thus the (monopole) magnetic charge conservation law.

---

[13] By the same token is "đ$Q$," in thermodynamics in general, a nontrivial 1-form and not an exact differential.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

***Comment 5.6*** *The fact that the existence of (monopole) magnetic charges and currents obstructs the expression of the electromagnetic field $F_{\mu\nu}$ as an antisymmetric derivative of an* ***unambiguously*** *specified 4-vector potential $A_\mu$ points to a significant difference between electric and magnetic charges and currents – in spite of the fact that the Maxwell equations (5.72)* ***look*** *"symmetric." This "symmetry" – a* ***duality****, more precisely – is the mapping*

$$\varpi_{EM} \,:\, F^{\mu\nu} \;\longleftrightarrow\; (*\mathbf{F})^{\mu\nu} = [\tfrac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}] = \begin{bmatrix} 0 & -cB_1 & -cB_2 & -cB_3 \\ cB_1 & 0 & E_3 & -E_2 \\ cB_2 & -E_3 & 0 & E_1 \\ cB_3 & E_2 & -E_1 & 0 \end{bmatrix}, \tag{5.85}$$

*which swaps the roles of $\vec{E}$ and $c\vec{B}$. This implies that the vanishing of $\rho_e$ and $\vec{j}_e$ is a necessary and sufficient condition for the existence of some unambiguously specified 1-form $\widetilde{\mathbf{A}}$ such that $*\mathbf{F} = \mathbf{d}\wedge\widetilde{\mathbf{A}}$; here, $\widetilde{\mathbf{A}} = dx^\mu \widetilde{A}_\mu$ is the 1-form of the* ***dual*** *4-vector of gauge potentials.*

***Conclusion 5.4*** *The difference between $\mathbf{F}$ and $*\mathbf{F}$, i.e., $F_{\mu\nu}$ and $\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}$, i.e., $\vec{E}$ and $c\vec{B}$, i.e., $(c\rho_e, \vec{j}_e)$ and $(\rho_m/c, \vec{j}_m/c^2)$ – and so also the whole formalism – is however fully* ***conventional****.*

*The discrete transformation (5.85) is equivalent to $\varpi_{EM} : (\vec{E}, c\vec{B}) \to (c\vec{B}, -\vec{E})$. Since $\varpi_{EM}^2 = -\mathbb{1}$ and $\varpi_{EM}^4 = \mathbb{1}$, $\varpi_{EM}$ is equivalent to a 90°-rotation. In fact, one may define even a continuous duality rotation*

$$\varpi_{EM}(\vartheta) : \begin{bmatrix} \vec{E} \\ c\vec{B} \end{bmatrix} \to \begin{bmatrix} \vec{E}' \\ c\vec{B}' \end{bmatrix} = \begin{bmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{bmatrix} \begin{bmatrix} \vec{E} \\ c\vec{B} \end{bmatrix} \tag{5.86}$$

*and correspondingly for electric and magnetic charges and currents. The statement that there are no magnetic monopoles is then equivalent to stating that, using this "rotation," the variables $\vec{E}$ and $c\vec{B}$ (i.e., $F_{\mu\nu}$) may always be chosen so that $\rho_m = 0 = \vec{j}_m$, so that $\mathbf{F} = \mathbf{d}\wedge\mathbf{A}$, i.e., $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ – simultaneously in the whole universe and for all particles in Nature.*

#### The standard electrodynamics

In agreement with experiments, we set $\rho_m = 0 = \vec{j}_m$, so that the relations (5.15) and (5.73b) hold, as does the so-called Bianchi identity, as a consequence of the now applicable definition (5.73b),

$$\varepsilon^{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma} = 0, \tag{5.87}$$

and instead of equation (5.79); equations (5.74)–(5.78) remain unchanged.

In classical electrodynamics, one primarily uses the electromagnetic field $F_{\mu\nu}$, i.e., $\vec{E}$ and $\vec{B}$, and the potentials are secondary. However, in the non-relativistic formulation of the interaction (5.19) of the electromagnetic field with substance in quantum theory, the potentials had already been proved to be the fundamental quantities. Besides, the assumption that the electromagnetic field is *defined* in relations (5.73b) makes the relation (5.87) – and then also the laws (5.72b) – a trivial consequence. Thus, in electrodynamics expressed in terms of the 4-vector potential $A_\mu$, the dynamics reduces to the equation (5.78):

$$\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = \partial_\mu\partial^\mu A^\nu - \partial^\nu(\partial_\mu A^\mu) = \frac{1}{4\pi\epsilon_0}\frac{4\pi}{c}j_e^\nu. \tag{5.88}$$

The number of independent degrees of freedom in the electromagnetic field is thereby reduced from six in the rank-2 tensor $F_{\mu\nu}$ (the components of electric and magnetic field) to four in the 4-vector $A_\mu$.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

However, the 4-vector potential, $A_\mu$, is well known not to be unambiguously determined, as we are free to change

$$A_\mu \quad \rightarrow \quad A'_\mu = A_\mu - c\partial_\mu \lambda. \tag{5.89}$$

This is precisely the gauge transformation of the scalar and vector potential (5.14a), as it was derived in Section 5.1. The physical meaning of the transformation (5.89) may be seen from the Fourier transform:

$$A_\mu \quad \rightarrow \quad A'_\mu = A_\mu - c\partial_\mu \lambda \qquad \xrightarrow{\mathscr{F}} \qquad \widetilde{A}'_\mu = \widetilde{A}_\mu + ick_\mu \widetilde{\lambda}, \tag{5.90}$$

where $k_\mu := p_\mu/\hbar$ is the wave 4-vector of electromagnetic radiation. The component of the 4-vector potential in the direction of motion (in the 4-dimensional spacetime!) of the electromagnetic beam is arbitrary, and may be cancelled by a judicious choice of the gauge function $\lambda$. In that sense one frequently imposes the Lorenz gauge[14]:

$$\partial_\mu A^\mu = 0 \qquad \leftrightarrow \qquad k_\mu \widetilde{A}^\mu = 0. \tag{5.91}$$

Notice that this gauge is Lorentz-invariant. Using it, the dynamical part of the Maxwell equations (5.88), simplifies to

$$\Box A^\mu = \frac{1}{4\pi\epsilon_0}\frac{4\pi}{c} j_e^\mu, \tag{5.92}$$

which is the wave-equation for the gauge potentials $A^\mu(\mathrm{x})$, with the *sources* $j_e^\mu$.

The gauge (5.91) reduces the number of degrees of freedom in the electromagnetic field (which is determined by the relation (5.73b) in terms of the 4-vector potential) from four to three. But, that's not all: the FitzGerald–Lorentz length contraction applies to all physical quantities, and so also to the components of the 4-vector potential. Since, in vacuum, the changes in the electromagnetic field propagate at the speed of light, it follows that the longitudinal component of the 4-vector potential $A_\mu(\mathrm{x})$ equals zero, that is, its Fourier transform satisfies $\vec{k}\cdot\vec{\widetilde{A}} = 0$. The inverse transformation then gives $\vec{\nabla}\cdot\vec{A} = 0$, the so-called Coulomb gauge. The combination of the Lorenz and the Coulomb gauge produces $\dot{A}_0 = 0$, so that the temporal component of the 4-vector gauge potential is an arbitrary constant.

This reduction of the number of degrees of freedom from three to two cannot be described in a Lorentz-invariant way, so there are essentially two different approaches:

1. in addition to the Lorentz-invariant gauge, impose another gauge – such as the Coulomb gauge $\vec{\nabla}\cdot\vec{A} = 0$, which explicitly violates Lorentz symmetry, or
2. leave $A_\mu$ "ungauged" and having more than two degrees of freedom. Subsequently, systematically track and subtract the contributions of the nonphysical degrees of freedom in the 4-vector $A_\mu$.

In the absence of free carriers of electric charge, $j_e^\mu = 0$. The equation (5.92) then becomes

$$\Box A^\mu = 0, \tag{5.93}$$

which is the d'Alembert equation, i.e., the Klein–Gordon equation with $m_\gamma = 0$. The solutions are found in the form

$$A^\mu(\mathrm{x}) = a\, e^{-(i/\hbar)\mathrm{p}\cdot\mathrm{x}}\epsilon^\mu(\mathrm{p}), \qquad \begin{cases} p_\mu p^\mu = 0 & \Rightarrow \quad E = |\vec{p}|c, \\ p_\mu \epsilon^\mu = 0 & \rightarrow \quad \epsilon^0 = 0 = \vec{p}\cdot\vec{\epsilon}, \end{cases} \tag{5.94}$$

[14] This *gauge* (as in "condition," of "specification") bears the name of Ludvig Valentin Lorenz (1829–91), not of Hendrik Antoon Lorentz (1853–1928) after whom the Lorentz transformations, (FitzGerald–)Lorentz length contraction, and Lorentz group were named.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

where in the second row we see the joint effect of the (Lorentz-invariant) Lorenz and (Lorentz-violating) Coulomb gauge, and where $a$ is the photon amplitude.

In quantum theory, $A^\mu(\mathrm{x})$ could serve as the wave-function of the photon: the 4-vector potential that has two physical degrees of freedom, which are transversal to the direction of the photon's propagation. In a Cartesian coordinate system where the photon moves along the $(x^3 = z)$-axis, the two transversal polarizations are

$$\vec{\epsilon}_1 = (1,0,0) \qquad \text{and} \quad \vec{\epsilon}_2 = (0,1,0), \tag{5.95}$$

and

$$\vec{\epsilon}_+ = \tfrac{1}{\sqrt{2}}(\vec{\epsilon}_1 + i\vec{\epsilon}_2) \quad \text{and} \quad \vec{\epsilon}_- = \tfrac{1}{\sqrt{2}}(\vec{\epsilon}_1 - i\vec{\epsilon}_2) \tag{5.96}$$

are the so-called right- and left-circular polarizations, the eigenvectors of the rotation generator, $J_3$, with the eigenvalues $\pm 1$, respectively.

### 5.2.3 The magnetic monopole sneaks in

The immediate interpretation of Conclusion 5.3 on p. 184, notwithstanding, Paul Dirac found in 1931:

1. There does exist a way to include magnetic monopole charges and currents into the standard electrodynamics, i.e., the physical system described by the equations

$$\partial_\mu F^{\mu\nu} = \frac{1}{4\pi\epsilon_0}\frac{4\pi}{c} j_e^\nu, \qquad \tfrac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\partial_\mu F_{\nu\rho} = 0, \qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{5.97}$$

2. The quantum nature of Nature forces the magnetic and the electric charges to satisfy a mutual, so-called Dirac (dual charge), quantization law:

$$q_e\, q_m = 2\pi\hbar\, n, \qquad n \in \mathbb{Z}. \tag{5.98}$$

From here,

$$\alpha_e := \frac{1}{4\pi\epsilon_0}\frac{e^2}{\hbar c} \approx \frac{1}{137} \qquad \Rightarrow \qquad \alpha_m := \frac{1}{4\pi\mu_0}\frac{g^2}{\hbar c} = \frac{n^2}{4}\frac{4\pi\epsilon_0\hbar c}{e^2} \approx \frac{137}{4}n^2, \tag{5.99}$$

so that the interaction intensity with magnetic monopole charges and their currents must be very large ($\frac{\alpha_m}{\alpha_e} \approx 4,690\, n^2$), reciprocally to the relatively weak interaction with (electric) monopole charges and their currents, $\alpha_e \approx 1/137$.

#### The magnetic monopole gauge potential

Dirac's quasi-realistic model of a magnetic monopole stems from the very well known fact about magnets, that the magnetic field is strongest near the ends of a magnetic (physical) dipole and weakest near its middle. Take one such magnet – a cylindrical solenoid, for example – and affix the coordinate origin to the "north" pole of the magnet, squeeze the cross-section of the solenoid and stretch it so that the "south" pole is pulled out towards $z \to -\infty$. In the limit when the cross-section of the solenoid is negligible and the "south" pole is infinitely far, the magnetic field of such a magnet is spherically symmetric and has a source (the "north" pole) at the coordinate origin, with the "south" pole nowhere in sight.

This thought-construction evidently shows that part of the space (the negative $z$-semi-axis) is physically inaccessible: Every test-magnet detects a spherically symmetric (Coulomb-esque) magnetic field $\vec{B} \propto q_m\vec{r}/r^3$ in all of space around the coordinate origin – except along the negative $z$-semi-axis, where the test magnet cannot be placed as that is where the infinitely long and infinitely thin solenoid is. This "forbidden zone" is called the Dirac string.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

Dirac showed that the vector potential [296]

$$\vec{A}(\vec{r}) \;:\quad \text{so that} \qquad \vec{\nabla}\times\vec{A}=\vec{B}=\frac{q_m}{4\pi}\frac{\vec{r}}{r^3} \tag{5.100}$$

must be singular, as a function of the position $\vec{r}$, along *some* line (the Dirac string) that begins at the coordinate origin and extends out to infinity – which is the location of the infinitely thin Dirac solenoid. However, in 1975, T. T. Wu and C. N. Yang showed that there is no need to exclude this line from the physically accessible space – paying the price in accepting that the vector potential $\vec{A}$ then cannot be an unambiguously specified (vector) function. However, since the vector potential is not directly measurable, this ambiguity (non-single valuedness) has no physically measurable repercussion.

Indeed, define [536, 210]

$$\vec{A}_N=\frac{q_m}{4\pi}\frac{x\,\hat{\mathrm{e}}_y-y\,\hat{\mathrm{e}}_x}{r(z+r)}, \qquad \vec{A}_S=\frac{q_m}{4\pi}\frac{x\,\hat{\mathrm{e}}_y-y\,\hat{\mathrm{e}}_x}{r(z-r)}, \tag{5.101a}$$

$$=-\frac{q_m}{4\pi}\frac{\cos(\theta)-1}{r\sin(\theta)}\,\hat{\mathrm{e}}_\phi, \qquad =-\frac{q_m}{4\pi}\frac{\cos(\theta)+1}{r\sin(\theta)}\,\hat{\mathrm{e}}_\phi, \tag{5.101b}$$

and notice that the function $\vec{A}_N$ is well defined everywhere except along the ("southern") $z$-semi-axis, while the function $\vec{A}_S$ is well defined everywhere except along the ("northern") $z$-semi-axis. Also, define

$$\vec{B}_N:=\vec{\nabla}\times\vec{A}_N=\frac{q_m}{4\pi}\frac{\vec{r}}{r^3}, \qquad \text{and} \qquad \vec{B}_S:=\vec{\nabla}\times\vec{A}_S=\frac{q_m}{4\pi}\frac{\vec{r}}{r^3}. \tag{5.102}$$

(except where $x=0=y$ and $z\leqslant 0$) (except where $x=0=y$ and $z\geqslant 0$)

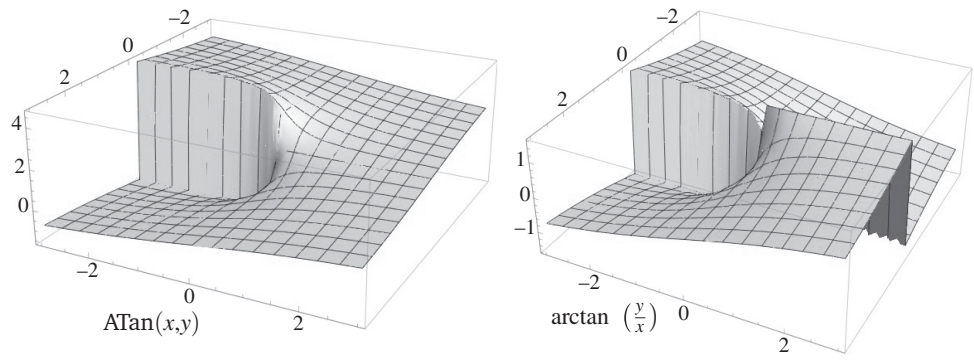
Since $\vec{B}_N$ and $\vec{B}_S$ perfectly coincide as functions everywhere where both are defined, the "true" magnetic field $\vec{B}$ is defined to be equal to $\vec{B}_N$ or $\vec{B}_S$, using that "auxiliary" magnetic field function that is well-defined in the region of interest.[15]

Since

$$\vec{A}_S-\vec{A}_N=2\frac{q_m}{4\pi}\left(\frac{y\hat{\mathrm{e}}_x-x\hat{\mathrm{e}}_y}{x^2+y^2}\right)=-2\frac{q_m}{4\pi}\vec{\nabla}\big[\,\mathrm{ATan}(x,y)\big], \tag{5.103}$$

where

$$\mathrm{ATan}(x,y):=\begin{cases}\arctan(y/x) & \text{for}\quad x\geqslant 0,\\ \pi+\arctan(y/x) & \text{for}\quad x\leqslant 0,\end{cases} \tag{5.104a}$$




(5.104b)

[15] This is the same "trick" that cartographers use when they carve up the map of the Earth's globe (which cannot be depicted accurately on a single flat sheet of paper) into a sufficiently large number of sufficiently small maps, each of which depicts adequately a sufficiently small region of the Earth surface. These maps are then bound into an atlas where "adjacent" maps overlap sufficiently to provide the traveller with connecting information along any – of course continuous – voyage.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

it follows that $\vec{A}_N$ and $\vec{A}_S$ differ by a gauge transformation (5.89) with the gauge parameter $\lambda(\mathrm{x}) = 2q_m\,\mathrm{ATan}(x,y)$. Since the potentials $\vec{A}_N$ and $\vec{A}_S$ are not themselves measurable, but provide the same (measurable) magnetic field, the gauge transformation

$$\vec{A}_N \to \vec{A}_S = \vec{A}_N + \vec{\nabla}\lambda_{NS}, \quad \lambda_{NS}(\mathrm{x}) = -2\frac{q_m}{4\pi}\,\mathrm{ATan}(x,y) \tag{5.105}$$

is then really a symmetry of the physical system.

#### Dirac's dual quantization of charges

As relations (5.14a)–(5.14b) show, the gauge transformation (5.105) induces the change in the phase of the electron wave-function:

$$\Psi(\mathrm{x}) \to \Psi'(\mathrm{x}) = e^{iq_e\lambda_{NS}(\mathrm{x})/\hbar}\Psi(\mathrm{x}). \tag{5.106}$$

As the value of the function $\mathrm{ATan}(x,y)$ is the azimuthal angle $\phi \simeq \phi + 2\pi$, the relation (5.105) yields

$$\exp\Big\{iq_e\lambda_{NS}(\mathrm{x})/\hbar\Big\} = \exp\Big\{-i\,\frac{q_m q_e}{2\pi\hbar}\,\phi\Big\}. \tag{5.107}$$

No gauge transformation – and so not this one – can *change* the single-valuedness of the wave-function, which of course is chosen single-valued to begin with. Thus the phase (5.107) also must be a single-valued function of $\phi$, and $\frac{q_m q_e}{2\pi\hbar}$ must be an integer:

$$\frac{q_e q_m}{2\pi\hbar} \stackrel{!}{=} n \in \mathbb{Z}, \qquad \text{i.e.,} \qquad q_m \stackrel{!}{=} n\Big(\frac{2\pi\hbar}{q_e}\Big), \tag{5.108}$$

which is called the Dirac (dual charge) quantization of the magnetic charge, and where $(2\pi\hbar/q_e)$ is the elementary (unit) amount of magnetic charge.

It will prove useful to rewrite this argument by direct integration of the relation (5.105):

$$\int_1^2 \mathrm{d}\vec{r}\cdot\vec{A}_N = \int_1^2 \mathrm{d}\vec{r}\cdot\vec{A}_S + \Big[\int_1^2 \mathrm{d}\vec{r}\cdot\vec{\nabla}\lambda_{NS} = \lambda_{NS}(\vec{r}_2) - \lambda_{NS}(\vec{r}_1)\Big], \tag{5.109}$$

that is,

$$\int_1^2 \mathrm{d}\vec{r}\cdot\vec{A}_N - \int_1^2 \mathrm{d}\vec{r}\cdot\vec{A}_S = \lambda_{NS}(\vec{r}_2) - \lambda_{NS}(\vec{r}_1). \tag{5.110}$$

Dirac's quantization of the magnetic charge thus stems from the requirement that $iq_e\int_C \mathrm{d}\vec{r}\cdot\vec{A}$ may depend on the choice of the concrete line integration contour only up to an integral multiple of $2\pi$:

$$e^{iq_e\int_{C_1}\mathrm{d}\vec{r}\cdot\vec{A} - q_e\int_{C_2}\mathrm{d}\vec{r}\cdot\vec{A}} = e^{iq_e\oint_{C_1-C_2}\mathrm{d}\vec{r}\cdot\vec{A}} \stackrel{!}{=} e^{2\pi i n} = 1, \tag{5.111}$$

where $(C_1 - C_2)$ is a closed contour since $C_1$ and $C_2$ have the same end-points: $\partial C_1 = \partial C_2$.[16] Using Stokes' theorem, $\oint_C \mathrm{d}\vec{r}\cdot\vec{A} = \int_S \mathrm{d}^2\vec{\sigma}\cdot(\vec{\nabla}\times\vec{A})$ where $S$ is some surface bounded by the contour $C$, i.e., $C = \partial S$, and the definition of the magnetic field, $\vec{B} := (\vec{\nabla}\times\vec{A})$, we have that

$$2\pi\, n \stackrel{!}{=} q_e\int_{C_1}\mathrm{d}\vec{r}\cdot\vec{A} - q_e\int_{C_2}\mathrm{d}\vec{r}\cdot\vec{A} = q_e\oint_{(C_1-C_2)=\partial S}\mathrm{d}\vec{r}\cdot\vec{A} = q_e\int_S \mathrm{d}^2\vec{\sigma}\cdot\vec{B}. \tag{5.112}$$

Applied to the magnetic field of a (hypothetical) magnetic monopole charge, this condition produces the quantization (5.108). However, the same condition also represents a reason for the existence of the so-called Aharonov–Bohm effect [☞ textbooks [407, 471, 480, 472, 29, 324],

[16] For any space $\mathscr{X}$, the symbol $\partial\mathscr{X}$ denotes the "boundary of $\mathscr{X}$."

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

for example], which is experimentally verified, and which should therefore be called the "Dirac–Aharonov–Bohm effect."

Today, several additional, alternative arguments are known to infer the same mutual quantization. One of them was published by Alfred S. Goldhaber in 1965. The magnetic field (5.100) of a magnetic monopole exerts a force upon a particle of electric charge $q_e$ that passes through the magnetic field at the velocity $\vec{v}$. This so-called Lorentz force,

$$\vec{F}_L = q_e\, \vec{v} \times \vec{B}, \tag{5.113}$$

is perpendicular to the plane containing $\vec{v}$ and $\vec{B}$. Select a coordinate system so that $\vec{v} = v\,\hat{e}_z$, where $\theta := \measuredangle(\vec{r},\vec{v}) = \measuredangle(\vec{B},\vec{v})$ since for the magnetic monopole $\vec{B} \propto \vec{r}$ [☞ relation (5.100)]. The distance $b := |r\sin(\theta)|$ is called the "impact parameter," just as in the set-up for the collision of two marbles, in Example 3.2 on p. 111. Select the $x$-axis to be in the direction of this parameter and $\vec{B}$ is in the $(x,z)$-plane. For sufficiently large values of $b$, the deflection (in the direction of the $y$-axis) from the trajectory (along the $z$-axis) will be small enough to be accurately estimated by the integral

$$(\triangle\vec{p})_y \approx \int_{-\infty}^{+\infty} \mathrm{d}t\ (\vec{F}_L)_y = \frac{q_e v q_m b}{4\pi} \int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{(b^2+v^2t^2)^{3/2}} = \frac{q_e\, q_m}{2\pi b}, \tag{5.114}$$

so that

$$(\triangle\vec{L})_z = b(\triangle\vec{p})_y = \frac{q_e q_m}{2\pi}. \tag{5.115}$$

It remains to conclude – because of the quantum nature of Nature – that the change in the angular momentum must be an integral multiple of $\hbar$. This immediately reproduces equation (5.108).

Finally, let us also mention the fact that the electromagnetic field has a linear momentum density $\epsilon_0\vec{E}\times\vec{B}$. For the field near point-like electric and magnetic charges that are separated by the vector $\vec{R}$, it may be shown that the total (integrated) linear momentum of the total field vanishes, whereby the total (integrated) angular momentum is independent of the choice of the coordinate origin and has the value [☞ [296]; this result was published by J. J. Thomson, in 1904]

$$\vec{L}_{EM} = \frac{q_e q_m}{4\pi}\frac{\vec{R}}{R}. \tag{5.116}$$

The quantization of this angular momentum in (integral) units of $\hbar$ also indicates a quantization of the magnetic charge in units that are inversely proportional to the elementary electric charge, but gives a value that is twice as large as the result (5.108). That is, the previous two arguments produce a stricter result. One could have obtained this as early as 1904 from equation (5.116), but only by adopting the quantization of angular momentum in *half-integral* units of $\hbar$ – thus foreshadowing spin-$\frac{1}{2}$ particles and systems. At the time, no one thought of it.

### 5.2.4 Exercises for Section 5.2

✎ ***5.2.1*** *Using the stated definitions of $J_i, K_j$ and the ensuing relations (5.45), prove equation (5.46).*

✎ ***5.2.2*** *Using the relation (5.45) with the choice $\varphi^i = 0$, $\beta^2 = 0 = \beta^3$ and $\beta^1 = \beta$, prove relation (5.48) by expanding the exponential function, then re-summing the result after using the relation (5.32).*

✎ ***5.2.3*** *Prove the equivalence of results (5.34) and (5.69), as well as that $\mathscr{H}_D = \overline{\Psi} H_D \Psi$.*

✎ ***5.2.4*** *Using the relation (5.74)–(5.75) and (3.13b), derive equations (5.77).*

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

## 5.3 Quantum electrodynamics with leptons

The description of electrodynamics in the previous section is classical. Quantum computations are consistently derived from quantum field theory – of photons and leptons – and this derivation is outside the scope of this book. Instead, following Ref. [243] and the introductory material in Section 3.3, we will consider several examples of computations with Feynman diagrams that depict interactions of charged leptons and photons.

### 5.3.1 Quantum electrodynamics calculation

We have already seen Feynman diagrams that depict electromagnetic processes: $O(e^4)$ contributions to the $e^- p^+$ scattering are depicted by the diagrams (4.39), and the two-photon $e^- e^+$ annihilation is depicted by diagram (4.48). Modeled on Section 3.3.4, we first assign a mathematical expression to every graphical element, and by adapting Procedure 3.1 on p. 116, we will compute the amplitude $\mathfrak{M}$, which we will then insert into the formulae (3.112) and (3.114) for decays and scattering, respectively.

Although we will not derive the Feynman rules for electrodynamics from the Lagrangian, we present this Lagrangian density. By combining the results (5.76) and (5.68a), changing

$$\partial_\mu \to D_\mu := \partial_\mu + \tfrac{i}{\hbar c} A_\mu Q \qquad \text{so that} \qquad D'_\mu \big(e^{i\varphi(\mathrm{x})}\Psi(\mathrm{x})\big) = e^{i\varphi(\mathrm{x})}\big(D_\mu \Psi(\mathrm{x})\big), \tag{5.117}$$

in accord with the definitions (5.13) and (5.73a), and where $Q\Psi = q_\Psi \Psi$ produces the electric charge of the particle represented by $\Psi$, we have

$$\begin{aligned}\mathscr{L}_{QED} &= \overline{\Psi}(\mathrm{x})\left[i\hbar c \not{D} - mc^2\right]\Psi(\mathrm{x}) - \tfrac{4\pi\epsilon_0}{4} F_{\mu\nu}F^{\mu\nu} \\ &= \overline{\Psi}(\mathrm{x})\left[\boldsymbol{\gamma}^\mu\big(\hbar c\, i\partial_\mu - q_\Psi A_\mu\big) - mc^2\right]\Psi(\mathrm{x}) \\ &\quad - \tfrac{4\pi\epsilon_0}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)\eta^{\mu\rho}\eta^{\nu\sigma}(\partial_\rho A_\sigma - \partial_\sigma A_\rho).\end{aligned} \tag{5.118}$$

By construction, this Lagrangian is invariant under the gauge transformation

$$A'_\mu(\mathrm{x}) = A_\mu(\mathrm{x}) - c\partial_\mu \varphi(\mathrm{x}) \qquad \text{and} \qquad \Psi'(\mathrm{x}) = e^{i\varphi Q/\hbar}\,\Psi(\mathrm{x}). \tag{5.119}$$

---

**Digression 5.9** The equation of motion for $A_\mu(\mathrm{x})$ is obtained by varying either the Lagrangian density (5.118) or the Hamilton action $\int \mathrm{d}^4\mathrm{x}\, \mathscr{L}_{QED}$ with respect to $A_\mu(\mathrm{x})$. Using so-called functional derivative generalization of partial derivatives:

$$\frac{\delta}{\delta A_\rho(\mathrm{y})}\mathscr{F}\big(A_\mu(\mathrm{x}), (\partial_\mu A_\nu(\mathrm{x})), \dots\big) := \delta^4(\mathrm{x}-\mathrm{y})\,\frac{\partial}{\partial A_\rho(\mathrm{x})}\mathscr{F}\big(A_\mu(\mathrm{x}), (\partial_\mu A_\nu(\mathrm{x})), \dots\big), \tag{5.120a}$$

$$\frac{\delta}{\delta(\partial_\rho A_\sigma(\mathrm{y}))}\mathscr{F}\big(A_\mu(\mathrm{x}), (\partial_\mu A_\nu(\mathrm{x})), \dots\big) = \delta^4(\mathrm{x}-\mathrm{y})\,\frac{\partial}{\partial(\partial_\rho A_\sigma(\mathrm{y}))}\mathscr{F}\big(A_\mu(\mathrm{x}), (\partial_\mu A_\nu(\mathrm{x})), \dots\big), \tag{5.120b}$$

we obtain the general result

$$\begin{aligned}\frac{\delta}{\delta A_\rho(\mathrm{x})}\int \mathrm{d}^4\mathrm{y}\, \mathscr{F}\big(A_\mu(\mathrm{y}), (\partial_\mu A_\nu(\mathrm{y}))\big) &= \int \mathrm{d}^4\mathrm{y}\, \frac{\delta A_\sigma(\mathrm{y})}{\delta A_\rho(\mathrm{x})}\frac{\partial}{\partial A_\sigma(\mathrm{y})}\mathscr{F}\big(A_\mu(\mathrm{y}), (\partial_\mu A_\nu(\mathrm{y}))\big) \\ &= \int \mathrm{d}^4\mathrm{y}\, \delta^4(\mathrm{x}-\mathrm{y})\delta^\sigma_\rho \frac{\partial}{\partial A_\sigma(\mathrm{y})}\mathscr{F}\big(A_\mu(\mathrm{y}), (\partial_\mu A_\nu(\mathrm{y}))\big) = \frac{\partial}{\partial A_\rho(\mathrm{x})}\mathscr{F}\big(A_\mu(\mathrm{x}), (\partial_\mu A_\nu(\mathrm{x}))\big).\end{aligned} \tag{5.120c}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

Using

$$\frac{\partial}{\partial A_\rho(\mathrm{x})} A_\mu(\mathrm{x}) = \delta^\rho_\mu, \qquad \frac{\partial}{\partial A_\rho(\mathrm{x})} \big(\partial_\mu A_\nu(\mathrm{x})\big) = 0, \tag{5.120d}$$

$$\frac{\partial}{\partial(\partial_\rho A_\sigma(\mathrm{x}))} A_\mu(\mathrm{x}) = 0, \qquad \frac{\partial}{\partial(\partial_\rho A_\sigma(\mathrm{x}))} \big(\partial_\mu A_\nu(\mathrm{x})\big) = \delta^\rho_\mu \delta^\sigma_\nu, \tag{5.120e}$$

where we need not write the arguments "(x)," we obtain

$$\partial_\mu \frac{\partial \mathscr{L}_{QED}}{\partial(\partial_\mu A_\nu)} = \frac{\partial \mathscr{L}_{QED}}{\partial A_\nu} \qquad \Rightarrow \qquad \partial_\mu F^{\mu\nu} = \frac{q_\Psi}{4\pi\epsilon_0} \overline{\Psi} \boldsymbol{\gamma}^\nu \Psi. \tag{5.120f}$$

Comparing result (5.120f) with equation (5.78) identifies

$$j^\mu_e := \frac{q_\Psi c}{4\pi} \overline{\Psi} \boldsymbol{\gamma}^\mu \Psi \tag{5.120g}$$

as the 4-vector of the electric current density. The combined Lagrangian density (5.118) shows that, while the dynamics of photons alone may be described in terms of the $F_{\mu\nu}$ field, i.e., $\vec{E}$ and $\vec{B}$, the Lagrangian description of the interaction with charged particles requires the use of the gauge 4-vector potential $A_\mu$ – although the derived equations of motion (5.120f) and the obvious (Bianchi) consequence (5.87) *may* be expressed fully in terms of the $\vec{E}$ and $\vec{B}$ fields.

**Digression 5.10** Varying the Lagrangian density $\mathscr{L}_{\mathrm{QED}}$, as in equation (5.118), with respect to $A_\mu$ and $\overline{\Psi}$ (from the left), we obtain the complementary and coupled system of Euler–Lagrange equations of motion:

$$\partial_\mu F^{\mu\nu} = \frac{q_\Psi}{4\pi\epsilon_0} \overline{\Psi} \boldsymbol{\gamma}^\nu \Psi, \qquad \Big[ i\hbar c\, \boldsymbol{\gamma}^\mu \partial_\mu - mc^2 \mathbb{1} \Big] \Psi = q_\Psi A_\mu \boldsymbol{\gamma}^\mu \Psi. \tag{5.121}$$

The procedure given in Digression 5.9 is equally applicable to interactions of arbitrary charged particles with photons: for a particle of a spin other than $\frac{1}{2}$, the Dirac Lagrangian density must be replaced by a corresponding Lagrangian density but where the "gauge covariant derivatives" $\partial_\mu \to D_\mu$ (5.117) are used. As an introduction and because of immediate application, the formulae will be written for a lepton/antilepton, i.e., electron/positron. The computations, however, are easy to adapt for other charged spin-$\frac{1}{2}$ particles – one should only substitute the appropriate charges and masses. Also, it should not be too hard to also adapt the computations to include charged particles without spin. This is usually called "scalar electrodynamics" in the literature, but we leave this aside.

Because of the difference in units and numerical simplification, the notation

$$g_e := \sqrt{4\pi\,\alpha_e} = \frac{|e|}{\sqrt{\epsilon_0 \hbar c}} \qquad (= |e|\sqrt{4\pi/\hbar c}, \text{ in Gauss's units}) \tag{5.122}$$

is useful. On one hand, $g_e$ gives a dimensionless measure of the interaction strength; on the other, many electrodynamics computations may then be relatively easily adapted for weak nuclear and chromodynamics computations by changing $g_e \to g_w$ and $g_e \to g_c$, respectively, and inserting a few additional factors [☞ Chapter 6].

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

The practical use of most concrete models in quantum field theory reduces to the prescription (see also Procedure 11.1 on p. 416 and discussion in Section 11.2.4):

***Procedure 5.1*** *Start with a concrete model defined within classical field theory.*

1. *For any considered process, list the possible sub-processes, as discussed in Section 3.3.1 and in the form of a sequence of Feynman diagrams partially ordered:*
   (a) *by the number of closed loops [☞ Comment 3.5 on p. 122],*
   (b) *by the powers of a characteristic interaction parameter,*
   (c) *by the powers of $\hbar$.*
2. *Compute the amplitude $\mathfrak{M}_i$ for each (sub)process, as described by the specific Feynman calculus rules of the model; see for example Procedures 5.2 on p. 193 and 6.1 on p. 232, below.*
3. *Add the amplitudes, with a negative relative sign between sub-processes that differ only by the exchange of two identical fermions.*
4. *Compute the corresponding scattering cross-section or decay constant as discussed in Section 3.3.3, and illustrated there for a simple toy-model.*

The specific Feynman calculus rules mentioned in step 2 above are derived from the same classical action and rely on the correspondences discussed in Section 3.3.1 and in particular the listing on page 106. As stated there, that task is deferred to proper field-theory texts [64, 63, 48, 257, 307, 221, 159, 422, 423, 538, 250, 389, 243, 45, 580, 238, 241, 239, 240].

For the particular case at hand, the model describing the interaction of electrically charged spin-$\frac{1}{2}$ fermions (such as electrons) and the electromagnetic field, the classical model is described by the Lagrangian (5.118), and the specific Feynman calculus rules are as follows:

***Procedure 5.2*** *The contribution to the amplitude $\mathfrak{M}$ corresponding to a given Feynman diagram for an electrodynamics process with electrons and positrons is computed following the algorithm [☞ textbooks [445, 425, 586] for a derivation]:*

1. ***Notation***
   (a) *Energy–momentum: Denote incoming and outgoing 4-momenta by $p_1, p_2, \ldots$, and the spins by $s_1, s_2, \ldots$ Denote the "internal" 4-momenta (assigned to lines that connect two vertices inside the diagram) by $q_1, q_2, \ldots$*
   (b) *Orientation: For a spin-$\frac{1}{2}$ particle, orient the line in the 4-momentum direction, oppositely for antiparticles. Orient external photon lines in the direction of time (herein, upward). Orient the internal photon lines arbitrarily, but use the so-chosen orientation consistently.*
   (c) *Polarization: Assign every external line the polarization factor:*

| | | | | |
|---|---|---|---|---|
| *Spin-$\frac{1}{2}$ particle* | *incoming* |  | $u^s$ | $s =$ *spin projection* $= \uparrow, \downarrow$ |
| | *outgoing* | | $\overline{u}_s$ | |
| *Spin-$\frac{1}{2}$ antiparticle* | *incoming* |  | $\overline{v}_s$ | ($\simeq$ *spin-$\frac{1}{2}$ particle, travels backwards in time*) |
| | *outgoing* | | $v^s$ | |
| *Photon* | *incoming* |  | $\epsilon^\mu$ | $\epsilon^\mu p_\mu = 0$ *and* $\epsilon^0 = 0$ |
| | *outgoing* | | $\epsilon^{\mu *}$ | |

(5.123)

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

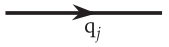
2. ***Vertices*** *To each vertex assign the factor*


$$\rightarrow \quad -ig_e\boldsymbol{\gamma}^\mu. \tag{5.124}$$

*Even without derivation, this factor clearly corresponds to the term* $-q_\Psi \overline{\Psi} \not{A} \Psi$ *in equation (5.118), and so represents the elementary interaction of the photon with the current of the charged particle that* $\Psi$ *represents.*

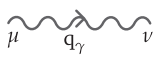
3. ***Propagators*** *To each internal line with the* $j$*th 4-momentum assign the factor:*

*spin-*$\frac{1}{2}$ *particle:*




$$\rightarrow \quad \frac{i}{\not{q}_j - m_j c} = i\frac{\not{q}_j + m_j c\mathbb{1}}{q_j^2 - m_j^2 c^2}, \tag{5.125}$$

*photon:*




$$\rightarrow \quad -i\frac{\eta_{\mu\nu}}{q_\gamma^2}. \tag{5.126}$$

*As internal lines depict virtual particles,* $\not{q}_j \neq m_j c$ *and* $q_\gamma^2 \neq 0$*, respectively [☞ Tables C.7 on p. 529 and C.8 on p. 529]. Up to multiplicative coefficients, these factors also stem from (5.118); these are Fourier transforms of the Green functions for the differential operators* $\not{\partial}$ *and* $D^{\mu\nu}$*, in* $\overline{\Psi}\not{\partial}\Psi := -\overline{\Psi}[i\hbar c\not{\partial} - mc^2]\Psi$ *and* $A_\mu D^{\mu\nu} A_\nu :\simeq -\frac{4\pi\epsilon_0}{4} F_{\mu\nu}F^{\mu\nu}$*, respectively, where "*$\simeq$*" denotes effective equality (equivalence) under the integral, after integration by parts and "*$:\simeq$*" defines the left-hand side by means of such an effective equality.*

**Digression 5.11** *Integration by parts is used rather often, so that, e.g.,*

$$\begin{aligned}\int \mathrm{d}^4\mathrm{x}\ (\partial_\mu A_\nu)(\partial^\mu A^\nu) &= \int \mathrm{d}^4\mathrm{x}\ \partial_\mu(A_\nu \partial^\mu A^\nu) - \int \mathrm{d}^4\mathrm{x}\ A_\nu(\partial_\mu \partial^\mu A^\nu) \\ &= \oint_{\mathscr{V}_{(\mu)}} (\mathrm{d}^3\mathrm{x})_\mu\ (A_\nu \partial^\mu A^\nu) - \int \mathrm{d}^4\mathrm{x}\ A_\nu(\partial_\mu \partial^\mu A^\nu),\end{aligned} \tag{5.127a}$$

*where* $\mathscr{V}_{(\mu)}$ *is a closed 3-dimensional hypersurface that bounds the 4-dimensional spacetime and* $(\mathrm{d}^3\mathrm{x})_\mu$ *is the volume element of* $\mathscr{V}_{(\mu)}$*. As the domain of 4-dimensional integrals is typically* ***all of*** *spacetime,* $\mathscr{V}_{(\mu)}$ *is a hypersurface "at infinity" where all fields are required to vanish, so the integrated term also vanishes. With this in mind, the relation (5.127a) is written as*

$$\int \mathrm{d}^4\mathrm{x}\ (\partial_\mu A_\nu)(\partial^\mu A^\nu) \simeq -\int \mathrm{d}^4\mathrm{x}\ A_\nu(\partial_\mu \partial^\mu A^\nu), \tag{5.127b}$$

*which defines the relation "*$\simeq$*," in this context, as "equality under spacetime integral up to integrated terms that are assumed to vanish," or "equivalence up to integrals of total derivatives."*

4. ***Energy–momentum conservation*** *To each vertex assign a factor* $(2\pi)^4\delta^4(\sum_j \mathrm{k}_j)$*, where* $\mathrm{k}_j$ *are 4-momenta that enter the vertex. 4-momenta that* ***leave*** *the vertex have a negative sign – except for external spin-*$\frac{1}{2}$ *antiparticles, since they are equivalent to particles that move backwards in time.*
5. ***Integration over 4-momenta*** *Internal lines correspond to virtual particles and their 4-momenta are unknown; these variables must be integrated:* $\int \frac{\mathrm{d}^4\mathrm{q}_j}{(2\pi)^4}$*.*

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

6. ***Reading off the amplitude*** *The foregoing procedure yields the result*

$$-i\,\mathfrak{M}\,(2\pi)^4\delta^4(\sum_j \mathrm{p}_j), \tag{5.128}$$

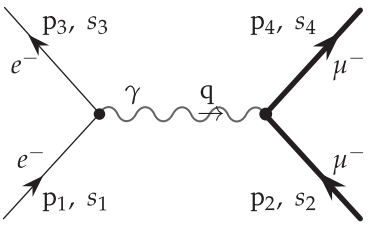
*where the factor $(2\pi)^4\delta^4(\sum_j \mathrm{p}_j)$ represents the 4-momentum conservation for the entire process, and where the amplitude (matrix element) $\mathfrak{M}$ is read off.*

7. ***Fermion loops*** *To each fermion loop (closed line) assign a factor $-1$. A mathematically rigorous derivation of this rule follows from Feynman's approach using path integrals, which is far beyond the scope of this book. See however Digression 2.4 on p. 52 and especially statement 4a therein; see also the booklet [166] for an intuitive albeit not entirely rigorous explanation, Ref. [434, Vol. 1, Appendix A] for a serious introduction, and Ref. [165] for the original reference.*
8. ***Antisymmetrization*** *Since the amplitude of the process must be antisymmetric in pairs of identical (external) fermions, the partial amplitudes that differ only in the exchange of two identical external fermions must have the relative sign $-1$.*

*As in Section 3.3.4, one draws all Feynman diagrams that contribute at the desired order in $g_e$, and then computes the (partial) amplitudes for each of the diagrams. The algebraic sum of these contributions yields the total amplitude, which is then inserted in formulae (3.112) and (3.114) for decays and scatterings, respectively.*

In the remaining part of Section 5.3, the contributions of the following 12 Feynman diagrams will be examined, where we follow the treatment in Refs. [243] [☞ also Refs. [64, 580, 241]]: Each of these diagrams depicts a separate contribution to some $O(g_e^2)$ process and, exceptionally, $O(g_e^4)$ for the last diagram. Processes are identified by the "external" particles, whereby diagram (a) in Figure 5.1, all by itself represents one process, while diagrams (b) and (c) in Figure 5.1 represent two contributions to the same process.

Denote the external lines so that incoming are bottom-left=1 and bottom-right=2, and outgoing are top-left=3 and top-right=4. So, e.g.,




(5.129)

depicts the elastic scattering of an electron and a muon via the exchange of a photon. In fact, the incoming (and so also the outgoing) pair of fermions in the diagrams in Figure 5.1 (a)–(d), p. 196, could be identified as any other pair of different spin-$\frac{1}{2}$ particles, including the electron–proton pair in the hydrogen atom. It is, however, important to keep in mind that the relativistic description in terms of the perturbative expansion in the degree of the interaction constant $g_e$ is appropriate for scatterings and for decays but not for bound states, the description of which is inherently non-perturbative in this sense.

To see this, note that the bound states of the hydrogen atom are determined by the Coulomb field, which results from summing over all possible exchange processes including one to infinitely many photons. The static electromagnetic field, known as the Coulomb field, may be identified with

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

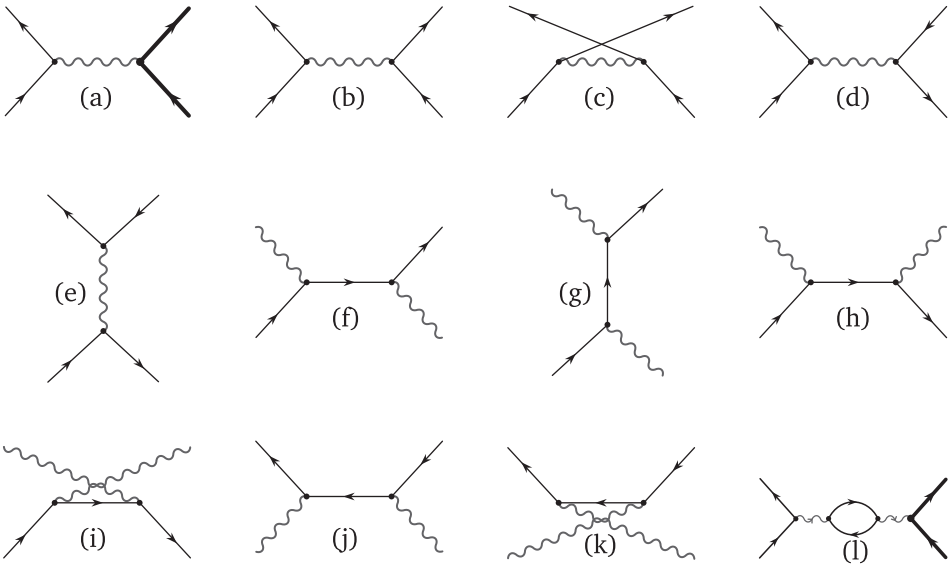



**Figure 5.1** The first 12 Feynman diagrams that depict the quantum-electrodynamical processes between spin-$\frac{1}{2}$ particles and the photon. The last diagram depicts one of the corrections for the process (a).

the (Bose) *condensation* of infinitely many photons,[17] and is a phenomenon that is inherently non-perturbative in the number of exchanged particles, and so inherently unreachable in the analysis using elementary processes depicted by Feynman diagrams.

In turn, in scatterings and decays, the exchange of a single photon produces the dominant contribution, while multi-particle exchanges produce ever smaller corrections: scatterings and decays are inherently perturbative in the number of exchanged particles.

#### Electron–muon scattering

Scatterings of the type $e^- + \mu^- \to e^- + \mu^-$, where the muon is a "target" that is significantly heavier than the "probe" (here, $e^-$), are called Mott scattering, after Sir Nevill Francis Mott. In the non-relativistic regime one obtains Rutherford scattering, named after Ernest Rutherford's experiment of bombarding a foil of gold with $\alpha$-particles. Reading off of the diagram in Figure 5.1, and following the Procedure 5.2 on p. 193, we get

$$\begin{aligned}
&\int \frac{\mathrm{d}^4\mathrm{q}}{(2\pi)^4}\,(2\pi)^4\delta^4(\mathrm{p}_1-\mathrm{p}_3-\mathrm{q})\,(2\pi)^4\delta^4(\mathrm{p}_2-\mathrm{p}_4+\mathrm{q})\\
&\qquad\times\left[\overline{u^{s_3}}_A(\mathrm{p}_3)(ig_e\,\gamma^{\mu A}{}_B)u^{s_1,B}(\mathrm{p}_1)\right]\left(\frac{-i\eta_{\mu\nu}}{\mathrm{q}^2}\right)\left[\overline{U^{s_4}}_C(\mathrm{p}_4)(ig_e\,\gamma^{\nu C}{}_D)U^{s_2,D}(\mathrm{p}_2)\right]\\
&=\frac{ig_e^2(2\pi)^4}{(\mathrm{p}_1-\mathrm{p}_3)^2}\,\delta^4(\mathrm{p}_2-\mathrm{p}_4+\mathrm{p}_1-\mathrm{p}_3)\left[\overline{u^{s_3}}_A(\mathrm{p}_3)\,\gamma^{\mu A}{}_B\,u^{s_1,B}(\mathrm{p}_1)\right]\left[\overline{U^{s_4}}_C(\mathrm{p}_4)\,\gamma_\mu{}^C{}_D\,U^{s_2,D}(\mathrm{p}_2)\right],
\end{aligned}\tag{5.130}$$

and comparison with the diagram (5.129) shows that $u^{s_1,b}(\mathrm{p}_1)$ represents the incoming electron, and $\overline{U^{s_4}}_C(\mathrm{p}_4)$ the outgoing muon, etc.

[17] And the other way round, photons are the quanta of the electromagnetic field in the sense that they are the smallest "packet" of a *change* in the electromagnetic field. These quanta – oscillations in the electromagnetic field – move at the speed of light; once established, the electrostatic or magnetostatic field does not move at all and extends through the whole available space.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

From there, using the expression (5.128), we get

$$\mathfrak{M}_{(a)} = -\frac{g_e^2}{(\mathrm{p}_1-\mathrm{p}_3)^2}\left[\overline{u^{s_3}}_A(\mathrm{p}_3)\,\gamma^{\mu\,A}{}_B\,u^{s_1,B}(\mathrm{p}_1)\right]\left[\overline{U^{s_4}}_C(\mathrm{p}_4)\,\gamma_\mu{}^C{}_D\,U^{s_2,D}(\mathrm{p}_2)\right]. \tag{5.131}$$

If the spins of the incoming and outgoing particles are known, the polarization spinors $u^{s_1,B}$, $U^{s_2,D}$, $\overline{u^{s_3}}_A$ and $\overline{U^{s_4}}_C$ are selected as given in equations (5.40), one computes the components of the 4-vectors $[\overline{u^{s_3}}\,\boldsymbol{\gamma}^\mu\,u^{s_1}]$ and $[\overline{U^{s_4}}\,\boldsymbol{\gamma}_\mu\,U^{s_2}]$, and then the sum of the products.

When in turn the spins of the interacting particles are not measured, and we are interested in the *inclusive effective cross-section* of the scattering, i.e., the *inclusive decay constant*, summing over all spins produces an important simplification. Indeed, the formulae (3.112) and (3.114) need $|\mathfrak{M}|^2 = \overline{\mathfrak{M}}\,\mathfrak{M}$. On the other hand, $\overline{\mathfrak{M}}_{(a)}$ contains a factor

$$\begin{aligned}\left[\overline{u}_A(\mathrm{p}_3)\,\gamma^{\mu\,A}{}_B\,u^B(\mathrm{p}_1)\right]^\dagger &= \left[u^\dagger(\mathrm{p}_3)\,\boldsymbol{\gamma}^0\,\boldsymbol{\gamma}^\mu\,u(\mathrm{p}_1)\right]^\dagger = \left[u^\dagger(\mathrm{p}_1)\,(\boldsymbol{\gamma}^\mu)^\dagger\,(\boldsymbol{\gamma}^0)^\dagger\,u(\mathrm{p}_3)\right]\\ &= \left[u^\dagger(\mathrm{p}_1)\mathbb{1}\,(\boldsymbol{\gamma}^\mu)^\dagger\,\boldsymbol{\gamma}^0\,u(\mathrm{p}_3)\right] = \left[u^\dagger(\mathrm{p}_1)\boldsymbol{\gamma}^0\,\boldsymbol{\gamma}^0\,(\boldsymbol{\gamma}^\mu)^\dagger\,\boldsymbol{\gamma}^0\,u(\mathrm{p}_3)\right]\\ &= \left[\overline{u}(\mathrm{p}_1)\,\overline{\boldsymbol{\gamma}}^\mu\,u(\mathrm{p}_3)\right], \qquad \overline{\boldsymbol{\gamma}}^\mu := \boldsymbol{\gamma}^0(\boldsymbol{\gamma}^\mu)^\dagger\boldsymbol{\gamma}^0,\end{aligned} \tag{5.132}$$

so that $|\mathfrak{M}_{(a)}|^2$ contains the factor

$$\left[\overline{u}_A(\mathrm{p}_3)\,\gamma^{\mu\,A}{}_B\,u^b(\mathrm{p}_1)\right]\left[\overline{u}_C(\mathrm{p}_1)\,\gamma^{\nu\,C}{}_D\,u^D(\mathrm{p}_3)\right]. \tag{5.133}$$

**Digression 5.12** The *physical* requirement (A.127) implies that

$$\overline{\boldsymbol{\gamma}}^\mu := \boldsymbol{\gamma}^0(\boldsymbol{\gamma}^\mu)^\dagger\boldsymbol{\gamma}^0 \overset{(A.127)}{=} \boldsymbol{\gamma}^\mu. \tag{5.134}$$

Finally, summing over spins permits using the relations (5.54):

$$\begin{aligned}&\sum_{s_1,s_3}\left[\overline{u^{s_3}}_A(\mathrm{p}_3)\,\gamma^{\mu\,A}{}_B\,u^{s_1\,B}(\mathrm{p}_1)\right]\left[\overline{u^{s_1}}_C(\mathrm{p}_1)\,\overline{\gamma}^{\nu\,C}{}_D\,u^{s_3\,D}(\mathrm{p}_3)\right]\\ &\quad= \sum_{s_3}\gamma^{\mu\,A}{}_B\left[\sum_{s_1}u^{s_1\,B}(\mathrm{p}_1)\,\overline{u^{s_1}}_C(\mathrm{p}_1)\right]\overline{\gamma}^{\nu\,C}{}_D\left[u^{s_3\,D}(\mathrm{p}_3)\,\overline{u^{s_3}}_A(\mathrm{p}_3)\right]\\ &\quad\overset{(5.54)}{=} \gamma^{\mu\,A}{}_B\,(\not{\mathrm{p}}_1+m_e c\mathbb{1})^B{}_C\,\overline{\gamma}^{\nu\,C}{}_D\,(\not{\mathrm{p}}_3+m_e c\mathbb{1})^D{}_A\\ &\quad= \mathrm{Tr}\left[\boldsymbol{\gamma}^\mu\,(\not{\mathrm{p}}_1+m_e c\mathbb{1})\,\overline{\boldsymbol{\gamma}}^\nu\,(\not{\mathrm{p}}_3+m_e c\mathbb{1})\right],\end{aligned} \tag{5.135}$$

which is independent of the spins $s_1, s_3$ that are not being measured.

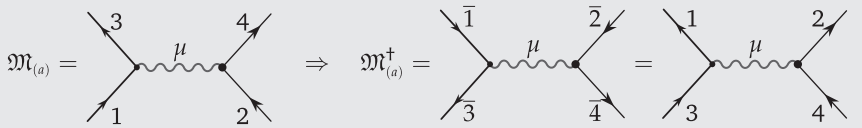
It then follows that

$$\begin{aligned}\langle|\mathfrak{M}_{(a)}|^2\rangle &= \frac{g_e^4}{(\mathrm{p}_1-\mathrm{p}_3)^4}\sum_{s_1,s_3}\mathrm{Tr}\left[\overline{u^{s_3}}(\mathrm{p}_3)\,\boldsymbol{\gamma}^\mu\,u^{s_1}(\mathrm{p}_1)\right]\mathrm{Tr}\left[\overline{u^{s_1}}(\mathrm{p}_1)\,\overline{\boldsymbol{\gamma}}^\nu\,u^{s_3}(\mathrm{p}_3)\right]\\ &\quad\times\sum_{s_2,s_4}\mathrm{Tr}\left[\overline{U^{s_4}}(\mathrm{p}_4)\,\boldsymbol{\gamma}_\mu\,U^{s_2}(\mathrm{p}_2)\right]\mathrm{Tr}\left[\overline{U^{s_2}}(\mathrm{p}_2)\,\overline{\boldsymbol{\gamma}}_\nu\,U^{s_4}(\mathrm{p}_4)\right]\\ &= \frac{g_e^4}{(\mathrm{p}_1-\mathrm{p}_3)^4}\mathrm{Tr}\left[\boldsymbol{\gamma}^\mu\,(\not{\mathrm{p}}_1+m_e c\mathbb{1})\,\overline{\boldsymbol{\gamma}}^\nu\,(\not{\mathrm{p}}_3+m_e c\mathbb{1})\right]\\ &\quad\times\mathrm{Tr}\left[\boldsymbol{\gamma}_\mu\,(\not{\mathrm{p}}_2+m_\mu c\mathbb{1})\,\overline{\boldsymbol{\gamma}}_\nu\,(\not{\mathrm{p}}_4+m_\mu c\mathbb{1})\right]\end{aligned} \tag{5.136}$$

$$= \frac{g_e^4}{(\mathrm{p}_1-\mathrm{p}_3)^4}\,X^{\mu\nu}(1,3;e^-)\,X_{\mu\nu}(2,4;\mu^-). \tag{5.137}$$

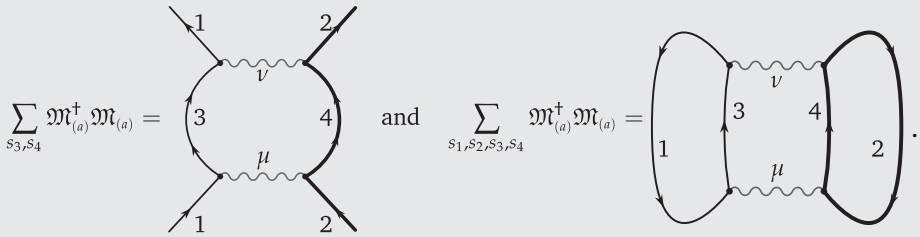
https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

**Digression 5.13** This result – in fact, the entire procedure (5.131)–(5.137) – may also be depicted graphically:



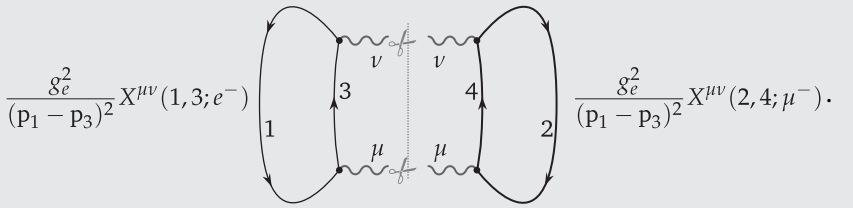

(5.138a)

where the diagram labels were simplified, so "1" stands for $u^{s_1,a}(\mathrm{p}_1)$ and "$\bar{1}$" for $\overline{u}^{s_1}_a(\mathrm{p}_1)$, etc. The product $\mathfrak{M}^\dagger_{(a)}\mathfrak{M}_{(a)}$ is then simply depicted by putting two diagrams next to each other. However, the summation of the product $\mathfrak{M}^\dagger_{(a)}\mathfrak{M}_{(a)}$ over spin (and, in general, all other unmeasured degrees of freedom) of, say, particle 1 is graphically depicted by connecting (concatenating) the two lines labeled "1" into a single line. Thus,




(5.138b)

By cutting the photon lines, we arrive at the graphical depiction (5.137):




(5.138c)

This graphical rendition of the computation of $\langle|\mathfrak{M}|^2\rangle$ is further detailed in Ref. [64].

The computation of the tensors

$$X^{\mu\nu}(1,3;e^-) := \mathrm{Tr}\left[\boldsymbol{\gamma}^\mu\,(\not{\mathrm{p}}_1 + m_e c\mathbb{1})\,\overline{\boldsymbol{\gamma}}^\nu\,(\not{\mathrm{p}}_3 + m_e c\mathbb{1})\right], \tag{5.139a}$$

$$X_{\mu\nu}(2,4;\mu^-) := \mathrm{Tr}\left[\boldsymbol{\gamma}_\mu\,(\not{\mathrm{p}}_2 + m_\mu c\mathbb{1})\,\overline{\boldsymbol{\gamma}}_\nu\,(\not{\mathrm{p}}_4 + m_\mu c\mathbb{1})\right] \tag{5.139b}$$

reduces to writing out the $\boldsymbol{\gamma}$-polynomials in the square brackets (5.136), and simplifying using the identities (A.125). The final result is [☞ Ref. [241] for the "factorization" (5.137) and derivation]

$$\begin{aligned}\langle|\mathfrak{M}_{(a)}|^2\rangle = \frac{8g_e^4}{(\mathrm{p}_1-\mathrm{p}_3)^4}\big[&(\mathrm{p}_1{\cdot}\mathrm{p}_2)(\mathrm{p}_3{\cdot}\mathrm{p}_4) + (\mathrm{p}_1{\cdot}\mathrm{p}_4)(\mathrm{p}_3{\cdot}\mathrm{p}_2) + 2(m_e m_\mu c^2)^2\\ &- (m_\mu c)^2(\mathrm{p}_1{\cdot}\mathrm{p}_3) - (m_e c)^2(\mathrm{p}_2{\cdot}\mathrm{p}_4)\big].\end{aligned} \tag{5.140}$$

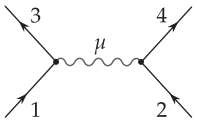
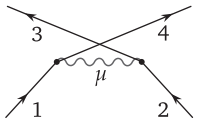
#### Electron–electron scattering

The computation (5.140) used that $e^- \neq \mu^-$. However, the result may be adapted also to the elastic $e^- + e^- \to e^- + e^-$ scattering, named after Christian Møller. This, however, does not reduce to a simple replacement $m_\mu \to m_e$ in the final expression (5.140), since when the two outgoing particles are identical, we must take into account another, equally possible process, depicted by the

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

Feynman diagram (c) in Figure 5.1 on p. 196. The total amplitude is then the *difference* between the amplitudes for (b) and for (c) in Figure 5.1 on p. 196. Indeed, since the electrons are fermions, the total amplitude must be antisymmetric with respect to the exchange of any two, and so also the two outgoing electrons. Thus,

$$\begin{aligned}\mathfrak{M}_{2e^-\to 2e^-} &= \mathfrak{M}_{(b)} - \mathfrak{M}_{(c)} \\ &= -\frac{g_e^2}{(\mathrm{p}_1-\mathrm{p}_3)^2}\left[\overline{u_3}\,\boldsymbol{\gamma}^\mu\,u_1\right]\left[\overline{u_4}\,\boldsymbol{\gamma}_\mu\,u_2\right] + \frac{g_e^2}{(\mathrm{p}_1-\mathrm{p}_4)^2}\left[\overline{u_4}\,\boldsymbol{\gamma}^\mu\,u_1\right]\left[\overline{u_3}\,\boldsymbol{\gamma}_\mu\,u_2\right] \qquad (5.141a)\end{aligned}$$

$=$  $-$  (5.141b)



where the expression is simplified by not writing the indices indicating the spin or those for the Dirac spinor components, and all these arguments are denoted by a single subscript: $u^{s_i,A}(\mathrm{p}_i) \to u_i$. Computing $\langle|\mathfrak{M}|^2\rangle$ in this case complicates as compared to (5.131)–(5.140): squaring the expression (5.141) by absolute value, we obtain

$$|\mathfrak{M}_{2e^-\to 2e^-}|^2 = |\mathfrak{M}_{(b)}|^2 + |\mathfrak{M}_{(c)}|^2 - 2\,\Re e\left(\mathfrak{M}_{(b)}^\dagger\,\mathfrak{M}_{(c)}\right). \qquad (5.142)$$

The first two summands may be copied from equation (5.140), upon changing $m_\mu \to m_e$ and swapping $3 \leftrightarrow 4$ for $|\mathfrak{M}_{(c)}|^2$. The remaining, "interference" summand[18] is "a little" more complicated:

$$\begin{aligned}\mathfrak{M}_{(b)}^\dagger\mathfrak{M}_{(c)} &\propto [\overline{u}_2\overline{\boldsymbol{\gamma}}^\mu u_4][\overline{u}_1\overline{\boldsymbol{\gamma}}_\mu u_3][\overline{u}_4\boldsymbol{\gamma}^\nu u_1][\overline{u}_3\boldsymbol{\gamma}_\nu u_2] = [\overline{u}_2\overline{\boldsymbol{\gamma}}^\mu u_4][\overline{u}_4\boldsymbol{\gamma}^\nu u_1][\overline{u}_1\overline{\boldsymbol{\gamma}}_\mu u_3][\overline{u}_3\boldsymbol{\gamma}_\nu u_2] \\ &= \left[\overline{u}_2\overline{\boldsymbol{\gamma}}^\mu u_4\overline{u}_4\boldsymbol{\gamma}^\nu u_1\overline{u}_1\overline{\boldsymbol{\gamma}}_\mu u_3\overline{u}_3\boldsymbol{\gamma}_\nu u_2\right], \qquad (5.143)\end{aligned}$$

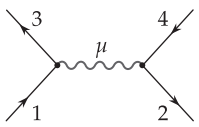
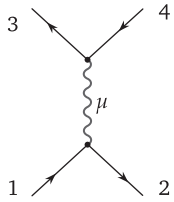
summing over spins produces

$$\begin{aligned}\langle\mathfrak{M}_{(b)}^\dagger\mathfrak{M}_{(c)}\rangle &\propto \left\langle\left[\overline{u}_2\overline{\boldsymbol{\gamma}}^\mu u_4\overline{u}_4\boldsymbol{\gamma}^\nu u_1\overline{u}_1\overline{\boldsymbol{\gamma}}_\mu u_3\overline{u}_3\boldsymbol{\gamma}_\nu u_2\right]\right\rangle = \left\langle\mathrm{Tr}\left[\overline{\boldsymbol{\gamma}}^\mu u_4\overline{u}_4\boldsymbol{\gamma}^\nu u_1\overline{u}_1\overline{\boldsymbol{\gamma}}_\mu u_3\overline{u}_3\boldsymbol{\gamma}_\nu u_2\overline{u}_2\right]\right\rangle \\ &= \mathrm{Tr}\left[\overline{\boldsymbol{\gamma}}^\mu(\not{\mathrm{p}}_4 + m_e c\mathbb{1})\boldsymbol{\gamma}^\nu(\not{\mathrm{p}}_1 + m_e c\mathbb{1})\overline{\boldsymbol{\gamma}}_\mu(\not{\mathrm{p}}_3 + m_e c\mathbb{1})\boldsymbol{\gamma}_\nu(\not{\mathrm{p}}_2 + m_e c\mathbb{1})\right], \qquad (5.144)\end{aligned}$$

for the computation of which one needs identities like (A.125), but up to and including the eighth degree in the $\boldsymbol{\gamma}$-matrices. However, using the *matrix* identities (A.121)–(A.122) these may always be reduced to the listed identities (A.125) [☞ Theorem A.5 on p. 487].

#### Electron–positron scattering

The elastic scattering $e^- + e^+ \to e^- + e^+$ is known as Bhabha scattering, after Homi Jehangir Bhabha. Again there are contributions from two sub-processes:

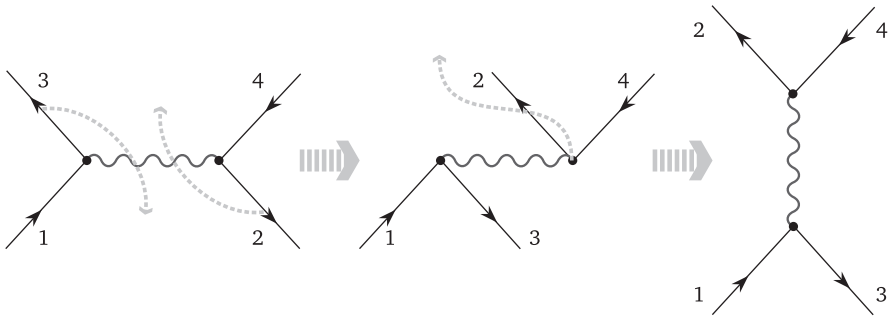
$\mathfrak{M}_{(d)} =$  and $\mathfrak{M}_{(e)} =$  (5.145)



[18] The existence of such interference summands is the hallmark of quantum mechanics: the basic principle is that in classical physics one adds probabilities of the partial contributions to a process, whereas in quantum physics one adds the amplitudes of those probabilities and then squares this sum to obtain the probability.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

It is not hard to show that $\mathfrak{M}_{(d)}(1,2,3,4) = \mathfrak{M}_{(c)}(1,3,2,4)$; i.e., by exchanging the incoming positron with the outgoing electron: the incoming positron, labeled "2," is equivalent to the outgoing electron, labeled "3," together with the $p_2 \leftrightarrow -p_3$ swap:




(5.146)

Antisymmetrizing with respect to this exchange of two fermions, we then have

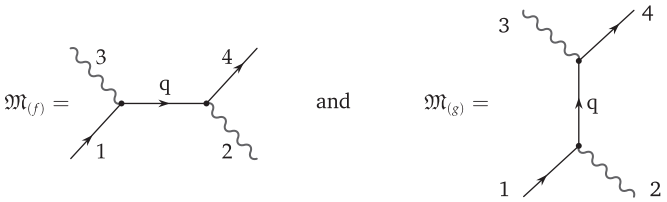
$$\mathfrak{M}_{e^-e^+\to e^-e^+} = \mathfrak{M}_{(d)} - \mathfrak{M}_{(e)} \tag{5.147a}$$

$$= -\frac{g_e^2}{(p_1-p_3)^2}[\overline{u}_3\boldsymbol{\gamma}^\mu u_1][\overline{v}_2\boldsymbol{\gamma}_\mu v_4] + \frac{g_e^2}{(p_1+p_2)^2}[\overline{v}_2\boldsymbol{\gamma}^\mu u_1][\overline{u}_3\boldsymbol{\gamma}_\mu v_4]. \tag{5.147b}$$

The expression for $\mathfrak{M}_{(d)}$ was obtained from equation (5.131), swapping $U_2 \to v_4$: incoming muon into the incoming (backwards in time!) positron, as well as $\overline{U}_4 \to \overline{v}_2$: outgoing muon into the outgoing (backwards in time!) positron.

#### Compton scattering

For electron–photon scattering, there are again two diagrams:




(5.148)

This time, the diagrams do not differ in an exchange of two fermions – we evidently do differentiate between the incoming and the outgoing electron, so there is no antisymmetrization; these amplitudes are therefore being added,

$$\mathfrak{M}_{e^-\gamma\to e^-\gamma} = \mathfrak{M}_{(f)} + \mathfrak{M}_{(g)}, \tag{5.149a}$$

where, following Procedure 5.2 on p. 193, we obtain

$$-i\mathfrak{M}_{(f)}(2\pi)^4\delta^4(p_1+p_2-p_3-p_4)$$

$$= \int \frac{d^4q}{(2\pi)^4}(2\pi)^4\delta^4(p_1-p_3-q)(2\pi)^4\delta^4(p_2-p_4+q)$$

$$\times \overline{u}_4\,\epsilon_2^\mu(-ig_e\boldsymbol{\gamma}_\mu)\,\frac{i(\not{q}+m_ec\mathbb{1})}{q^2-m_e^2c^2}(-ig_e\boldsymbol{\gamma}_\nu)\epsilon_3^{\nu*}\,u_1, \tag{5.149b}$$

$$\mathfrak{M}_{(f)} = \frac{g_e^2}{(p_1-p_3)^2-m_e^2c^2}\left(\epsilon_2^\mu[\overline{u}_4\,\boldsymbol{\gamma}_\mu(\not{p}_1-\not{p}_3+m_ec\mathbb{1})\boldsymbol{\gamma}_\nu\,u_1]\epsilon_3^{\nu*}\right), \tag{5.149c}$$

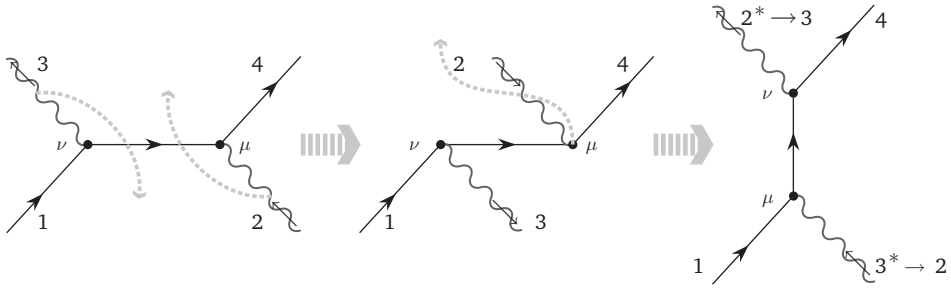
https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

$$\mathfrak{M}_{(g)} = \frac{g_e^2}{(\mathrm{p}_1+\mathrm{p}_2)^2 - m_e^2c^2}\,(\epsilon_3^{\mu*}[\overline{u}_4\,\boldsymbol{\gamma}_\mu(\not{\mathrm{p}}_1 - \not{\mathrm{p}}_3 + m_e c\mathbb{1})\boldsymbol{\gamma}_\nu\,u_1]\epsilon_2^\nu). \tag{5.149d}$$

We obtained the amplitude $\mathfrak{M}_{(g)}$ from $\mathfrak{M}_{(f)}$ by swapping

$$\mathrm{p}_2 \leftrightarrow -\mathrm{p}_3 \quad \text{and} \quad (\epsilon_2^\mu, \epsilon_3^{\nu*}) \leftrightarrow (\epsilon_3^{\mu*}, \epsilon_2^\nu). \tag{5.150}$$

This is easy to depict diagrammatically:

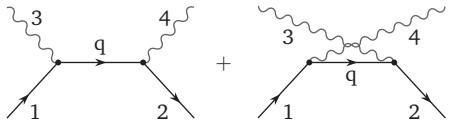



(5.151)

**Electron–positron pair annihilation and creation**

For inelastic scattering $e^- + e^+ \to 2\gamma$, there are again two diagrams:

$$\mathfrak{M}_{e^-+e^+\to 2\gamma} = \mathfrak{M}_{(h)} + \mathfrak{M}_{(i)} =$$



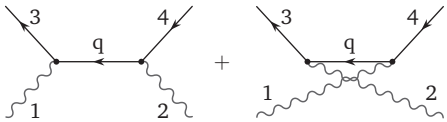

(5.152)

which are being added: They differ in the exchange of two photons $3 \leftrightarrow 4$, which are bosons, so the total amplitude is being symmetrized. As in the previous examples, we obtain

$$\begin{aligned} &= \frac{g_e^2}{(\mathrm{p}_1-\mathrm{p}_3)^2 - m_e^2c^2}\,(\epsilon_4^{\nu*}[\overline{v}_2\boldsymbol{\gamma}_\nu(\not{\mathrm{p}}_1 - \not{\mathrm{p}}_3 + m_e c\mathbb{1})\boldsymbol{\gamma}_\mu u_1]\epsilon_3^{\mu*}) \\ &\quad + \frac{g_e^2}{(\mathrm{p}_1-\mathrm{p}_4)^2 - m_e^2c^2}\,(\epsilon_3^{\nu*}[\overline{v}_2\boldsymbol{\gamma}_\nu(\not{\mathrm{p}}_1 - \not{\mathrm{p}}_4 + m_e c\mathbb{1})\boldsymbol{\gamma}_\mu u_1]\epsilon_4^{\mu*}). \end{aligned} \tag{5.153}$$

For the process of *pair-creation* $2\gamma \to e^- + e^+$, there are again two diagrams:

$$\mathfrak{M}_{2\gamma\to e^-+e^+} =$$




(5.154)

These contributions to the amplitude are being added as they again differ in the exchange of the two incoming photons. Owing to the evident (time-reversal) symmetry between the results (5.154) and (5.152), we have that $\mathfrak{M}_{2\gamma\to e^-+e^+} = \mathfrak{M}^\dagger_{e^-+e^+\to 2\gamma}$.

### 5.3.2 Effective cross-sections and lifetimes

The results for $\mathfrak{M}$ and $\langle|\mathfrak{M}|^2\rangle$ from the previous section may now be used in the above formulae (3.112) and (3.114), (3.122) and (3.127), as well as (3.159), (3.161) and (3.162).

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

**Mott and Rutherford scattering**

For the scattering of a light "probe" on a heavy "target" where the spins are not measured, we may use the results (5.136). In addition, in the approximation where the target mass ($m_B = M$) is sufficiently larger than the probe mass ($m_A = m$) so that the target recoil is negligible – which is easily realistic if the target is affixed in the lab – we use the result (3.161):

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \approx \left(\frac{\hbar}{8\pi Mc}\right)^2 \left\langle |\mathfrak{M}|^2 \right\rangle . \tag{5.155}$$

Since the target is immovable, we have

$$\mathrm{p}_1 = (-E/c, \vec{p}_1), \qquad \mathrm{p}_2 = (-Mc, \vec{0}), \qquad \mathrm{p}_3 \approx (-E/c, \vec{p}_3), \qquad \mathrm{p}_4 \approx (-Mc, \vec{0}), \tag{5.156}$$

where we used the conservation law of energy, i.e., the 0th component of 4-momentum, and have approximated $E_4 \approx Mc^2$ and $\vec{p}_4 \approx \vec{0}$, so that $E_3 \approx E_1 = E$. It follows that the angle in the relation $\vec{p}_1 \cdot \vec{p}_3 = \vec{p}^2 \cos\theta$ is small, $\theta \approx 0$, so that $|\vec{p}_1| \approx |\vec{p}_3| =: |\vec{p}|$. In this approximation,

$$(\mathrm{p}_1 - \mathrm{p}_3)^2 \approx -(\vec{p}_1 - \vec{p}_3)^2 = -\vec{p}_1{}^2 - \vec{p}_3{}^2 + 2\vec{p}_1 \cdot \vec{p}_3 = -4\vec{p}^2 \sin^2\left(\frac{\theta}{2}\right), \tag{5.157a}$$

$$(\mathrm{p}_1 \cdot \mathrm{p}_3) \approx \frac{E^2}{c^2} - \vec{p}_1 \cdot \vec{p}_3 = \vec{p}^2 + m^2c^2 - \vec{p}^2 \cos\theta \ = m^2c^2 + 2\vec{p}^2 \sin^2\left(\frac{\theta}{2}\right), \tag{5.157b}$$

$$(\mathrm{p}_1 \cdot \mathrm{p}_2) = ME \approx (\mathrm{p}_2 \cdot \mathrm{p}_3) \approx (\mathrm{p}_1 \cdot \mathrm{p}_4) \approx (\mathrm{p}_3 \cdot \mathrm{p}_4), \qquad (\mathrm{p}_2 \cdot \mathrm{p}_4) \approx M^2c^2. \tag{5.157c}$$

Thus,

$$\left\langle |\mathfrak{M}|^2 \right\rangle \approx \left(\frac{g_e^2\, Mc}{\vec{p}^{\,2} \sin^2(\frac{\theta}{2})}\right)^2 \left(m^2c^2 + \vec{p}^{\,2} \cos^2\left(\frac{\theta}{2}\right)\right), \tag{5.158}$$

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \approx \left(\frac{\alpha\hbar}{2\,\vec{p}^{\,2} \sin^2(\frac{\theta}{2})}\right)^2 \left(m^2c^2 + \vec{p}^{\,2} \cos^2\left(\frac{\theta}{2}\right)\right). \tag{5.159}$$

This is Mott's formula, which is a very good approximation of the differential cross-section for $e^-$–$p^+$ scattering, and even better for electron scattering on heavy ions. In the approximation where $\vec{p}^2 \ll m^2c^2$, we obtain

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \approx \left(\frac{\alpha\hbar}{2\,\vec{p}^{\,2} \sin^2(\frac{\theta}{2})}\right)^2 m^2c^2 = \left(\frac{\alpha\hbar c}{2\,m\vec{v}^{\,2} \sin^2(\frac{\theta}{2})}\right)^2, \tag{5.160}$$

which is the classical Rutherford formula (2.3).

The system of equations (4-momentum conservation)

$$(-E_1/c, \vec{p}_1) + (-Mc, \vec{0}) = (-E_3/c, \vec{p}_3) + (-E_4/c, \vec{p}_4) \tag{5.161a}$$

produces, denoting $p_i := |\vec{p}_i|$,

$$\sqrt{m^2c^2 + p_1^2} + Mc = \sqrt{m^2c^2 + p_3^2} + \sqrt{M^2c^2 + p_4^2}, \tag{5.161b}$$

$$p_1 = p_3 \cos\theta + p_4 \cos\phi, \tag{5.161c}$$

$$0 = p_3 \sin\theta - p_4 \sin\phi. \tag{5.161d}$$

Eliminating the angle $\phi$ from the last two equations produces

$$p_4 = |\vec{p}_1 - \vec{p}_3| = \sqrt{p_1{}^2 - 2p_1p_3\cos(\theta) + p_3{}^2} \tag{5.161e}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

which, together with relation (5.161b) gives $\left(\text{with } E_1 = c\sqrt{m^2c^2 + \vec{p}_1^{\,2}}\right)$

$$p_{3\pm} = \frac{p_1(E_1M + m^2c^2)\cos(\theta) \pm p_1(E_1 + Mc^2)\sqrt{M^2 - m^2\sin^2(\theta)}}{(M^2+m^2)c^2 + 2E_1M + p_1^{\,2}\sin^2(\theta)} \tag{5.162}$$

This is why simplifying approximations such as Mott's are convenient.

#### Electron–positron pair annihilation

In a model that has only electrons (and positrons) and photons, the decay – strictly speaking – is not possible: neither can a fermion (an electron or a positron) decay into any number of photons, nor can a photon decay into a real electron–positron pair [☞ Exercise 5.3.6]. However, the well-studied decay $\pi^0 \to 2\gamma$ is actually the process $\pi^0 = (q + \bar{q}) \to \gamma + \gamma$, which is in fact an inelastic scattering of a quark–antiquark pair that were, originally, bound into the state $\pi^0$. This process has contributions not only from the electromagnetic interaction, but also from weak and strong nuclear interactions, which complicates the estimate.

Instead, consider the decay of positronium, which is most conveniently computed in the positronium rest-frame, i.e., in the electron–positron CM system. It is known that in this system the electron and the positron move rather slowly, so we compute in the approximation where the electron and the positron are static immediately before their annihilation. The two photons created in the annihilation carry the same energy and so have linear momenta of the same magnitude and opposite direction. Thus we choose

$$\mathrm{p}_{e^-} = \mathrm{p}_1 = m_e c(-1,0,0,0), \qquad \mathrm{p}_{e^+} = \mathrm{p}_2 = m_e c(-1,0,0,0), \tag{5.163a}$$

$$\mathrm{p}_{\gamma_1} = \mathrm{p}_3 = m_e c(-1,0,0,1), \qquad \mathrm{p}_{\gamma_2} = \mathrm{p}_4 = m_e c(-1,0,0,-1), \tag{5.163b}$$

whereby it follows that

$$(\mathrm{p}_1 - \mathrm{p}_3)^2 - m_e^2c^2 = -2m_e^2c^2 = (\mathrm{p}_2 - \mathrm{p}_4)^2 - m_e^2c^2. \tag{5.164}$$

Besides, for the photons we use both the Lorenz gauge (5.91), whereby

$$\epsilon_3 \cdot \mathrm{p}_3 = 0 = \epsilon_4 \cdot \mathrm{p}_4, \tag{5.165}$$

as well as the Coulomb gauge, whereby the polarization 4-vectors $\epsilon_3$ and $\epsilon_4$ have no temporal component. Since $\mathrm{p}_1$ and $\mathrm{p}_2$ only have temporal components, it follows that

$$\epsilon_3 \cdot \mathrm{p}_1 = 0 = \epsilon_4 \cdot \mathrm{p}_1 \quad \text{and} \quad \epsilon_3 \cdot \mathrm{p}_2 = 0 = \epsilon_4 \cdot \mathrm{p}_2. \tag{5.166}$$

The expressions (5.153) may, after a little simplifying, be written as

$$\mathfrak{M}_{(h)} = \frac{g_e^2}{(\mathrm{p}_1 - \mathrm{p}_3)^2 - m_e^2c^2}\left([\bar{v}_2 \not{\epsilon}_4^*(\not{\mathrm{p}}_1 - \not{\mathrm{p}}_3 + m_e c\mathbb{1})\not{\epsilon}_3^* u_1]\right), \tag{5.167}$$

$$\mathfrak{M}_{(i)} = \frac{g_e^2}{(\mathrm{p}_1 - \mathrm{p}_4)^2 - m_e^2c^2}\left([\bar{v}_2 \not{\epsilon}_3^*(\not{\mathrm{p}}_1 - \not{\mathrm{p}}_4 + m_e c\mathbb{1})\not{\epsilon}_4^* u_1]\right), \tag{5.168}$$

where $\not{\epsilon}_i^* := \epsilon_i^{*\mu}\boldsymbol{\gamma}_\mu$ – the gamma-matrix is not conjugated. Consider first $\mathfrak{M}_{(h)}$, where

$$\not{\mathrm{p}}_1\not{\epsilon}_3^* \overset{(A.126a)}{=} -\not{\epsilon}_3^*\not{\mathrm{p}}_1 + 2\epsilon_3^* \cdot \mathrm{p}_1 \overset{(5.166)}{=} -\not{\epsilon}_3^*\not{\mathrm{p}}_1, \tag{5.169a}$$

$$\not{\mathrm{p}}_3\not{\epsilon}_3^* \overset{(A.126a)}{=} -\not{\epsilon}_3^*\not{\mathrm{p}}_3 + 2\epsilon_3^* \cdot \mathrm{p}_3 \overset{(5.165)}{=} -\not{\epsilon}_3^*\not{\mathrm{p}}_3, \tag{5.169b}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

$$(\not{p}_1 - \not{p}_3 + m_e c\mathbb{1})\not{\epsilon}_3^* u_1 = \not{\epsilon}_3^*(-\not{p}_1 + \not{p}_3 + m_e c\mathbb{1})u_1 = \not{\epsilon}_3^*\not{p}_3 u_1, \tag{5.169c}$$

where the last equality holds as the incoming electron, $u_1$, is on-shell, i.e., it satisfies the Dirac equation, $(\not{p}_1 - m_e c\mathbb{1})u_1 = 0$. $\mathfrak{M}_{(i)}$ is similarly simplified so that, using the choices (5.163b), we obtain

$$\begin{aligned}
\mathfrak{M}_{e^-+e^+\to 2\gamma} &= -\frac{g_e^2}{2m_e^2c^2}\overline{v}_2[\not{\epsilon}_4^*\not{\epsilon}_3^*\not{p}_3 + \not{\epsilon}_3^*\not{\epsilon}_4^*\not{p}_4]u_1 \qquad (5.170)\\
&= -\frac{g_e^2}{2m_e c}\overline{v}_2[\not{\epsilon}_4^*\not{\epsilon}_3^*(\boldsymbol{\gamma}^0 + \boldsymbol{\gamma}^3) + \not{\epsilon}_3^*\not{\epsilon}_4^*(\boldsymbol{\gamma}^0 - \boldsymbol{\gamma}^3)]u_1\\
&= -\frac{g_e^2}{2m_e c}\overline{v}_2[(\not{\epsilon}_4^*\not{\epsilon}_3^* + \not{\epsilon}_3^*\not{\epsilon}_4^*)\boldsymbol{\gamma}^0 + (\not{\epsilon}_4^*\not{\epsilon}_3^* - \not{\epsilon}_3^*\not{\epsilon}_4^*)\boldsymbol{\gamma}^3]u_1\\
&= -\frac{g_e^2}{2m_e c}\overline{v}_2[2\epsilon_{4\mu}^*\eta^{\mu\nu}\epsilon_{3\nu}^*\boldsymbol{\gamma}^0 + 4i\epsilon_{4\mu}^*\boldsymbol{\gamma}^{\mu\nu}\epsilon_{3\nu}^*\boldsymbol{\gamma}^3]u_1\\
&= -\frac{g_e^2}{2m_e c}\overline{v}_2[-2\vec{\epsilon}_4^*\cdot\vec{\epsilon}_3^*\boldsymbol{\gamma}^0 + i(\vec{\epsilon}_4^*\times\vec{\epsilon}_3^*)\cdot\vec{\boldsymbol{\Sigma}}\boldsymbol{\gamma}^3]u_1, \qquad (5.171)
\end{aligned}$$

where we used again that $\epsilon_i^0 = 0$ and where we defined

$$\boldsymbol{\Sigma}_i := 2\varepsilon_{ijk}\boldsymbol{\gamma}^{jk} = \tfrac{i}{2}\varepsilon_{ijk}[\boldsymbol{\gamma}^j, \boldsymbol{\gamma}^k]. \tag{5.172}$$

Finally, we use that the spins of the electron and the positron are antiparallel, and use

$$u_1^{\uparrow} = \sqrt{2mc}\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \quad u_1^{\downarrow} = \sqrt{2mc}\begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \quad \overline{v}_2^{\downarrow} = \sqrt{2mc}[0\ 0\ 1\ 0], \quad \overline{v}_2^{\uparrow} = \sqrt{2mc}[0\ 0\ 0\ 1]. \tag{5.173}$$

Thus, using the concrete matrices in Appendix A.6.1,

$$\mathfrak{M}_{\uparrow\downarrow} = -2ig_e^2(\vec{\epsilon}_3^* \times \vec{\epsilon}_4^*)_z = -\mathfrak{M}_{\downarrow\uparrow}, \tag{5.174}$$

from which it follows that the symmetric state of the electron–positron system, $(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)/\sqrt{2}$, cannot decay into two photons. However, since the process $e^+ + e^- \to \gamma$ is kinematically forbidden, it follows that the symmetric state of positronium may only decay into three or more photons.

On the other hand, the *antisymmetric* state, $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$, *can* decay into two photons. Thus we have

$$\mathfrak{M}_{|0,0\rangle} = \tfrac{1}{\sqrt{2}}(\mathfrak{M}_{\uparrow\downarrow} - \mathfrak{M}_{\downarrow\uparrow}), \tag{5.175}$$

where $|0,0\rangle = (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$ is the so-called *singlet* state of positronium before decay.

Next, re-insert the polarization vectors[19]

$$\vec{\epsilon}_{|1,+1\rangle} = -\tfrac{1}{\sqrt{2}}(1, i, 0) \qquad \text{and} \qquad \vec{\epsilon}_{|1,-1\rangle} = \tfrac{1}{\sqrt{2}}(1, -i, 0), \tag{5.176}$$

so that

$$(\vec{\epsilon}_3^* \times \vec{\epsilon}_4^*)_{\uparrow\downarrow} = \big(\vec{\epsilon}_{3,|1,+1\rangle}^* \times \vec{\epsilon}_{4,|1,-1\rangle}^*\big) = -\frac{1}{2}\begin{vmatrix}\hat{\mathrm{e}}_1 & \hat{\mathrm{e}}_2 & \hat{\mathrm{e}}_3\\ 1 & -i & 0\\ 1 & i & 0\end{vmatrix} = -i\,\hat{\mathrm{e}}_3 = -(\vec{\epsilon}_3^* \times \vec{\epsilon}_4^*)_{\downarrow\uparrow}, \tag{5.177}$$

whereby the photon polarization too must be in the antisymmetric superposition

$$\tfrac{1}{\sqrt{2}}\big(|1,+1\rangle_3|1,-1\rangle_4 - |1,-1\rangle_3|1,+1\rangle_4\big). \tag{5.178}$$

[19] The signs are chosen so that $\{\vec{\epsilon}_{|1,+1\rangle}, \vec{\epsilon}_{|1,-1\rangle}, \hat{\mathrm{e}}_3\}$ would form a right-handed coordinate system.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

Finally, adding the contributions to the amplitude as in the superposition (5.178),

$$\mathfrak{M}_{e^-+e^+\to 2\gamma} = -4g_e^2. \tag{5.179}$$

Although the final numeric value of this result seems disproportionately simple in comparison with the length and details of the derivation, note that we have also derived that the antiparallel spins in the electron–positron system imply that:

1. the positronium spin before the two-photon decay equals zero, i.e., the positronium is in the so-called singlet state $|0,0\rangle = (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$;
2. the spin of the two-photon state produced in the positronium decay equals zero, and the state itself is the antisymmetric superposition (5.178);
3. the triplet state of positronium, $|1,0\rangle = (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)/\sqrt{2}$, may only decay into three or more photons.

Given the amplitude (5.179), we may compute: First of all, using the result (3.127), we have the effective cross-section of the electron–positron annihilation in the CM system:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \Big(\frac{\hbar c}{8\pi(E_1+E_2)}\Big)^2 \frac{|\vec{p}_f|}{|\vec{p}_i|}|\mathfrak{M}|^2 = \Big(\frac{\hbar c}{16\pi(m_e c)}\Big)^2 \frac{|E_\gamma/c|}{|m_e v|}|-4g_e^2|^2, \tag{5.180}$$

where we used that, because of (5.163a)–(5.163b), $E_1 = mc^2 = E_2$ and $E_\gamma = m_e c^2$. Simplifying, we obtain

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{cv}\Big(\frac{\hbar\alpha}{m_e}\Big)^2, \qquad \text{and} \qquad \sigma = \frac{4\pi}{cv}\Big(\frac{\hbar\alpha}{m_e}\Big)^2, \tag{5.181}$$

since $\mathrm{d}\sigma/\mathrm{d}\Omega$ does not depend on angles.

For the decay constant and the lifetime of positronium, use relation (3.109), where the total number of scatterings equals $N = L\sigma$, and luminosity is $L = v\rho$, with $\rho$ the probability density of finding the electron and the positron at the decay location. For an individual positronium "atom," $\rho = |\Psi(\vec{0},t)|^2$ and $N$ represents the decay probability in unit time, i.e., the decay constant. Thus,

$$\Gamma = v\,\sigma\,|\Psi(\vec{0},t)|^2 = \frac{4\pi}{c}\Big(\frac{\hbar\alpha}{m_e}\Big)^2 |\Psi(\vec{0},t)|^2, \tag{5.182}$$

in agreement with Conclusion 3.2 on p. 113, and the relation (3.110). Recall: $[|\Psi(\vec{0},t)|^2] = L^{-3}$; at the end of Section 4.1.5, the result from analyzing the hydrogen atom was adapted, $|\Psi(\vec{0},t)|^2 = \big(\frac{\alpha m_e c}{\hbar\, n}\big)^3$. With this result, we finally get

$$\Gamma = \frac{4\pi}{c}\Big(\frac{\hbar\alpha}{m_e}\Big)^2\Big[\frac{1}{\pi}\Big(\frac{\alpha(\frac{1}{2}m_e)c}{\hbar\, n}\Big)^3\Big] = \frac{\alpha^5\, m_e c^2}{2\,\hbar\, n^3}, \tag{5.183}$$

and the positronium lifetime becomes

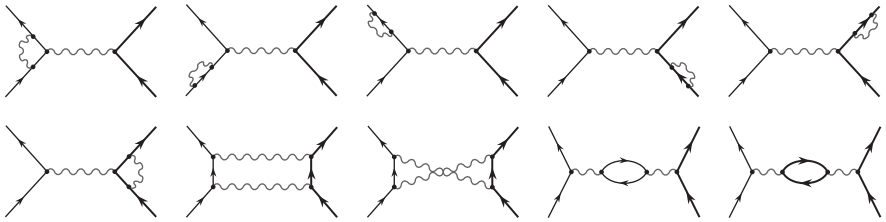
$$\tau = \frac{1}{\Gamma} = \frac{2\,\hbar\, n^3}{\alpha^5\, m_e c^2} \approx (1.24494\times 10^{-10}\,\mathrm{s}) \times n^3. \tag{4.51}$$

### 5.3.3 Renormalization

When discussing electron–muon scattering (5.129)–(5.140), we took into account only the Feynman diagram of lowest order in the $g_e$-, i.e., $\alpha$-power expansion. The results (5.131)–(5.140)

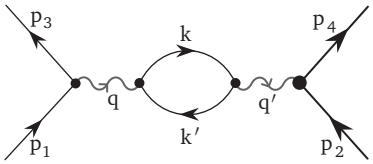
https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

produce $\mathfrak{M}_{(a)} = \sqrt{\langle|\mathfrak{M}_{(a)}|^2\rangle} = O(g_e^2) = O(\alpha)$. Next order corrections stem from the following diagrams:



(5.184)

of which we will consider the last two. Denote




(5.185)

**Calculation**

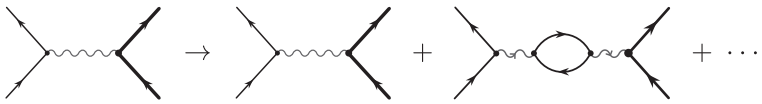
The amplitude for this process is obtained following Procedure 5.2 on p. 193:

$$\begin{aligned}
&\int \frac{\mathrm{d}^4\mathrm{q}}{(2\pi)^4}\frac{\mathrm{d}^4\mathrm{q}'}{(2\pi)^4}\frac{\mathrm{d}^4\mathrm{k}}{(2\pi)^4}\frac{\mathrm{d}^4\mathrm{k}'}{(2\pi)^4}\,(2\pi)^4\delta^4(\mathrm{p}_1-\mathrm{p}_3-\mathrm{q})\,(2\pi)^4\delta^4(\mathrm{q}-\mathrm{k}+\mathrm{k}')\,(2\pi)^4\delta^4(\mathrm{k}-\mathrm{k}'-\mathrm{q}')\\
&\quad\times(2\pi)^4\delta^4(\mathrm{p}_2-\mathrm{p}_4+\mathrm{q}')\,[\overline{u}_3(ig_e\boldsymbol{\gamma}^\mu)u_1]\left(\frac{-i\eta_{\mu\nu}}{\mathrm{q}^2}\right)\\
&\quad\times(-1)\,\mathrm{Tr}\left[(ig_e\boldsymbol{\gamma}^\nu)\frac{i}{\not{\mathrm{k}}-m_ec}(ig_e\boldsymbol{\gamma}^\rho)\frac{i}{\not{\mathrm{k}}'-m_ec}\right]\left(\frac{-i\eta_{\rho\sigma}}{(q')^2}\right)[\overline{U}_4(ig_e\boldsymbol{\gamma}^\sigma)U_2] \qquad (5.186a)\\
&=-g_e^4\int\frac{\mathrm{d}^4\mathrm{q}}{(2\pi)^4}\frac{\mathrm{d}^4\mathrm{k}}{(2\pi)^4}\,(2\pi)^4\delta^4(\mathrm{p}_1-\mathrm{p}_3-\mathrm{q})\,(2\pi)^4\delta^4(\mathrm{p}_2-\mathrm{p}_4+\mathrm{q})\\
&\quad\times[\overline{u}_3\,\boldsymbol{\gamma}^\mu\,u_1]\left(\frac{\eta_{\mu\nu}}{\mathrm{q}^2}\right)\mathrm{Tr}\left[\boldsymbol{\gamma}^\nu\frac{1}{\not{\mathrm{k}}-m_ec}\boldsymbol{\gamma}^\rho\frac{1}{\not{\mathrm{k}}-\not{\mathrm{q}}-m_ec}\right]\left(\frac{\eta_{\rho\sigma}}{\mathrm{q}^2}\right)[\overline{U}_4\,\boldsymbol{\gamma}^\sigma\,U_2]\\
&=-i(2\pi)^4\,\delta^4(\mathrm{p}_1+\mathrm{p}_2-\mathrm{p}_3-\mathrm{p}_4)\\
&\quad\times\left[\frac{-ig_e^4}{\mathrm{q}^4}\int\frac{\mathrm{d}^4\mathrm{k}}{(2\pi)^4}[\overline{u}_3\,\boldsymbol{\gamma}^\mu\,u_1]\frac{\mathrm{Tr}[\boldsymbol{\gamma}_\mu(\not{\mathrm{k}}+m_ec)\boldsymbol{\gamma}_\rho(\not{\mathrm{k}}-\not{\mathrm{q}}+m_ec)]}{(\mathrm{k}^2-m_e^2c^2)[(\mathrm{k}-\mathrm{q})^2-m_e^2c^2]}[\overline{U}_4\,\boldsymbol{\gamma}^\rho\,U_2]\right]_{\mathrm{q}=\mathrm{p}_1-\mathrm{p}_3}, \qquad (5.186b)
\end{aligned}$$

where the factor $(-1)$ in the expression (5.186a) reflects rule 7 in Procedure 5.2 on p. 193. With the abbreviation $\mathrm{q} := \mathrm{p}_1 - \mathrm{p}_3$, we have

$$\mathfrak{M}_{(a')} = \frac{-ig_e^4}{\mathrm{q}^4}[\overline{u}_3\,\boldsymbol{\gamma}^\mu\,u_1]\left\{\int\frac{\mathrm{d}^4\mathrm{k}}{(2\pi)^4}\frac{\mathrm{Tr}[\boldsymbol{\gamma}_\mu(\not{\mathrm{k}}+m_ec)\boldsymbol{\gamma}_\rho(\not{\mathrm{k}}-\not{\mathrm{q}}+m_ec)]}{(\mathrm{k}^2-m_e^2c^2)[(\mathrm{k}-\mathrm{q})^2-m_e^2c^2]}\right\}[\overline{U}_4\,\boldsymbol{\gamma}^\rho\,U_2]. \qquad (5.187)$$

Comparing with equation (5.131), we see that the inclusion of this $O(g_e^4)$ contribution[20]




$$\mathfrak{M}_{(a)} \;\to\; \mathfrak{M}_{(a;2)} + \mathfrak{M}_{(a';4)} + \cdots \qquad (5.188)$$

[20] For the complete result computed to $O(g_e^4)$, we of course must include all contributions (5.184) [☞ Refs. [243, 45, 580, 241]]; for brevity and pedagogical focus, only the last two are considered here.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

is equivalent to replacing the photon propagator in Procedure 5.2:

$$\frac{-i\eta_{\mu\rho}}{\mathrm{q}^2} \;\to\; \frac{-i\eta_{\mu\rho}}{\mathrm{q}^2} + \frac{-i\,H_{\mu\rho}}{\mathrm{q}^4} + \cdots \;=\; \frac{-i}{\mathrm{q}^2}\Big[\eta_{\mu\rho} + \frac{H_{\mu\rho}}{\mathrm{q}^2} + \cdots\Big], \tag{5.189}$$

where

$$H_{\mu\rho} \;:=\; i g_e^2 \int \frac{\mathrm{d}^4\mathrm{k}}{(2\pi)^4} \frac{\mathrm{Tr}[\boldsymbol{\gamma}_\mu(\not{\mathrm{k}} + m_e c)\boldsymbol{\gamma}_\rho(\not{\mathrm{k}} - \not{\mathrm{q}} + m_e c)]}{(\mathrm{k}^2 - m_e^2 c^2)[(\mathrm{k}-\mathrm{q})^2 - m_e^2 c^2]}. \tag{5.190}$$

Since $H_{\mu\rho}$ may only depend on the (rank-2) metric tensor $\eta_{\mu\nu}$ and the 4-momentum $q_\mu$, it must be that (as a rank-2 tensor and with $[H_{\mu\nu}] = 2[\mathrm{q}]$)

$$H_{\mu\rho} = -\eta_{\mu\rho}\,\mathrm{q}^2\,I(\mathrm{q}^2) + q_\mu\,q_\rho\,J(\mathrm{q}^2). \tag{5.191}$$

Here, $I(\mathrm{q}^2)$ and $J(\mathrm{q}^2)$ are two Lorentz-invariant functions of the 4-momentum q, so they must be functions of the Lorentz-invariant square $\mathrm{q}^2$. Since q is the 4-momentum of the *virtual* photon, $\mathrm{q}^2$ need not be restricted to the mass shell ($\mathrm{q}^2 = 0$, for the massless photon) and may attain arbitrary values.

The function $J(\mathrm{q}^2)$ contributes nothing to the final result, as it occurs, within the amplitude (5.187), only contracted with the 4-momentum:

$$[\overline{u}_3\boldsymbol{\gamma}^\mu\,u_1]q_\mu = [\overline{u}_3\,\not{q}\,u_1] = [\overline{u}_3\,(\not{p}_1 - \not{p}_3)\,u_1] = 0. \tag{5.192}$$

This last equality holds since both the incoming and the outgoing electrons are on the mass shell:

$$\not{p}_1\,u_1 = m_e c\,u_1 \qquad \text{and} \qquad \overline{u}_3\not{p}_3 = \overline{u}_3\,m_e c. \tag{5.193}$$

(Recall: $\not{p}_i = \boldsymbol{\gamma}^\mu p_{i\,\mu}$ are 4×4 matrices, $u_i$ 4-component column-matrices and $\overline{u}_i$ 4-component row-matrices.) It remains to compute the function $I(\mathrm{q}^2)$, which may be brought into the shape [243]:

$$I(\mathrm{q}^2) = \frac{g_e^2}{12\pi^2}\Big\{\int_{m_e^2}^{\infty}\frac{\mathrm{d}\xi}{\xi} - 6\int_0^1 \mathrm{d}\zeta\;\zeta(1-\zeta)\ln\Big(1 - \tfrac{\mathrm{q}^2}{m_e^2c^2}\zeta(1-\zeta)\Big)\Big\}. \tag{5.194}$$

While the first integral diverges logarithmically,

$$\int_{m_e^2}^{\infty}\frac{\mathrm{d}\xi}{\xi} = \lim_{\mu\to\infty}\int_{m_e^2}^{\mu^2}\frac{\mathrm{d}\xi}{\xi} = 2\lim_{\mu\to\infty}\ln\Big(\frac{\mu}{m_e}\Big) \;= \infty, \tag{5.195a}$$

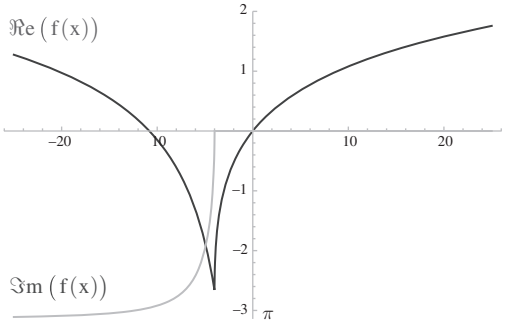
the second term, in curly brackets in relation (5.194), equals

$$\begin{aligned} f(x) &:= \frac{(12-5x)\sqrt{x(x+4)} - 6(x-2)(x+4)\tan^{-1}\Big(\sqrt{\frac{x}{x+4}}\Big)}{3\sqrt{x^3(x+4)}} \\ &= \frac{4}{x} - \frac{5}{3} - \frac{2(x-2)}{x}\sqrt{\frac{x+4}{x}}\tan^{-1}\Big(\sqrt{\frac{x}{x+4}}\Big), \end{aligned} \tag{5.195b}$$

$$x := -\frac{\mathrm{q}^2}{m_e^2c^2} = -\frac{(\mathrm{p}_1 - \mathrm{p}_3)^2}{m_e^2c^2} \approx 4\frac{\vec{p}^{\,2}_{e,\mathrm{in}}}{m_e^2c^2}\sin^2\big(\tfrac{\theta}{2}\big) = 4\frac{\vec{v}^{\,2}_{e,\mathrm{in}}}{c^2}\gamma_e^2\sin^2\big(\tfrac{\theta}{2}\big), \tag{5.195c}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

where $\vec{p}_{e,\mathrm{in}}$ and $\vec{v}_{e,\mathrm{in}}$ are, respectively, the 3-vectors of linear momentum and the velocity of the incoming electron, $\gamma_e$ the corresponding relativistic factor, and $\theta$ their deflection angle. Note that $f(x)$ varies relatively slowly:




$$
\begin{aligned}
&f(x) \sim \ln(x) \text{ for } |x| \gg 1\\
&f(x) \sim x/5 \text{ for } |x| \ll 1\\
&\min\left(\Re e(f(x))\right) = -8/3\\
&\lim_{x\to-\infty} \Im m\left(f(x)\right) = -\pi\\
&\Im m\left(f(x)\right) = 0 \text{ for } x \geqslant -4
\end{aligned}
\tag{5.196}
$$

The total amplitude is

$$\mathfrak{M}_{(a)} = \lim_{\mu\to\infty} \mathfrak{M}_{(a)}(\mathrm{q}^2,\mu) + \cdots, \tag{5.197a}$$

where

$$\mathfrak{M}_{(a)}(\mathrm{q}^2,\mu) = -g_R^2(\mu)\left[\overline{u}_3\,\boldsymbol{\gamma}^\mu\,u_1\right]\left(\frac{\eta_{\mu\nu}}{\mathrm{q}^2}\right)\left\{1+\frac{g_R^2(\mu)}{12\pi^2}\,f\left(\frac{-\mathrm{q}^2}{m_e^2c^2}\right)\right\}\left[\overline{U}_4\,\boldsymbol{\gamma}^\nu\,U_2\right] + \cdots, \tag{5.197b}$$

$$g_{e,R}(\mu) := g_e\sqrt{1-\frac{g_e^2}{6\pi^2}\ln\left(\frac{\mu}{m_e}\right)}, \tag{5.197c}$$

and where "$\cdots$" denotes omitted contributions from the other diagrams (5.184), as well as $O(g_e^6)$ contributions, and the equality (5.197b) with definition (5.197c) holds up to $O(g_e^6)$ corrections.

#### Physical meaning

The definition (5.197a) actually uncovers a conceptual error in the original set-up of the computation: The identification of the $e^- + \mu^- \to e^- + \mu^-$ elastic scattering amplitude of course depends on the strength of the interaction of the electron and the muon with the photons that mediate the electromagnetic interaction. The measure of the strength of that interaction was initially identified [☞ definition (5.122)] with the dimensionless parameter $g_e = \frac{e}{\sqrt{\epsilon_0\hbar c}}$ used in the assignment (5.124), which in turn is derived (within a field theory course) from the *classical* Lagrangian (5.118), with $q_\Psi \to -e$ (for the electron). However, the electric charge is of course a *measured* parameter, and elastic scatterings such as $e^- + \mu^- \to e^- + \mu^-$ in fact *define* the quantity that we call the (physical) electric charge. In other words, the original parameter $q_\Psi \to -e = -g_e\sqrt{\epsilon_0\hbar c}$ used in (5.118) is neither independently nor directly measurable, and should never have been identified identically with the physical electric charge of the electron.

> ***Conclusion 5.5*** *The quantity that* ***is*** *measurable and which* ***is*** *being compared with experimental data may in turn be identified with the symbol $g_{e,R}$, as defined by the relation (5.197c) as a function of the auxiliary (intermediate and, essentially, arbitrary) parameters $g_e, \mu$ – and up to $O(g_e^4)$ contributions, which were omitted in the expansion (5.197a).*

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

From the form of equation (5.197c), taking the $\mu \to \infty$ limit and the physical fact that the measurable charge $g_{e,R}$ is of course a finite quantity, it follows that the original and unmeasurable variable $g_e$ must be a function of the variable $\mu$, so that

$$g_{e,R} := \lim_{\mu\to\infty} g_e(\mu)\sqrt{1 - \frac{g_e^2(\mu)}{6\pi^2}\ln\left(\frac{\mu}{m_e}\right) + \cdots} \;<\; \infty \tag{5.198}$$

This perhaps fussy "detailing" is in fact logical, given that both $g_e$ and $\mu$ are auxiliary (intermediate) variables that serve only to connect the mathematical model (developed in a perturbative way from the classical physics model) to the physical quantities that this model describes.

In field theory (a course that should follow this introduction), the rules in Procedure 5.2 on p. 193 are *derived* from the Lagrangian for electrodynamics of charged spin-$\frac{1}{2}$ particles. The parameter $g_e$ should show up in this Lagrangian. However, just like that Lagrangian, the parameter by itself is not measurable, but defines the measurable charge by means of the iterative relation the beginning of which is given by equation (5.198). Thus, relation (5.197b) may be written as

$$\mathfrak{M}_{e^-+\mu^-\to e^-+\mu^-}(\mathrm{q}^2) = -g_{e,R}^2(\mathrm{q}^2)\left[\overline{u}_3\,\boldsymbol{\gamma}^\mu\,u_1\right]\left(\frac{\eta_{\mu\nu}}{\mathrm{q}^2}\right)\left[\overline{U}_4\,\boldsymbol{\gamma}^\nu\,U_2\right] + \cdots, \tag{5.199a}$$

$$g_{e,R}(\mathrm{q}^2) = g_{e,R}(0)\sqrt{1 + \frac{g_{e,R}^2(0)}{12\pi^2}\,f\!\left(\frac{-\mathrm{q}^2}{m_e^2c^2}\right)}, \tag{5.199b}$$

that is,

$$\begin{aligned}\alpha_{e,R}(\mathrm{q}^2) &= \alpha_{e,R}(0)\left\{1 + \frac{\alpha_{e,R}(0)}{3\pi}\,f\!\left(\frac{-\mathrm{q}^2}{m_e^2c^2}\right)\right\},\\ &\approx \alpha_{e,R}(0)\left\{1 + \frac{\alpha_{e,R}(0)}{3\pi}\,\ln\left(\frac{\mathrm{q}^2}{m_e^2c^2}\right)\right\}, \qquad \mathrm{q}^2 \gg m_e^2c^2,\end{aligned} \tag{5.199c}$$

where the electric charge is defined as the renormalized parameter of the electromagnetic interaction, $g_{e,R}(\mathrm{q}^2)$, as is then defined the parameter of the electromagnetic fine structure, $\alpha_{e,R}(\mathrm{q}^2)$, as a function of the Lorentz-invariant intensity of the 4-momentum transfer, from the "probe" (here $e^-$) to the "target" (here $\mu^-$).

The quantities $g_{e,R}(0)$ and $\alpha_{e,R}(0)$ are the limiting values of the functions $g_{e,R}(\mathrm{q}^2)$ and $\alpha_{e,R}(\mathrm{q}^2)$, when the 4-momentum transfer between the "probe" and "target" is negligible, and in that limit we have $\alpha_{e,R}(0) \approx \frac{1}{137}$. The numerical values of the corrections (5.199c) are relatively small, e.g., $O(6\times10^{-6})$ for a direct collision at $c/10$ speed, so that the value $\frac{1}{137}$ is used as a first approximation for $\alpha_{e,R}(\mathrm{q}^2)$ as if it were a constant. However, precise measurements of electromagnetic processes, such as in the Lamb shift (1.38a), indeed verify the corrections (5.199c).

The fact (5.199c) that the numerical value of the electric charge depends on the 4-momentum of the interaction with which that electric charge is being measured indicates the conceptual error in classical physics, where the parameters in the model of the physical system or process have *a priori* identified physical meaning and concrete value. The quantum nature of Nature teaches us that only those particular combinations and functions of the model parameters for which the values really can be measured must in fact have concrete (real and finite) values.

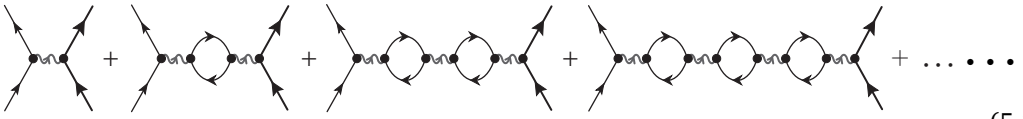
The contribution of the last diagram (5.184) equals the second term in equation (5.199c), only with a virtual muon in the central closed loop. That induces the replacement $m_e \to m_\mu$ in the result (5.199c), which reduces the contribution since

$$\ln\left(\frac{\mathrm{q}^2}{m_\mu^2c^2}\right) = \ln\left(\frac{\mathrm{q}^2}{m_e^2c^2}\right) - \left[\,2\ln(206) \approx 10.6558\,\right]. \tag{5.200}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

In fact, since the fermionic closed loop in the center of the diagram (5.185) depicts a *virtual* spin-$\frac{1}{2}$ fermion – which by definition is not observed – the contribution of the same diagram should be summed over all electrically charged spin-$\frac{1}{2}$ fermions. The electron's contribution is however dominant, since the electron is the lightest of all electrically charged spin-$\frac{1}{2}$ fermions, and the corrections (5.199c) [☞ relations (5.195b)–(5.196)] are scaled by the logarithm of the inverse mass of the particle in this central loop.

Finally, the relation (5.199c) is a result of the $O(g_e^4)$ contributions, which is depicted by the diagrams (5.188). It is not hard to show that the infinite series of diagrams of growing order:



(5.201)

summing the geometric series result in

$$\alpha_{e,R}(|\mathrm{q}^2|) \approx \frac{\alpha_{e,R}(0)}{1 - \frac{\alpha_{e,R}(0)}{3\pi} \ln\left(\frac{|\mathrm{q}^2|}{m_e^2 c^2}\right)}, \qquad |\mathrm{q}^2| \gg m_e^2 c^2. \tag{5.202}$$

In the domain $m_e^2 c^2 \ll \mathrm{q}^2 \ll m_e^2 c^2 \exp\left\{\frac{3\pi}{2\alpha(0)}\right\}$, $\alpha_{e,R}(\mathrm{q}^2)$ is a very slowly growing function, and the approximation $\alpha_{e,R}(\mathrm{q}^2) \approx \alpha_{e,R}(0) \approx \frac{1}{137}$ is very good.

The various diagrams that are not shown in the series (5.201) [☞ collection (5.184)] either provide significantly smaller contributions than those shown (comparing diagrams of the same order in $g_e^2$) or their contribution may be absorbed by *renormalizing* parameters such as the mass of the electron, $m_e$. The contributions (5.201) are usually called the "leading logarithm" contributions.

#### The renormalization group

Note that the result (5.197c) was obtained by including the quantum correction of only the lowest order, and the result (5.202) includes the dominant corrections. Evidently, these corrections – computed iteratively and sequentially – may be organized in a quantitative sequence, so that from one iteration to the next one there is a "flow":

$$\left(\alpha_{e,R}^{(0)}(|\mathrm{q}^2|) := \alpha_{e,R}(0)\right) \;\mapsto\; \cdots \;\mapsto\; \alpha_{e,R}^{(k)}(|\mathrm{q}^2|) \;\mapsto\; \alpha_{e,R}^{(k+1)}(|\mathrm{q}^2|) \;\mapsto\; \cdots \;\mapsto\; \alpha_{e,R}^{(\infty)}(|\mathrm{q}^2|), \tag{5.203}$$

where only the limiting result, $\alpha_{e,R}^{(\infty)}(|\mathrm{q}^2|)$, may be identified with the real physical quantity. The precise specification of the ordering of this renormalization "flow" depends on the concrete application – and this is one of those conceptual ideas that are applied in almost all branches of physics! The formal transformations that lead from one step in this renormalization flow into the next form a structure called the "renormalization group" – although it in fact does not satisfy the group axioms: The transformation $\mathcal{R}_{(k)}^{(k+1)}$ that takes the $k$th into the $(k+1)$th "step" has no binary operation defined with most other such transformations; only the consecutive "products" of the form $\mathcal{R}_{(k)}^{(k+1)} \circ \mathcal{R}_{(k-1)}^{(k)}$ are defined [☞ Comment 9.2 on p. 323].

In field theory, the application of this procedure and its structure was discovered by Ernst Stückelberg and Andre Petermann back in 1953 [502, 146]. The contemporary practice in field theory varies, but by now mostly relies on Kenneth Wilson's approach (1982 Nobel Prize), further developed by Joseph Polchinski [431]; see also Ref. [425] by Michael Peskin, who was Wilson's student, and who in turn was Gell-Mann's student. In this approach, the renormalization flow is organized by means of a varying upper limit in otherwise divergent integrals, i.e., by the energy/mass values up to which particles and excitations are included. The earlier approach, after

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

Murray Gell-Mann and Frank James Low, varies the value of the renormalization 4-momentum $\mu$ in the computations that lead to the results such as (5.198) and is still being used.

Essentially two types of behavior can result from this renormalization flow, which motivates:

> ***Definition 5.1*** *A quantum system is* ***renormalizable*** *if merely the parameters used in the classical Lagrangian of a system change owing to quantum corrections, but the* ***functional form*** *of this Lagrangian remains the same. Otherwise, a system is* ***non-renormalizable.***

See also Definition 11.1 on p. 419 for a more precise statement. Suffice it here to say that all possible Yang–Mills type gauge theory models interacting with any spin-0 and spin-$\frac{1}{2}$ matter fields – including the Standard Model – are renormalizable.

### 5.3.4 Exercises for Section 5.3

✎ ***5.3.1*** *For* $\Gamma \in \{\mathbb{1}, \boldsymbol{\gamma}^\mu, \boldsymbol{\gamma}^{\mu\nu}, (\boldsymbol{\gamma}^\mu\widehat{\boldsymbol{\gamma}}), \widehat{\boldsymbol{\gamma}}\}$*, compute for which* $\Gamma$ *is* $\overline{\Gamma} := (\boldsymbol{\gamma}^0\Gamma^\dagger\boldsymbol{\gamma}^0) = \Gamma$*, and for which* $\Gamma$ *is* $\overline{\Gamma} = -\Gamma$*.*

✎ ***5.3.2*** *Derive the equations of motion (5.120f).*

✎ ***5.3.3*** *Derive equation (5.140).*

✎ ***5.3.4*** *Derive equation (5.144), and then equation (5.142).*

✎ ***5.3.5*** *Find the diagram that, as (5.138c), depicts the result (5.144) and represents it by a graphical depiction as in Digression 5.13 on p. 198.*

✎ ***5.3.6*** *From the 4-vector equation* $\mathrm{p}_\gamma = \mathrm{p}_{e^-} + \mathrm{p}_{e^+}$ *and the symmetry* $\vec{p}_\gamma\cdot\vec{p}_{e^-} = \vec{p}_\gamma\cdot\vec{p}_{e^+}$*, as well as* $|\vec{p}_{e^-}| = |\vec{p}_{e^+}|$*, show that the decay of a real photon into a real electron–positron pair is kinematically forbidden.*

✎ ***5.3.7*** *Prove relation (5.195c), i.e., that* $\mathrm{q}^2 = -4\vec{p}_i{}^2\sin^2\left(\frac{\theta}{2}\right)$*.*

✎ ***5.3.8*** *Compute the collision energy for which the expression (5.202) diverges.*

## 5.4 Quantum electrodynamics of hadrons

The interaction between photons and quarks is described by the same theory as the interaction between photons and leptons, discussed in Sections 5.2.1 and 5.3.1. However, individual quarks are not available for experimenting. They are always within bound states, so-called hadrons: mesons, which are $(q\overline{q})$-systems, and baryons, which are $(qqq)$-systems. The interaction between a lepton (as a "probe") and a hadron (as the "target"), as well as between two hadrons, reduces to the interaction with individual (anti)quarks *within* the hadron, and so necessarily depends on the distribution of these individual (anti)quarks within the hadron. This distribution is described by so-called form-factors, which effectively[21] describe the *strong nuclear interactions* that bind the (anti)quarks into hadron bound states.

The second difficulty stems from the fact that the number of new hadrons is limited only by the available energy: as the collision energy grows, more and more new hadrons may be produced

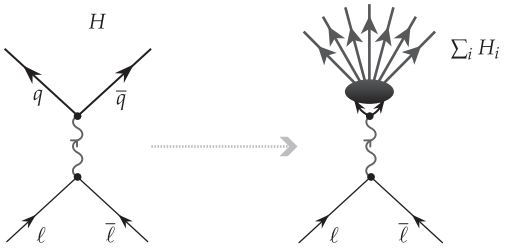
[21] Here, "effective" means "successfully and with no a-priori detailed fundamental basis/derivation"; the adjective "phenomenological" is used in the literature, in the same sense.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

in inelastic collisions, and the analysis very quickly becomes a combinatorially growing nightmare. Catalogues of hadrons such as Ref. [293] provide data about hundreds and hundreds of hadrons.

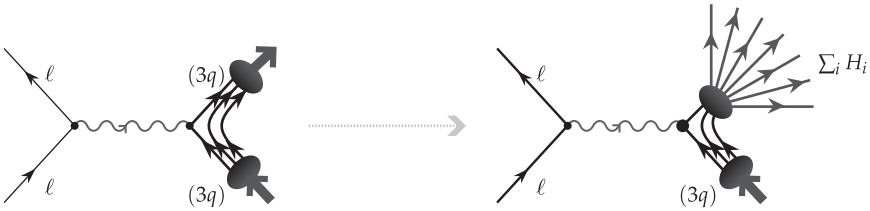
In principle, in the interaction between leptons and hadrons as well as between two hadrons, there are also contributions from weak nuclear interactions. However, that (third) source of complications is in most cases negligible, as the weak nuclear interaction is much weaker [☞ discussion on p. 67].

The two types of processes that are significant in hadronic experiments are production from electron–positron annihilation,




(5.204)

and so-called deep inelastic lepton–hadron collisions,




(5.205)

In both cases, increasing collision energy (indicated by the dotted arrow) gives rise to the production of a large number of outgoing hadrons. The strong nuclear interaction dominates this "hadronization," which in these diagrams is represented by the dark oval where the quark–antiquark pair (i.e., the three quarks) bind into a palette of bound states (hadrons). However, the electromagnetic part of the interaction may be separated as the interaction between the lepton and the individual (anti)quarks. Thus, computations of the amplitudes from the previous sections may be adapted also to these collisions, but the kinematic part of the analysis is significantly more complicated. Herein, we consider only the part of this analysis that is determined by the symmetries and general requirements.

### 5.4.1 Hadron production in electron–positron annihilation

Even with enough energy for the final collision results to include many hadrons, the electromagnetic part of the process (5.204) primarily reduces to transforming a $\ell\bar{\ell}$-pair, by way of a virtual photon, into a $q\bar{q}$-pair. That quark and antiquark then decay into lighter quarks, emit gluons and so produce a palette of various hadrons. This second stage of the process contains all the complications from strong interactions.

To describe the first stage, let $m$ be the lepton and antilepton mass, $M$ the mass of the produced quark and antiquark, and $Q$ the electric charge of the quark (so $-Q$ is the electric charge of the antiquark) in units of elementary electric charge, $e$, so that $Q(u) = +\frac{2}{3}$, $Q(d) = -\frac{1}{3}$, etc. Adapting the second term in the result (5.147b) we have

$$\mathfrak{M}_{\ell\bar{\ell}\to q\bar{q}} = \frac{Q\, g_e^2}{(\mathrm{p}_1+\mathrm{p}_2)^2}\,[\overline{v}_2\boldsymbol{\gamma}^\mu u_1][\overline{U}_3\boldsymbol{\gamma}_\mu V_4], \tag{5.206}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

where $u_1$, $\overline{v}_2$, $\overline{U}_3$ and $V_4$ are Dirac spinors for the incoming lepton and antilepton, and outgoing quark and antiquark, respectively. Averaging as in (5.132)–(5.137) we obtain

$$\begin{aligned}\left\langle |\mathfrak{M}_{\ell\bar{\ell}\to q\bar{q}}|^2 \right\rangle &= \frac{1}{4}\Big(\frac{Q g_e^2}{(\mathrm{p}_1+\mathrm{p}_2)^2}\Big)^2 \operatorname{Tr}\left[\boldsymbol{\gamma}^\mu(\not{\mathrm{p}}_1 + mc)\boldsymbol{\gamma}^\nu(\not{\mathrm{p}}_2 - mc)\right] \\ &\qquad\qquad \times \operatorname{Tr}\left[\boldsymbol{\gamma}^\mu(\not{\mathrm{p}}_4 - Mc)\boldsymbol{\gamma}^\nu(\not{\mathrm{p}}_3 + Mc)\right] \\ &= 8\Big(\frac{Q g_e^2}{(\mathrm{p}_1+\mathrm{p}_2)^2}\Big)^2 \left[(\mathrm{p}_1\cdot\mathrm{p}_3)(\mathrm{p}_2\cdot\mathrm{p}_4) + (\mathrm{p}_1\cdot\mathrm{p}_4)(\mathrm{p}_2\cdot\mathrm{p}_3) + 2(mc)^2(Mc)^2\right. \\ &\qquad\qquad \left. + (mc)^2(\mathrm{p}_3\cdot\mathrm{p}_4) + (Mc)^2(\mathrm{p}_1\cdot\mathrm{p}_2)\right] \\ &= Q^2 g_e^4\left\{1 + \Big(\frac{mc^2}{E}\Big)^2 + \Big(\frac{Mc^2}{E}\Big)^2 + \Big[1-\Big(\frac{mc^2}{E}\Big)^2\Big]\Big[1-\Big(\frac{Mc^2}{E}\Big)^2\Big]\cos^2\theta\right\}, \end{aligned} \tag{5.207}$$

where $E$ is the energy of the incoming lepton in the CM system and $\theta$ the angle between the incoming lepton and the outgoing quark. Treating the outgoing quark–antiquark as if they were free particles, the differential effective cross-section is given by the relation (3.127), whereupon angular integration yields

$$\sigma = \frac{\pi}{3}\Big(\frac{Q\hbar c\alpha}{E}\Big)^2 \sqrt{\frac{1-(Mc^2/E)^2}{1-(mc^2/E)^2}}\Big[1+\frac{1}{2}\Big(\frac{mc^2}{E}\Big)^2\Big]\Big[1+\frac{1}{2}\Big(\frac{Mc^2}{E}\Big)^2\Big]. \tag{5.208}$$

For energies below $Mc^2$, the effective cross-section becomes imaginary, i.e., the process is kinematically forbidden: $E < Mc^2$ is not enough energy to produce a quark–antiquark pair of mass $M$ each. In turn, if $E > Mc^2 \gg mc^2$, expanding the square-roots and multiplying the factors yields

$$\sigma = \frac{\pi}{3}\Big(\frac{Q\hbar c\alpha}{E}\Big)^2 F(m,M,E), \tag{5.209a}$$

$$\begin{aligned} F(m,M,E) &= \Re e\sqrt{\frac{1-(Mc^2/E)^2}{1-(mc^2/E)^2}}\Big[1+\frac{1}{2}\Big(\frac{mc^2}{E}\Big)^2\Big]\Big[1+\frac{1}{2}\Big(\frac{Mc^2}{E}\Big)^2\Big] \\ &\approx \Re e\Big[1+\Big(\frac{mc^2}{E}\Big)^2+\frac{5}{8}\Big(\frac{mc^2}{E}\Big)^4+\cdots\Big]\Big[1-\frac{3}{8}\Big(\frac{Mc^2}{E}\Big)^4+\cdots\Big]. \end{aligned} \tag{5.209b}$$

In typical experiments $mc^2 \lll E$, so that the first factor in $F(m,M,E)$ is negligibly different from 1. The second factor, however, gives a significant contribution when the energy suffices to produce a quark of mass $M$ but is not much larger than $Mc^2$. The behavior of the step-like function $F(m,M,E)$ near a threshold $E \sim Mc^2$, where the approximating condition $E > Mc^2$ of the expansion (5.209b) is not satisfied, is shown in Figure 5.2.

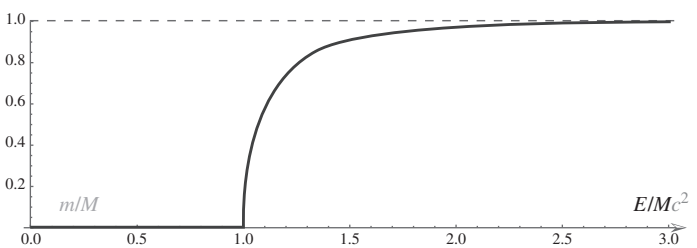



**Figure 5.2** A sketch of the function $\Re e\left[F(m,M,E)\right]$ near the value $E \sim Mc^2$.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

As the collision energy is increased, heavier and heavier quarks may be produced in the process. These quarks are in fact virtual particles, in the sense that they cannot be observed directly in the detectors, since they decay, emit gluons and finally bind into hadrons – dominated by strong interactions, and so very fast, $\sim 10^{-23}$ s. To avoid the need for estimating the details of this "hadronization," consider the ratio

$$R(E) := \frac{\sigma(e^- + e^+ \to \text{hadrons})}{\sigma(e^- + e^+ \to \mu^- + \mu^+)} \quad \approx \quad \Big[3 \sum_i {Q_i}^2\Big]_{M_i < E/c^2}. \tag{5.210}$$

Here the universal factors such as $\frac{\pi}{3}(\frac{\hbar c \alpha}{E})^2$ cancel, and the contributions to the function $F(m, M_i, E)$ from individually created quarks give only small corrections to the result given. As a function of the collision energy, $R(E)$ is step-like,[22] increasing suddenly when the energy $E$ reaches a threshold to produce a new quark, and is approximately constant between these thresholds. For example,

$$R(E) \approx 3[(\tfrac{2}{3})^2 + (-\tfrac{1}{3})^2] = \tfrac{5}{3}, \qquad E \leqslant M_{u,d}c^2, \tag{5.211a}$$
$$R(E) \approx 3[(\tfrac{2}{3})^2 + (-\tfrac{1}{3})^2 + (-\tfrac{1}{3})^2] = 2, \qquad E \leqslant M_s c^2, \tag{5.211b}$$
$$R(E) \approx 3[(\tfrac{2}{3})^2 + (-\tfrac{1}{3})^2 + (-\tfrac{1}{3})^2 + (\tfrac{2}{3})^2] = \tfrac{10}{3}, \qquad E \leqslant M_c c^2, \tag{5.211c}$$

and so forth. At energies between $M_c c^2 \approx 1{,}270$ MeV and about 2,000 MeV, a significant discrepancy from this simple form of $R(E)$ showed. However, it was soon discovered that this was due to resonance effects related to the production of the third (and mostly unexpected) electrically charged lepton ($m_\tau = 1{,}784\,\text{MeV}/c^2$), which decays mostly into hadrons. Considering the simplicity of the approximation and when the $\tau$-lepton contributions are correctly accounted for, the agreement of the simple relation (5.210) with experiments is very good.

Besides, the overall factor of 3 in the relation (5.210) stems from the fact that every quark has three colors, i.e., that for every mass and electric charge there actually exist three quarks – one of each color.

***Conclusion 5.6*** *The very good agreement of the simple approximation (5.210) with experiments is then the direct experimental proof of the existence of color – or at least the fact that every quark exists in (otherwise unexplained) triplicate.*

---

**Digression 5.14** At least one curious Student asked why do quarks have to have fractional electric charges, as they have been standardly assigned since 1974–5.

The original model by Han and Nambu indeed proposed quarks that were to have integral charges, dependent on color. For example, we may choose, following result (5.206),

$$Q(u^r) = +1, \;\; Q(u^y) = +1, \;\; Q(u^b) = 0,$$
$$Q(d^r) = 0, \;\; Q(d^y) = 0, \;\; Q(d^b) = -1, \tag{5.212a}$$

and similarly for the $s$-, $c$-, $b$- and $t$-quarks. Since the average electric charge for each quark "flavor" equals the standard (fractional) charge, no process where the amplitude is linearly proportional to the charges – such as hadron production from lepton collisions (5.206) – can possibly distinguish between the integrally charged Han–Nambu model and the fractionally charged Gell-Mann–Zweig model.

[22] The shape of these "steps" is described by the function $F(m, M, E)$, but it also exhibits resonant effects in the form of very narrow peaks of large intensity $\gg 1$, immediately above the $E = Mc^2$ threshold, which is not shown in the sketch in Figure 5.2 on p. 213 and the analysis of which is omitted herein.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

However, a process that depends on the *square* of the electric charge – such as (5.185), which contributes to the renormalization and so ultimately to the dependence of the interaction intensity on the interaction energy (5.197c) – can distinguish the integrally charged quarks from the fractionally charged quarks. Indeed, the "corrections" under the square-root symbol are in the result (5.197c) given for a single electrically charged particle – the electron – since we assumed that the virtual particle in the closed loop in the diagram (5.185) is in fact the electron. However, with energies $\mu c^2$, one must add the contributions from all the quarks and charged leptons with masses not larger than $\mu$, and which are proportional to the cumulative factor

$$\Big[\sum_i Q_i^2 \ln\big(\tfrac{\mu}{m_i}\big)\Big]_{m_i \leqslant \mu}, \tag{5.212b}$$

which grows differently for integrally charged quarks than for fractionally charged ones, and which give (one possible) experimentally measurable difference. Similarly, the results (5.211) would differ quantitatively for integrally charged quarks:

$$\widetilde{R}(E)\big|_{E\leqslant M_{u,d}c^2} \approx 3, \quad \widetilde{R}(E)\big|_{E\leqslant M_s c^2} \approx 4, \quad \widetilde{R}(E)\big|_{E\leqslant M_c c^2} \approx 6, \quad \text{etc.} \tag{5.212c}$$

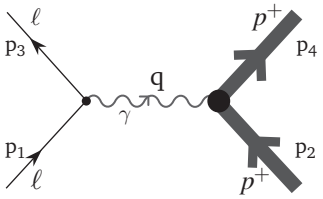
In the early 1970s, such comparisons with experiments confirmed the fractional electric charges of the quarks in the model of Gell-Mann and Zweig; see also Footnote *25* on p. 220.

### 5.4.2 The electrodynamics contribution in lepton–hadron scattering

A lepton–hadron collision occurs, to a first approximation, between the lepton and one of the (anti)quarks in the hadron, and by way of exchanging a single photon. Of course, when the hadron is a baryon, one must sum over all three quarks in the baryon, and if the hadron is a meson, one must sum the contributions from the interaction with the quark and with the antiquark. In this process, the strong nuclear force field, which keeps the (anti)quark state bound, receives part of the 4-momentum transfer, but this usually produces minor corrections to this initial approximation.

#### Elastic lepton–hadron scattering

If the proton were a point-like spin-$\frac{1}{2}$ Dirac spinor with no additional structure, relations (5.131) and (5.137) would be valid for *elastic* collisions




(5.213)

with only the small change, $m_\mu \to M := m_p$ in the function $X_{\mu\nu}(2,4;\mu^-) \to K_{\mu\nu}(2,4;p^+)$:

$$\Big\langle |\mathfrak{M}_{\ell p\to \ell p}|^2 \Big\rangle = \frac{g_e^4}{(\mathrm{p}_1-\mathrm{p}_3)^4} X^{\mu\nu}(1,3;\ell)\, K_{\mu\nu}(2,4;p^+), \tag{5.214}$$

$$\begin{aligned} X^{\mu\nu}(1,3;\ell) &= \mathrm{Tr}\left[\boldsymbol{\gamma}^\mu \left(\not{\mathrm{p}}_1 + m_\ell c\mathbb{1}\right) \overline{\boldsymbol{\gamma}}^\nu \left(\not{\mathrm{p}}_3 + m_\ell c\mathbb{1}\right)\right] \\ &= 2\big[p_1^\mu p_3^\nu + p_1^\nu p_3^\mu + \eta^{\mu\nu}[m_\ell^2 c^2 - (\mathrm{p}_1\cdot\mathrm{p}_3)]\big]. \end{aligned} \tag{5.215}$$

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

The tensors $X^{\mu\nu}$ and $X_{\mu\nu}$ were computed for point-like (elementary) spin-$\frac{1}{2}$ Dirac spinors, but protons are not point-like (elementary) and their structure causes deviations from the results (5.136)–(5.140), which produce the function $K_{\mu\nu}(2,4;p^+)$. These deviations reflect the effects of strong interaction that bind the quarks into the proton. A description of that structure in quantum chromodynamics [☞ Section 6.1] is too complex for a serious analysis here. However, we do know that $K_{\mu\nu}(2,4;p^+)$ is a rank-2 tensor, and may depend only on the 4-vectors $\mathrm{p}_2$, $\mathrm{p}_4$ and $\mathrm{q} := (\mathrm{p}_1 - \mathrm{p}_3) = (\mathrm{p}_4 - \mathrm{p}_2)$ and, of course, the metric $\eta_{\mu\nu}$. Following tradition, we'll use the following 4-vectors: incoming probe 4-momentum, $\mathrm{p} \equiv \mathrm{p}_2$, and the transfer 4-momentum, q, and write $\mathrm{p}_4 = (\mathrm{p}_2 + \mathrm{q})$. It is therefore possible to parametrize this corrected tensor as

$$K_{\mu\nu}(2,4;p^+) = -K_1\,\eta_{\mu\nu} + \frac{K_2}{M^2c^2}\,p_\mu p_\nu + \frac{K_4}{M^2c^2}\,q_\mu q_\nu + \frac{K_5}{M^2c^2}\,(p_\mu q_\nu + q_\mu p_\nu), \tag{5.216}$$

where $K_i$ are functions of the only scalar variable,[23] $\mathrm{q}^2 = (\mathrm{p}_4 - \mathrm{p}_2)^2$. Since $X^{\mu\nu}(1,3;\ell)$ is a symmetric tensor (5.139a), the antisymmetric part in $K_{\mu\nu}(2,4;p^+)$ – if it even exists – does not contribute to the expression (5.214). This restricts the expansion (5.216) to be symmetric with respect to the $\mu \leftrightarrow \nu$ exchange.

Next, it may be shown [243, p. 277] and [257, Sections 8.2–8.3] that $q^\mu K_{\mu\nu} = 0$, so that [☞ Exercise 5.4.1]

$$K_4 = \frac{M^2c^2}{\mathrm{q}^2}K_1 + \tfrac{1}{4}K_2 \qquad \text{and} \qquad K_5 = \tfrac{1}{2}K_2. \tag{5.217}$$

Thus, $K^{\mu\nu}$ may be parametrized by only two form-factors:

$$K^{\mu\nu}(2,4;p^+) = -K_1(\mathrm{q}^2)\Big(\eta^{\mu\nu} - \frac{q^\mu q^\nu}{\mathrm{q}^2}\Big) + \frac{K_2(\mathrm{q}^2)}{M^2c^2}\Big(p^\mu + \tfrac{1}{2}q^\mu\Big)\Big(p^\nu + \tfrac{1}{2}q^\nu\Big). \tag{5.218}$$

Combining the results (5.214), (5.215) and (5.218), we arrive at

$$\Big\langle |\mathfrak{M}_{\ell p\to\ell p}|^2 \Big\rangle = \Big(\frac{2g_e{}^2}{\mathrm{q}^2}\Big)^2\Big[K_1[(\mathrm{p}_1\cdot\mathrm{p}_3) - 2m_\ell^2c^2] + K_2\Big(\frac{(\mathrm{p}_1\cdot\mathrm{p})(\mathrm{p}_3\cdot\mathrm{p})}{M^2c^2} + \frac{\mathrm{q}^2}{4}\Big)\Big] \tag{5.219}$$

$$\approx \frac{g_e{}^4c^2}{4EE'\sin^4(\theta/2)}\left(2K_1\sin^2\left(\frac{\theta}{2}\right) + K_2\cos^2\left(\frac{\theta}{2}\right)\right) \tag{5.220}$$

where we switched to the lab frame, where the proton is initially at rest, $\mathrm{p} = (Mc,0,0,0)$, the lepton has initial energy $E$ and is deflected, with energy $E'$, at an angle $\theta$ from its initial direction of motion. We have also assumed that $E, E' \gg m_\ell c^2$, and have approximated $m_\ell \approx 0$. Then $\mathrm{p}_1 = E(1,\hat{p}_i)/c$, $\mathrm{p}_3 = E(1,\hat{p}_f)/c$, and $\hat{p}_i\cdot\hat{p}_f = \cos\theta$.

The outgoing lepton energy $E'$ is kinematically determined:

$$E' = \frac{E}{1 + (2E/Mc^2)\sin^2(\frac{\theta}{2})}. \tag{5.221}$$

Besides, in the approximation $m_\ell \approx 0$ we have the result (3.160), and so

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \Big(\frac{\alpha\hbar}{4ME\sin^2(\frac{\theta}{2})}\Big)^2\frac{E'}{E}\left(2K_1\sin^2\left(\frac{\theta}{2}\right) + K_2\cos^2\left(\frac{\theta}{2}\right)\right)$$

$$= \Big(\frac{\alpha\hbar}{4ME\sin^2\left(\frac{\theta}{2}\right)}\Big)^2\frac{2K_1\sin^2(\frac{\theta}{2}) + K_2\cos^2(\frac{\theta}{2})}{1 + (2E/Mc^2)\sin^2(\frac{\theta}{2})}, \tag{5.222}$$

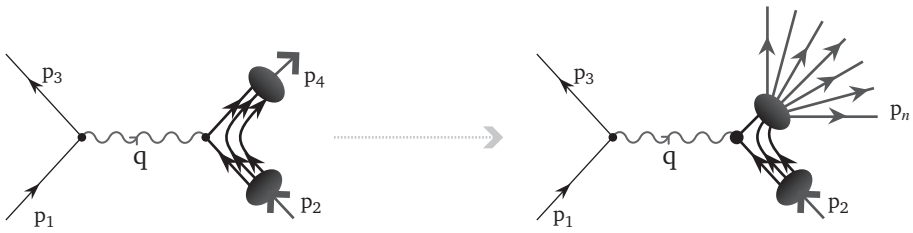
[23] Indeed, $\mathrm{p}_2^2 = \mathrm{p}_4^2 = M^2c^2$ is a constant, and $\mathrm{q}\cdot\mathrm{p}_2 = -\frac{1}{2}\mathrm{q}^2$. Also, $K_3$ is the standard notation for the term that appears in the analysis of neutrino–proton collisions, but not for electrically charged leptons.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

which is the so-called (Marshall Nicholas) Rosenbluth formula from 1950. By measuring the angular dependence of the electrons scattered *elastically* on initially stationary protons (or heavy positive ions), one determines experimentally the form-factors $K_1(\mathrm{q}^2)$ and $K_2(\mathrm{q}^2)$, also known as "structure functions."

**Deep inelastic (light) lepton–hadron scattering**

In the case of the *inelastic* collisions




(5.223)

especially where more than one hadron emerges from the collision, the analysis must be adapted more thoroughly. Fermi's golden rule [☞ p. 113] yields

$$\mathrm{d}\sigma = \frac{\hbar^2 \left\langle |\mathfrak{M}|^2_{\ell p\to \ell X}\right\rangle}{4\sqrt{(\mathrm{p}_1\cdot\mathrm{p}_2)-(m_1m_2c^2)^2}} \prod_{i=3}^{n}\left(\frac{c\,\mathrm{d}^3\vec{p}_i}{(2\pi)^3 2E_i}\right)(2\pi)^4\delta^4\left(\mathrm{p}_1+\mathrm{p}_2-\sum_{j=3}^{n}\mathrm{p}_j\right), \quad (5.224a)$$

$$\left\langle |\mathfrak{M}|^2_{\ell p\to \ell X}\right\rangle = \frac{{g_e}^4}{\mathrm{q}^4} X^{\mu\nu}(1,3;\ell\ \text{(lepton)}) K_{\mu\nu}(2,4;X\ \text{(hadrons)}). \quad (5.224b)$$

If collisions of this type are taken inclusively and we only measure the deflection angle of the scattered lepton and its energy (and so effectively know $\mathrm{p}_3$), the result (5.224) must be summed over all possible hadron results and their momenta, so we have

$$\mathrm{d}\sigma = \frac{4\pi M\hbar^2 {g_e}^4\, X^{\mu\nu}(1,3;\ell)}{4\mathrm{q}^4\sqrt{(\mathrm{p}_1\cdot\mathrm{p}_2)^2-(m_1m_2c^2)^2}}\left(\frac{c\,\mathrm{d}^3\vec{p}_3}{(2\pi)^3 2E_3}\right) W_{\mu\nu}, \quad (5.225)$$

$$W_{\mu\nu} := \frac{1}{4\pi M}\sum_X \int\cdots\int \prod_{i=4}^{n}\left(\frac{c\,\mathrm{d}^3\vec{p}_i}{(2\pi)^3 2E_i}\right)(2\pi)^4\delta^4\left(\mathrm{p}_1+\mathrm{p}_2-\sum_{j=3}^{n}\mathrm{p}_j\right) K_{\mu\nu}(2,4;X). \quad (5.226)$$

For an initially stationary proton, $\mathrm{p}_2 \equiv \mathrm{p} = (-Mc,\vec{0})$ and the incoming lepton energy $E$, we have $\mathrm{p}_1 = (-E/c,\vec{p}_i)$. Therefore,

$$(\mathrm{p}_1\cdot\mathrm{p}_2) = ME, \quad \Rightarrow \quad \sqrt{(\mathrm{p}_1\cdot\mathrm{p}_2)^2-(m_1m_2c^2)^2} = \sqrt{M^2(E^2-m_\ell^2c^4)} \approx ME \quad (5.227)$$

since, in typical experiments of this type and with $\ell = e^\pm$, we have $m_\ell \ll E/c^2$. We therefore approximate $m_\ell \approx 0$, so that $\mathrm{p}_1 = E(-1,\hat{p}_i)/c$ and $\mathrm{p}_3 = E'(-1,\hat{p}_f)/c$. Then,

$$\mathrm{d}^3\vec{p}_f = |\vec{p}_f|^2\mathrm{d}|\vec{p}_f|\mathrm{d}\Omega \approx c^{-3}(E')^2\,\mathrm{d}E'\mathrm{d}\Omega, \quad (5.228)$$

and

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E'\,\mathrm{d}\Omega} = \left(\frac{\alpha\hbar}{c\mathrm{q}^2}\right)^2 \frac{E'}{E}\, X^{\mu\nu}(1,3;\ell)\, W_{\mu\nu}. \quad (5.229)$$

Unlike in elastic collisions, $\mathrm{p}_{\mathrm{tot}} = \sum_{i=4}^{n}\mathrm{p}_i$ with $n > 4$ in inelastic collisions where multiple hadrons emerge, so $\mathrm{p}_{\mathrm{tot}}^2 \neq M^2c^2$ [☞ Tables C.7 on p. 529 and C.8 on p. 529]. There then exists no relation like (5.221) between $E'$ and $E,\theta$ for inelastic collisions; $E'$ is independent of $E$ and $\theta$. The

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

result (5.229) then provides the differential effective cross-section in the span of outgoing lepton energies $[E', E'+\mathrm{d}E']$, as is reasonable for a free and continuous variable $E'$.

The second consequence of $\mathrm{p}_{\mathrm{tot}}^2 \neq M^2c^2$ is that also $\mathrm{q}\cdot\mathrm{p} \neq -\mathrm{q}^2/2$, and one defines the variable

$$x := -\frac{\mathrm{q}^2}{2\mathrm{q}\cdot\mathrm{p}}, \quad x \in [0,1]. \tag{5.230}$$

The general dependence of the form-factor $W_{\mu\nu}$ on the transfer 4-momentum q is parametrized the same way as for $K_{\mu\nu}(2,4;p^+)$, but we now have

$$W^{\mu\nu} = W_1(\mathrm{q}^2,x)\Big(-\eta^{\mu\nu}+\frac{q^\mu q^\nu}{\mathrm{q}^2}\Big)+\frac{W_2(\mathrm{q}^2,x)}{M^2c^2}\Big(p^\mu+\tfrac{1}{2x}q^\mu\Big)\Big(p^\nu+\tfrac{1}{2x}q^\nu\Big), \tag{5.231}$$

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E'\mathrm{d}\Omega} = \Big(\frac{\alpha\hbar}{2ME\sin^2(\frac{\theta}{2})}\Big)^2\frac{E'}{E}\left(2W_1\sin^2\left(\frac{\theta}{2}\right)+W_2\cos^2\left(\frac{\theta}{2}\right)\right). \tag{5.232}$$

The Rosenbluth formula (5.222) is the special case obtained by substituting

$$W_i(\mathrm{q}^2,x) = -\frac{K_i(\mathrm{q}^2)}{2M\mathrm{q}^2}\,\delta(x-1), \qquad i=1,2. \tag{5.233}$$

Note that the $\delta(x-1)$ factor not only formally fixes $x = -\mathrm{q}^2/2\mathrm{q}\cdot\mathrm{p} \stackrel{!}{=} 1$, but also implies the relation (5.221).

Finally, the Rosenbluth formula (5.222) may further be specialized to an ideally point-like (elementary) proton by substituting

$$K_1 \to -\mathrm{q}^2 \qquad \text{and} \qquad K_2 \to 4M^2c^2. \tag{5.234}$$

This idealization is not a bad approximation when the electron energy is sufficiently small and the electron does not come too close to the proton ($\sim 10^{-15}$ m), so that the proton internal structure has negligible influence on the scattering.

#### Experimental verification of the parton model

For the elastic collision $A+B \to A'+B'$, we have in the lab frame

$$\mathrm{p}_1 = (-E/c, \vec{p}_i), \quad \mathrm{p}_2 = (-Mc, \vec{0}), \quad \mathrm{p}_3 = (-E'/c, \vec{p}_f), \quad \mathrm{p}_4 = (-E''/c, \vec{P}_f). \tag{5.235}$$

Then, with $m_A = m_{A'} = m$ and $m_B = m_{B'} = M$,

$$\mathrm{q} = (\mathrm{p}_1-\mathrm{p}_3) = \Big((E'-E)/c, (\vec{p}_i-\vec{p}_f)\Big) = (\mathrm{p}_4-\mathrm{p}_2) = \Big(Mc-E''/c, \vec{P}_f\Big), \tag{5.236}$$

$$\mathrm{q}\cdot\mathrm{p}_2 = M(E-E'), \tag{5.237}$$

$$\mathrm{q}^2 \approx -4\frac{EE'}{c^2}\sin^2\left(\frac{\theta}{2}\right) \qquad \text{when} \qquad mc^2 \ll E, E'. \tag{5.238}$$

Note that $\mathrm{q}^2/c^2$ is proportional to the Mandelstam variable $t$ [☞ definitions (3.62)].

In the late 1960s, James Bjorken computed within the quark model that the expressions

$$F_1(x) := M\,W_1(\mathrm{q}^2,x) \qquad \text{and} \qquad F_2(x) := \frac{-\mathrm{q}^2}{2Mc^2x}W_2(\mathrm{q}^2,x) \tag{5.239}$$

become asymptotically independent of $\mathrm{q}^2$ at very high energies (5.238). Here the values of both $|\mathrm{q}^2|$ and $|\mathrm{q}\cdot\mathrm{p}|$ are large (as compared to $mc^2$), but their ratio (5.230) remains small; $x \in [0,1]$.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

This asymptotic independence from the magnitude of the 4-momentum transfer, $\sqrt{\mathrm{q}^2}$, is called "Bjorken scaling," and was soon confirmed in deep inelastic collisions, mostly of electrons and protons. May it suffice here to mention that this phenomenon confirms that in the deep inelastic collisions the 4-momentum transfer is mostly to one of the three quarks, that these quarks are much smaller than the proton and that they may be treated as point-like particles.

In 1969, Curtis Callan and David Gross proved an additional relationship between Bjorken's functions:

$$2x\,F_1(x) = F_2(x). \tag{5.240}$$

This relation was also quickly confirmed experimentally. Suffice it to say, this relation between $F_1(x)$ and $F_2(x)$ depends on the quark spins, and the Callan–Gross relation (5.240) indicates that quarks have spin $\frac{1}{2}$.

**Digression 5.15** Both relations (5.239) and (5.240) may be derived [☞ Ref. [243], str. 271–277] by treating the quarks as point-like particles, and writing $f_i(x)$ for the probability that the $i$th quark receives the $x$th fraction of the 4-momentum transfer. Using the assumptions (5.233) and (5.234), write

$$W_1^i := \frac{Q_i^2}{2m_i}\delta(x_i{-}1), \quad W_2^i := -\frac{2m_i c^2 Q_i^2}{\mathrm{q}^2}\delta(x_i{-}1), \quad x_i := -\frac{\mathrm{q}^2}{2\mathrm{q}{\cdot}\mathrm{p}_i}, \tag{5.241a}$$

where $m_i$ is the mass of the $i$th quark, and $\mathrm{p}_i$ its 4-momentum. Since the quarks *mostly* move together as the proton, suppose that

$$\mathrm{p}_i =: z_i\,\mathrm{p} \tag{5.241b}$$

is the 4-momentum of the $i$th quark, and equals the $z_i$th fraction of the 4-momentum of the whole proton. It follows that

$$\mathrm{p}_i{}^2 = m_i{}^2 c^2 \quad \text{and} \quad \mathrm{p}^2 = M^2 c^2 \qquad \Rightarrow \qquad m_i = z_i M. \tag{5.241c}$$

So, if $z_i$ varies depending on the dynamics within the proton as a bound state of three quarks, then so do the *effective* quark masses.[24] Relation (5.241b) implies that $x_i = x/z_i$, so that

$$W_1^i = \frac{Q_i^2}{2M}\,\delta(x{-}z_i) \quad \text{and} \quad W_2^i = -\frac{2Mc^2x^2Q_i^2}{\mathrm{q}^2}\,\delta\Big(x{-}z_i\Big). \tag{5.241d}$$

From this, we have

$$W_1 = \sum_i \int_0^1 \mathrm{d}z_i\; W_1^i = \frac{1}{2M}\sum_i Q_i{}^2 f_i(x) \qquad \Rightarrow \quad F_1(x) = \frac{1}{2}\sum_i Q_i{}^2 f_i(x), \tag{5.241e}$$

$$W_2 = \sum_i \int_0^1 \mathrm{d}z_i\; W_2^i = -\frac{2Mc^2}{\mathrm{q}^2}x^2\sum_i Q_i{}^2 f_i(x) \quad \Rightarrow \quad F_2(x) = x\sum_i Q_i{}^2 f_i(x), \tag{5.241f}$$

which agrees with Bjorken's assertion that $F_1(x)$ and $F_2(x)$ are independent of $\mathrm{q}^2$. Besides, the results (5.241e)–(5.241f) clearly also imply the Callan–Gross relation (5.240).

[24] This definition of the masses intuitively takes into account that quarks are bound within the proton, whereby their inertia, i.e., their response to interaction, differs from what it would be were they free particles.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

Besides the facts that:

1. Bjorken (asymptotic) independence of the functions (5.239) from the magnitude of the 4-momentum transfer, $\sqrt{q^2}$, provides experimental confirmation that the proton consists of three electrically charged point-like (elementary) "ingredients,"
2. the Callan–Gross relation (5.240) – also experimentally confirmed – indicates that these "ingredients" have spin $\frac{1}{2}$,

the analysis in Digression 5.15 also leads to the experimental confirmation of the existence of so-called gluons, the particles that mediate the strong nuclear interaction and bind the quarks into bound states, the hadrons.

Indeed, the relation (5.241c) is too naive: quarks are bound within the hadron and their parameters are not directly measurable, whereby the quarks within the hadrons (and free ones do not exist!) are virtual particles. As such, they need not satisfy the equations of motion, i.e., they are not on the mass shell. However, the *average* fraction of 4-momentum carried by the $u$-quarks would have to be about twice larger than the average fraction of 4-momentum carried by the $d$-quarks simply because there are twice as many $u$-quarks as $d$-quarks in the proton, and their masses are approximately the same. Therefore,

$$\int_0^1 \mathrm{d}x\; x\, f_u(x) = 2 \int_0^1 \mathrm{d}x\; x\, f_d(x). \tag{5.242}$$

Using the result (5.241f) as a first approximation, we obtain

$$F_2(x) \approx x\big[(\tfrac{2}{3})^2 f_u(x) + (-\tfrac{1}{3})^2 f_d(x)\big] \quad \overset{(5.242)}{\Longrightarrow} \quad \int_0^1 \mathrm{d}x\; F_2(x) \approx \int_0^1 \mathrm{d}x\; x f_d(x). \tag{5.243}$$

However, the *measured* average values of the form-factor $F_2(x)$ give

$$\int_0^1 \mathrm{d}x\; x f_d(x) \approx 0.18 \quad \text{and} \quad \int_0^1 \mathrm{d}x\; x f_u(x) \approx 0.36. \tag{5.244}$$

The sum of these average values – the fraction of the 4-momentum carried, on average, by *either one of the two u-quarks or the d-quark* – adds up to 0.54. In other words, this analysis indicates that the proton must also contain some electrically neutral "ingredients," which carry 46% of the transfer 4-momentum.

Quantum chromodynamics gives a much better estimate for the form-factors $W_1$ and $W_2$, and then also of Bjorken's functions $F_1(x)$ and $F_2(x)$, and thereby also the probabilities $f_u(x)$ and $f_d(x)$. However, the essence of the conclusion remains unchanged: A non-negligible fraction of the transfer 4-momentum is not carried by the quarks, but by electrically neutral "ingredients" of the proton. These "ingredients" of the proton must interact with the quarks by means of the strong nuclear interaction simply because that is the strongest type of interaction and, being electrically neutral, they cannot interact electromagnetically. The transfer 4-momentum is thus equally fast and uniformly shared among the three quarks, as well as these electrically neutral "ingredients."

On the other hand, the strong nuclear interaction may be described analogously to electromagnetic interactions – by means of a mediating particle. Analogously to the exchange of photons in electromagnetic interactions, strong nuclear interactions are mediated by *gluons*. As the strong nuclear interaction is independent of the electromagnetic interaction and the electric charge of the particles that interact with strong nuclear interactions, it follows that gluons must be electromagnetically neutral. This then permits the electrically neutral "ingredients" of the proton that carry about 46% of the transfer 4-momentum to be identified with gluons[25] [☞ Section 6.1].

[25] This additionally [☞ Digression 5.14 on p. 214] rules out the Han–Nambu model with integrally charged quarks: As those charge assignments are color-dependent, the gluons of the Han–Nambu model must be electrically charged and would also have to interact electromagnetically.

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press

Detailed estimates based on quantum chromodynamics indeed provide a very good agreement with experimental data, and this then is the third significant result of deep inelastic scattering: Besides the experimental confirmation that the proton consists of three point-like spin-$\frac{1}{2}$ electrically charged quarks with charges $Q_u = \frac{2}{3}$ and $Q_d = -\frac{1}{3}$, deep inelastic scattering also confirms experimentally the existence of electromagnetically neutral gluons, which interact with quarks by means of the strong nuclear interaction.

### 5.4.3 Exercises for Section 5.4

✎ ***5.4.1*** *Derive the relation (5.217), using that* $\mathrm{q} := (\mathrm{p}_1 - \mathrm{p}_3) = (\mathrm{p}_4 - \mathrm{p}_2)$ *and* $\mathrm{p}_2^2 = \mathrm{p}_4^2 = -M^2c^2$*, as well as that* $\mathrm{p} := \mathrm{p}_2$ *and* $\mathrm{q}$ *are two linearly independent 4-vectors. (Hint: it should prove useful to first prove that* $\mathrm{q}^2 = -2\,\mathrm{q}\cdot\mathrm{p}$*.)*

✎ ***5.4.2*** *Compare the result (5.229) with the Rutherford (5.160), Mott (5.159) and Rosenbluth (5.222) formulae, as well as the limiting (5.234) case of the latter, and its reduction under the additional condition* $|\mathrm{q}^2| \ll M^2c^2$*. Exhibit the hierarchy of approximations (and their physical meaning) that relate these results.*

✎ ***5.4.3*** *Derive equation (5.221).*

✎ ***5.4.4*** *Derive equation (5.241d).*

https://doi.org/10.1017/9781009291507.008 Published online by Cambridge University Press