Physics in spacetime

High-energy collisions, scatterings and most decays are, for the most part, relativistic quantum processes. It is therefore imperative to recall relativistic kinematics and the basic rules of tensor calculus; see Appendix B.2 for a more complete introduction. However, this chapter neither replaces nor competes with complete treatments of the special theory of relativity such as the illustrative but perfectly detailed text [512] or the *first and original* text [566]. Instead, the purpose of this chapter is to provide an introduction and the results that will be useful in following the subsequent material.

3.1 The Lorentz transformations and tensors

When describing physical processes, one necessarily uses a mathematical apparatus such as a (reference[1]) coordinate system – equipped with a specific and appropriate collection of coordinates. The choice of any one such coordinate system is arbitrary and should not affect the characteristics of the natural laws.

The basic idea (oft cited as one of the two postulates) of Einstein's theory of *relativity* is that the change in the choice of the *coordinate system* and corresponding *coordinates* used to describe spacetime must not change the meaning of natural laws – and so must not have any measurable, i.e., observable consequences.

Digression 3.1 Chapter 5 will show that the so-called gauge principle is simply the generalization of this relativity to the spaces of so-called internal degrees of freedom (also a type of coordinates), such as the phase of the complex wave-function of any charged particle.

In the *special theory of relativity*, this idea of relativity is limited to so-called *inertial* (coordinate) systems, which are usually defined as coordinate systems that differ from each other

[1] The term "coordinate system" is used instead of "reference system" to remind the Reader that its choice necessarily includes a choice of a particular system of four variables, and the specification of the metric tensor in the space of those variables; [☞ Chapter 9 for the general case, and here the relations (3.15)–(3.19)].

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

only in moving one with respect to the other with a constant relative velocity.[2] However, against all intentions, this "definition" does not exclude, for example, a pair of coordinate systems that co-rotate about the same axis with the same angular velocity, but move with a constant relative velocity along the co-rotating axis. Intuitively, the actual intention was to define a class of coordinate systems that move with a constant velocity with respect to a system. . . *at rest!* This shows that our intuition was infiltrated by the Newtonian idea of absolute space and time; it would be inconsistent to define the relativistic inertial frames using Newtonian ideas.

A definition that relies on a relative property between two members of the class being defined must imply that at least one member of this class can be unambiguously identified, so that other members of the class would be defined by comparison with this chosen reference. However, the very *essence* of the theory of relativity is that no such singled-out reference system can exist, which makes such a relative definition essentially insufficient.

The practical property of all inertial coordinate systems in non-relativistic physics is that Newton's first law holds in them. As it was our intention anyway for this law to hold, following Griffiths [243], we adopt it as a definition:

Definition 3.1 *A coordinate system is* ***inertial*** *if Newton's first law, i.e., the law of inertia, holds in it: every body moves at a constant velocity in a straight line if and only if the sum total of all forces that act upon it vanishes.*

Comment 3.1 *It is not hard to show that Definition 3.1 implies that the relative velocity between any two inertial systems is constant. Thus, Definition 3.1 implies the usually assumed property of inertial coordinate systems, which were meant to be selected. It also excludes the non-inertial systems such as the above-mentioned co-axially translating co-rotating systems, which are known to be accelerated.*

Comment 3.2 *Evidently, for a specified coordinate system, we must know what is a "straight line" and what qualifies as a "constant velocity." As the first notion is purely geometrical, and the second requires differential calculus in the specified coordinate system, Definition 3.1 presupposes this level of mathematical knowledge of the specified coordinate system. However, this requirement is logically acceptable and even to be expected. Also, this definition of an inertial system depends on a presupposed familiarity with the concept of force; Chapter 9 will show that amongst* ***all*** *coordinate systems the concept of force may be exchanged for the concept of curvature of the coordinate system. In contradistinction then, all inertial coordinate systems are also* ***flat****, i.e., have no curvature. To be pedantic, one must also require a "trivial topology," i.e., no globally nontrivial features such as multiple connectedness.*

Comment 3.3 *Finally, note that Definition 3.1 implies the testing of certain numerical values: deviation from a straight line, the magnitude and direction of the vector of acceleration. Since every measurement is subject to error, both criteria are subject to the limitations of real, physical measurements. Definition 3.1 is therefore a truly* ***physical*** *one. For example: in nearly all experiments, the "laboratory system" is considered to be inertial, although it is not really so. The laboratory is on the surface of the planet Earth, the gravitational field of which bends the trajectories of objects and accelerates them. Also, the Earth rotates about its axis, so that there are also Coriolis-type forces. Furthermore, the Earth is in the*

[2] It should be kept in mind that the special theory of relativity [179, 55, 69, to name but a few textbooks] is only the linear (*flat* spacetime) approximation to the general theory of relativity [508, 62, 367, 548, 66, 96]. It is possible to extend the use of special relativity so as to include relatively accelerated systems where the incurred nonlinear effects may be consistently neglected. The Reader is expected to have used the formalism of the special theory of relativity at the level of standard texts in electrodynamics [296].

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

gravitational field of the Moon and the Sun; the Earth rotates about the Sun; together with the solar system, they also rotate about the galactic center, etc. For all practical purposes – and to the precision needed in most experiments – these effects are either negligible or can be accounted for by computation. Only if all these (both conceptual and computational) corrections are negligible may the "laboratory system" be regarded as inertial – to within the stated tolerance. The same applies to all other practical applications of Definition 3.1.

Typically cited as the second postulate is the statement that the speed of propagation of light in vacuum, c, is constant. In the "particle system of units" that we adopt herein, c and $\hbar$ are used as basic units, and both quantities are automatically regarded as universal *constants* of Nature.

3.1.1 Space and time mixing

The next step is the realization that *space* and *time* in relativistic physics are not independently specifiable quantities. The transition from one inertial coordinate system S into another inertial coordinate system S', one that moves with the constant velocity $\vec{v}$ with respect to S, is achieved by means of the so-called Lorentz *boosts*[3] [☞ Exercise 3.1.1]:

$$\vec{r}\,' = \vec{r} + (\gamma - 1)(\hat{v}\cdot\vec{r})\,\hat{v} - \gamma\vec{v}t, \qquad \vec{r} = \vec{r}\,' + (\gamma - 1)(\hat{v}\cdot\vec{r}\,')\,\hat{v} + \gamma\vec{v}t', \tag{3.1a}$$

$$t' = \gamma\Big(t - \frac{\vec{v}\cdot\vec{r}}{c^2}\Big), \qquad t = \gamma\Big(t' + \frac{\vec{v}\cdot\vec{r}\,'}{c^2}\Big), \tag{3.1b}$$

$$\gamma := \Big(1 - \frac{\vec{v}^{\,2}}{c^2}\Big)^{-\frac{1}{2}}, \qquad \hat{v} := \frac{\vec{v}}{\sqrt{\vec{v}^{\,2}}}. \tag{3.1c}$$

The inverse transformation (in the right-hand column) is formally identical to the original (in the left-hand column), only with a flipped sign of the relative velocity between the two inertial systems S and S'. Also, note that the formulae for the corresponding Galilean transformation in non-relativistic physics emerge in the formal limit $c \to +\infty$, where $\gamma \to 1$.

It is essential to understand that all relativistic effects stem from boosts (3.1) – which after all are the novelty of Lorentz transformations. For a swift motivation for Lorentz symmetries with the benefit of hindsight of a transpired century, see Digression 8.1 on p. 295. Suffice it to say, Lorentz transformations are the correct symmetry of the Maxwell equations, and therefore also of any matter system that interacts with the electromagnetic field.

Relativity of simultaneity If two events A and B are simultaneous in system S so $t_A = t_B$, they need not be simultaneous in system S':

$$t'_i = \gamma\Big(t_i - \frac{\vec{v}\cdot\vec{r}_i}{c^2}\Big), \quad i = A, B, \qquad \Rightarrow \qquad t'_A - t'_B = \gamma\,\frac{\vec{v}\cdot(\vec{r}_B - \vec{r}_A)}{c^2}, \tag{3.2}$$

which vanishes only if $\vec{v}$ is orthogonal to the difference vector (the extent) $\vec{r}_B - \vec{r}_A$, but not otherwise.

[3] Herein, "boost" denotes the mathematical change of coordinates from one inertial system into another, and which moves with a *constant* velocity $\vec{v}$ with respect to the former. The physical process implementing this change would of course require acceleration, to which the special theory of relativity is explicitly *not* applicable. By "Lorentz transformation," some earlier texts [326] mean only boosts, which leads to contradictory-sounding statements such as "Lorentz transformations do not form a group": indeed, boosts alone do not form a group, as their combination may also be a simple rotation. To avoid this nonsense, by "Lorentz transformation" I mean an arbitrary element of the so-called Lorentz group, which contains both rotations and boosts; see Appendix A.5.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Relativity of distance/extent Although a tad trivial, notice that by the length of an object (as measured in an inertial coordinate system S) we mean the extent between the positions of the end-points (A and B) of that object, $L = |\triangle\vec{r}| = |\vec{r}_B - \vec{r}_A|$, as measured *simultaneously*. Since simultaneity is not absolute – see equation (3.2) – neither can we expect length to be.

Consider the two positions $\vec{r}_A$ and $\vec{r}_B$ in the system S, spanning the extent $\triangle\vec{r} := (\vec{r}_B - \vec{r}_A)$. Using equation (3.1a) in the inertial system S', this extent measures

$$\triangle\vec{r}\,' = \triangle\vec{r} + (\gamma - 1)(\hat{v}\cdot\triangle\vec{r})\hat{v} - \gamma\,\vec{v}(t_B - t_A). \tag{3.3}$$

If the two positions $\vec{r}_A$ and $\vec{r}_B$ have been established simultaneously in the system S (such as the case of measuring the extent between end-points, i.e., the length of an object), then $t_B - t_A = 0$, and we have that

$$\triangle\vec{r}\,' = \triangle\vec{r} + (\gamma - 1)(\hat{v}\cdot\triangle\vec{r})\hat{v} = \triangle\vec{r}_\perp + \gamma\triangle\vec{r}_\parallel, \tag{3.4}$$

where the special cases are

$$\triangle\vec{r}\,'_\parallel = \gamma\,\triangle\vec{r}_\parallel, \qquad \triangle\vec{r}_\parallel := (\hat{v}\cdot\triangle\vec{r})\,\hat{v}, \tag{3.5a}$$

$$\triangle\vec{r}\,'_\perp = \triangle\vec{r}_\perp, \qquad \triangle\vec{r}_\perp := \vec{r} - (\hat{v}\cdot\triangle\vec{r})\,\hat{v}. \tag{3.5b}$$

Formula (3.5a) is the well-known FitzGerald–Lorentz contraction: For an object (and its system S') that moves lengthwise with velocity $\vec{v}$ with respect to the system S, the measurement of the length of the object in the latter system is $L = \triangle\vec{r}_\parallel = L'/\gamma = \gamma^{-1}\triangle\vec{r}\,'_\parallel$. Since $\gamma \geqslant 1$, it follows that $L \leqslant L'$. In turn, formula (3.5b) shows that there is no FitzGerald–Lorentz contraction in directions perpendicular to the relative velocity of the two coordinate systems.

Relativity of the duration of time Consider two moments of time t'_A and t'_B in the inertial system S', which moves with velocity $\vec{v}$ with respect to the inertial system S. Using equation (3.1a) then gives

$$t_B - t_A = \gamma(t'_B - t'_A) + \gamma\frac{\vec{v}\cdot(\vec{r}\,'_B - \vec{r}\,'_A)}{c^2}. \tag{3.6}$$

If the two moments of time t'_A and t'_B have been measured in the same place in system S' (such as the case when the duration of a localized process is observed within the system S'), then $\vec{r}\,'_A = \vec{r}\,'_B$, and $\triangle t' := t'_B - t'_A$ is the duration of time in this "moving" system S'. Then

$$\triangle t = \gamma\,\triangle t', \tag{3.7}$$

is the well-known time dilation formula: $\triangle t \geqslant \triangle t'$. The S-measurement of the duration of time between the events A and B is longer than measured in system S', where A and B are in the same place.

For elementary particle physics, this effect is priceless: in any system with respect to which they move, particles "live" longer than as measured in the system where they are at rest. Thus, a muon created in the higher layers of the Earth's atmosphere nevertheless arrives at the Earth's surface, although its lifetime is only 2.197 μs in its rest-frame. Equivalently, from the muon's point of view, the trip through the Earth's atmosphere is, owing to the FitzGerald–Lorentz contraction, shorter and allows the muon to arrive at the Earth's surface within its lifetime of only 2.197 μs. This also explains how particles with lifetimes of only $\sim 10^{-23}$ s are nevertheless observable.

Addition of velocities For an object that moves with respect to an inertial system S so that it traverses the extent $\triangle\vec{r}$ during the duration of time $\triangle t$, the (average) velocity is $\vec{u} = \triangle\vec{r}/\triangle t$. In the inertial system S', which moves with the constant velocity $\vec{v}$ with respect to S, for that same object one measures the velocity $\vec{u}\,' = \triangle\vec{r}\,'/\triangle t'$, so that

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

$$\vec{u} := \frac{\triangle\vec{r}}{\triangle t} = \frac{\triangle\vec{r}' + (\gamma-1)(\hat{v}\cdot\triangle\vec{r}')\hat{v} + \gamma\vec{v}\,\triangle t'}{\gamma\big(\triangle t' + \frac{(\vec{v}\cdot\triangle\vec{r}')}{c^2}\big)} = \frac{\gamma^{-1}\vec{u}' + (1-\gamma^{-1})(\hat{v}\cdot\vec{u}')\hat{v} + \vec{v}}{\big(1+\frac{(\vec{v}\cdot\vec{u}')}{c^2}\big)}$$
$$= \frac{\vec{u}'_{\parallel} + \vec{v}}{\big(1+\frac{(\vec{v}\cdot\vec{u}')}{c^2}\big)} + \frac{\vec{u}'_{\perp}}{\gamma\big(1+\frac{(\vec{v}\cdot\vec{u}')}{c^2}\big)}, \qquad \text{where} \qquad \vec{u}'_{\parallel} = (\vec{u}'\cdot\hat{v})\hat{v}, \quad \vec{u}'_{\perp}\cdot\hat{v} = 0. \tag{3.8}$$

The first term is the familiar formula for relativistic addition of collinear velocities, and the second term provides the lesser-known $\hat{v}$-orthogonal contribution to the velocity $\vec{u}'$. Notice that the bigger the velocity $\vec{v}$, the bigger the factor γ, and the lesser the contribution from the second (orthogonal) term. It induces an element of rotation – which is a consequence of the algebraic fact that two Lorentz boosts generate a rotation (A.103); see Appendix A.5 for more details.

— ❦ —

As they will be useful, consider the following approximations:

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \approx 1 + \frac{1}{2}\beta^2 + \frac{3}{8}\beta^4 + \frac{5}{16}\beta^6 + O(\beta^8), \qquad \beta := \frac{v}{c} \ll 1; \tag{3.9a}$$
$$\text{or} \qquad \approx \frac{1}{\sqrt{2\epsilon}}\Big[1 + \frac{1}{4}\epsilon + \frac{3}{32}\epsilon^2 + \frac{5}{128}\epsilon^3 + O(\epsilon^4)\Big], \qquad \epsilon := \Big(1 - \frac{|\vec{v}|}{c}\Big) \ll 1; \tag{3.9b}$$
$$\gamma^{-1} = \sqrt{1-\beta^2} \approx 1 - \frac{1}{2}\beta^2 - \frac{1}{8}\beta^4 - \frac{1}{16}\beta^6 + O(\beta^8); \tag{3.9c}$$
$$\text{or} \qquad \approx \sqrt{2\epsilon}\Big[1 - \frac{1}{4}\epsilon - \frac{1}{32}\epsilon^2 - \frac{1}{128}\epsilon^3 + O(\epsilon^4)\Big]. \tag{3.9d}$$

The expansions (3.9a) and (3.9c) are appropriate approximations for small (non-relativistic, $v \ll c$) velocities, while the expansions (3.9b) and (3.9d) are convenient for large (ultra-relativistic, $v \approx c$) velocities.

3.1.2 Spacetime and the index notation

Since the 3-vector $\vec{r}$ (spatial position) and the moment of time t were in the previous section shown to not be independently specifiable quantities, introduce the 4-vector spacetime[4]

$$\mathrm{x} := \sum_{\mu=0}^{3} x^\mu\,\hat{\mathrm{e}}_\mu, \qquad \text{where} \qquad x^0 = ct, \quad \vec{r} = \sum_{i=1}^{3} x^i\hat{\mathrm{e}}_i, \tag{3.10}$$

and where $\hat{\mathrm{e}}_1, \hat{\mathrm{e}}_2, \hat{\mathrm{e}}_3$ are usual unit vectors in some (e.g., Cartesian) inertial coordinate system, and $\hat{\mathrm{e}}_0$ is the additional, fourth unit vector in the direction of time. From now on, we will adopt the strict Einstein convention, whereby summation is implied over any pair of indices precisely if one is a superscript and the other a subscript; thus, $\sum$-symbols are no longer written except for emphasis. Also, Greek indices range over values $0, 1, 2, 3$, while Latin indices are restricted to range over $1, 2, 3$.

[4] It is customary in the literature to denote 4-vectors by a Latin letter without any index or arrow – just like scalars. *Usually*, the context clarifies which of the two is meant; however, without an explicit note, this convention leaves it unclear if a particular "a" denotes a scalar or a 4-vector. Since the purpose of this book is to *introduce* the Reader to the material, "upright" Latin letters will be used for 4-vectors: *herein*, "$\mathrm{a}, \mathrm{b}, \mathrm{c}, \ldots$" denote 4-vectors, while "$a, b, c, \ldots$" are scalars.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Digression 3.2 Note the difference in transformations:

$$\left.\begin{aligned} \mathrm{d}x^\mu &= \left(\frac{\partial x^\mu}{\partial y^\nu}\right) \mathrm{d}y^\nu, && (3.11a)\\ \frac{\partial}{\partial x^\mu} &= \left(\frac{\partial y^\nu}{\partial x^\mu}\right) \frac{\partial}{\partial y^\nu}, && (3.11b) \end{aligned}\right\} \text{mutually reciprocal transformations}$$

when changing coordinates $x^\mu \to y^\mu$. Taking a cue from the transformations (3.11a)–(3.11b), any 4-vectors the components of which transform:

$$A^\mu(\mathrm{x}) = \left(\frac{\partial x^\mu}{\partial y^\nu}\right) A^\nu(\mathrm{y}) \qquad \text{are called } \textbf{contravariant}; \tag{3.11c}$$

$$B_\mu(\mathrm{x}) = \left(\frac{\partial y^\nu}{\partial x^\mu}\right) B_\nu(\mathrm{y}) \qquad \text{are called } \textbf{covariant}. \tag{3.11d}$$

Digression 3.3 Note that the respectively reciprocal transformations automatically imply that combinations such as

$$A^\mu(\mathrm{x})\, B_\mu(\mathrm{x}), \quad A^\mu(\mathrm{x})\, \frac{\partial}{\partial x^\mu}, \quad B_\mu(\mathrm{x})\, \mathrm{d}x^\mu, \quad \text{etc.} \tag{3.12a}$$

are invariant with respect to coordinate transformations. Therefore, sums such as

$$\mathrm{A}(\mathrm{x}) := A^\mu(\mathrm{x})\, \hat{\mathrm{e}}_\mu \quad \text{and} \quad \mathrm{B}(\mathrm{x}) := B_\mu(\mathrm{x})\, \hat{\mathrm{e}}^\mu \tag{3.12b}$$

specify the vectors $\mathrm{A}(\mathrm{x})$ and $\mathrm{B}(\mathrm{x})$ *invariantly*. That is, no matter which coordinate system we select, the components $A^\mu(\mathrm{x})$ and $B_\mu(\mathrm{x})$ will transform oppositely from the basis vectors $\hat{\mathrm{e}}^\mu$ and $\hat{\mathrm{e}}_\mu$, respectively, leaving the expressions (3.12b) invariant; see Comment B.1 on p. 512.

Mathematical literature favors this invariant notation, but we will follow the physics notation, using components specified with respect to an implicitly chosen coordinate system, as done in Digression 3.2. Furthermore, a quick comparison of equations (3.12a) and (3.12b) shows that $\frac{\partial}{\partial x^\mu}$ and $\mathrm{d}x^\mu$, being natural vector quantities in any coordinate system, may well serve as explicit choices of basis vectors $\hat{\mathrm{e}}_\mu$ and $\hat{\mathrm{e}}^\mu$, respectively.

In this 4-vector notation, the general Lorentz transformations may be compactly written as

$$y^\mu = \mathrm{L}^\mu{}_\nu\, x^\nu \qquad \Leftrightarrow \qquad \mathrm{y} = \mathbf{L}\,\mathrm{x} \qquad \Leftrightarrow \qquad \begin{bmatrix} y^0 \\ y^1 \\ y^2 \\ y^3 \end{bmatrix} = \begin{bmatrix} \mathrm{L}^0{}_0 & \mathrm{L}^0{}_1 & \mathrm{L}^0{}_2 & \mathrm{L}^0{}_3 \\ \mathrm{L}^1{}_0 & \mathrm{L}^1{}_1 & \mathrm{L}^1{}_2 & \mathrm{L}^1{}_3 \\ \mathrm{L}^2{}_0 & \mathrm{L}^2{}_1 & \mathrm{L}^2{}_2 & \mathrm{L}^2{}_3 \\ \mathrm{L}^3{}_0 & \mathrm{L}^3{}_1 & \mathrm{L}^3{}_2 & \mathrm{L}^3{}_3 \end{bmatrix} \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix}. \tag{3.13a}$$

Comparing equation (3.13) with (3.1), rewriting as the analogous system of equations, all $4 \times 4 = 16$ matrix elements $\mathrm{L}^\mu{}_\nu$ for concrete boosts may be identified:

$$\mathbf{L} = \begin{bmatrix} \gamma & -\gamma\frac{v_x}{c} & -\gamma\frac{v_y}{c} & -\gamma\frac{v_z}{c} \\ -\gamma\frac{v_x}{c} & 1+(\gamma-1)\frac{v_x{}^2}{\vec{v}^{\,2}} & (\gamma-1)\frac{v_x v_y}{\vec{v}^{\,2}} & (\gamma-1)\frac{v_x v_z}{\vec{v}^{\,2}} \\ -\gamma\frac{v_y}{c} & (\gamma-1)\frac{v_y v_x}{\vec{v}^{\,2}} & 1+(\gamma-1)\frac{v_y{}^2}{\vec{v}^{\,2}} & (\gamma-1)\frac{v_y v_z}{\vec{v}^{\,2}} \\ -\gamma\frac{v_z}{c} & (\gamma-1)\frac{v_z v_x}{\vec{v}^{\,2}} & (\gamma-1)\frac{v_z v_z}{\vec{v}^{\,2}} & 1+(\gamma-1)\frac{v_z{}^2}{\vec{v}^{\,2}} \end{bmatrix}. \tag{3.13b}$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

In the general case, Lorentz transformations also include the familiar rotations in addition to the boosts (3.13) and are represented by constant (independent of spacetime coordinates) 4×4 matrices of unit determinant:

$$\frac{\partial \mathrm{L}^{\mu}{}_{\nu}}{\partial x^{\rho}} = 0, \quad (\mu,\nu,\rho = 0,1,2,3), \qquad \det(\mathbf{L}) = 1. \tag{3.14}$$

Digression 3.4 By comparison, the transformation (3.13) is seen to be the special case of the general case (3.11a), when the matrix $\frac{\partial x^{\mu}}{\partial y^{\nu}} = \mathrm{L}^{\mu}{}_{\nu}$ satisfies the additional conditions (3.14), turning the coordinate change $x^{\mu} \to y^{\nu}$ linear ($y^{\nu} = \mathrm{L}^{\mu}{}_{\nu}\, x^{\nu} + C^{\mu}$) and homogeneous ($C^{\mu} = 0$).

Now, just as the rotation group $SO(3)$ leaves the Euclidean length invariant, general Lorentz transformations leave the quantity

$$(c\,\tau)^2 := c^2\,t^2 - \vec{r}^{\,2} \;=\; c^2\,t^2 - \left[(x^1)^2 + (x^2)^2 + (x^3)^2\right] \tag{3.15}$$

invariant [☞ Appendix A.1.4]. Since c is constant, the quantity τ is also Lorentz-invariant and is called the "proper time." The name stems from the fact that, in any particular inertial system, for any two separate moments in time in the same place we have $\triangle t = t_B - t_A \neq 0$ and $\triangle \vec{r} = \vec{r}_B - \vec{r}_A = 0$, so that

$$\triangle\tau^2 := (t_B - t_A)^2 - c^{-2}\big[\underbrace{(x^1_B - x^1_A)^2 - (x^2_B - x^2_A)^2 - (x^3_B - x^3_A)^2}_{=0}\big] = (t_B - t_A)^2. \tag{3.16}$$

Note that time dilation (3.7) implies that the proper time for any process is always the shortest; in any other inertial system, the duration of that process can only be longer than the proper times or equal to it. Indeed, since $\triangle\tau$ is invariant, in any inertial system where $\triangle\vec{r} \neq 0$, and the events A and B do not happen in the same point in space, $\triangle t$ must be bigger so that $(\triangle t)^2 - c^{-2}(\triangle\vec{r})^2$ remains constant, i.e., invariant with respect to the transformation from that inertial system into the inertial (rest-)system where $\triangle\vec{r} = 0$.

The invariant quantity (3.15) may be more compactly written as[5]

$$c^2\,\tau^2 = \mathrm{x}^2 = \mathrm{x}\cdot\mathrm{x} := x^{\mu}\,\eta_{\mu\nu}\,x^{\nu}. \tag{3.17}$$

An operation "x·y" denotes the (Lorentzian) scalar product of 4-vectors:

Definition 3.2 *For 4-vectors* x *and* y*, the invariant (scalar) product is*

$$\mathrm{x}\cdot\mathrm{y} = x^{\mu}\,\eta_{\mu\nu}\,y^{\nu}. \tag{3.18}$$

The quantity $\mathrm{x}^2 :=$ x·x *is, simply, the "4-vector* x *square." The matrix*

$$\boldsymbol{\eta} \;=\; [\eta_{\mu\nu}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{3.19}$$

is the ***metric tensor*** *– the* ***metric*** *– of the empty (flat) spacetime. The number of positive and negative eigenvalues in the matrix* $[\eta_{\mu\nu}]$ *is called the* ***signature****, and spacetime and its metric are said to have signature* $(1,3)$.

[5] The scalar product of two n-vectors a and b is denoted "$a\cdot b$"; the Reader must understand from the context if this denotes the Euclidean, Lorentzian or some other scalar product. Following this tradition, note that the notation herein is unambiguous, as Euclidean 3-vectors are indicated by an over-arrow and Lorentzian 4-vectors are denoted by "upright" Latin letters. Therefore, $\vec{a}\cdot\vec{b}$ is the Euclidean scalar product, while a·b is Lorentzian.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

With this definition, it is possible to verify that the 4×4 matrices representing the general Lorentz transformations satisfy the $\boldsymbol{\eta}$-orthogonality condition

$$\mathbf{L}^T\boldsymbol{\eta} = \boldsymbol{\eta}\mathbf{L}^{-1}, \quad \text{i.e.,} \quad \mathbf{L}^T\boldsymbol{\eta}\mathbf{L} = \boldsymbol{\eta}, \quad \text{or} \quad \mathrm{L}^{\rho}{}_{\mu}\eta_{\rho\sigma}\mathrm{L}^{\sigma}{}_{\nu} = \eta_{\mu\nu}. \tag{3.20}$$

This generalizes the orthogonality relation $\boldsymbol{R}^T\mathbb{1}\boldsymbol{R} = \mathbb{1}$, satisfied by the usual rotation matrices, where the identity matrix serves as the metric for the Euclidean invariant scalar product, $\vec{r}\cdot\vec{r} = x^i\delta_{ij}x^j$. Just as the rotation group is denoted $SO(3)$, the Lorentz group is then denoted $SO(1,3)$ – reminding us that the signature of the metric $\boldsymbol{\eta}$ used to define the Lorentz-invariant scalar product (3.18) is $(1,3)$ [☞ Appendix A.5].

Also,

Definition 3.3 *A 4-vector* v *in spacetime with the metric tensor* $\eta_{\mu\nu}$ *is called*

$$\textbf{\textit{time-like (temporal)}}, \ \textit{if} \quad \mathrm{v}^2 > 0, \tag{3.21a}$$

$$\textbf{\textit{space-like (spatial)}}, \ \textit{if} \quad \mathrm{v}^2 < 0, \tag{3.21b}$$

$$\textbf{\textit{light-like (null)}}, \ \textit{if} \quad \mathrm{v}^2 = 0. \tag{3.21c}$$

It should be fairly straightforward that the replacement $t \to (it)$ changes the sign of η_{00}, the signature into $(0,4)$, and the boosts (3.13). The qualitative nature of this change is easiest to spot in the special case when the coordinate system is chosen so that $\vec{v} \to v\,\hat{\mathrm{e}}_1$ in the relation (3.13):

$$[\mathrm{L}^{\mu}{}_{\nu}] = \begin{bmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ -\gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \cosh(\phi) & -\sinh(\phi) & 0 & 0 \\ -\sinh(\phi) & \cosh(\phi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\text{hyperbolic ``rotation''}}, \tag{3.22}$$

where we defined the formal variable $\phi := \cosh^{-1}(\gamma)$, so that $v = c\tanh(\phi)$ and $\frac{v}{c}\gamma = \sinh(\phi)$. *[✎ Verify.]* Upon the replacement $t \to it$:

$$[\mathrm{L}^{\mu}{}_{\nu}] \xrightarrow[\varphi=-i\phi]{t\to it} \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) & 0 & 0 \\ \sin(\varphi) & \cos(\varphi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3.23}$$

so that Lorentz boosts in the x^1-direction become

$$\begin{bmatrix} (ict') \\ x'^1 \\ x'^2 \\ x'^3 \end{bmatrix} = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) & 0 & 0 \\ \sin(\varphi) & \cos(\varphi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} (ict) \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \tag{3.24}$$

rotations in the $((ict), x^1)$-plane in the so-called Wick-rotated spacetime $((ict), x^1, x^2, x^3)$. Although Henry Poincaré was the first to notice that the complex transformation

$$(ct, x^1, x^2, x^3) \to ((ict), x^1, x^2, x^3) \tag{3.25}$$

turns the group $SO(1,3)$ of Lorentz transformations[6] into the group of rotations $SO(4)$, this was first used by Hermann Minkowski to restate the Maxwell equations and the special theory

[6] We will see later that the Lorentz group is actually *Spin*$(1,3)$, the double covering of the $SO(1,3)$ group, for spinors to be describable by single-valued spacetime functions [☞ discussion around the relations (5.45)–(5.48)].

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

of relativity into the 4-dimensional notation. This result solidified the physical irreducibility of 4-dimensional *spacetime*, which is why it is often referred to as "Minkowski space." To emphasize the mixed signature of space+time, the term "spacetime" will be used throughout.

Digression 3.5 Following the example of Digression 3.2 on p. 88, the 4-vector with components x^μ is Lorentz-contravariant, whereby the vector with components $x_\nu := (x^\mu \eta_{\mu\nu})$ is Lorentz-*covariant*, as the quantity $x^\mu\, \eta_{\mu\nu}\, x^\nu$ is defined to be Lorentz-invariant:

$$x^\mu \to \widetilde{x}^\mu = \mathrm{L}^\mu{}_\nu\, x^\nu, \quad \text{contravariant 4-vector;} \tag{3.26a}$$

$$\Rightarrow \quad x_\mu := (x^\nu \eta_{\nu\mu}) \to (\widetilde{x}^\nu \widetilde{\eta}_{\nu\mu}) = \widetilde{x}_\mu = \mathrm{L}^{-1}{}_\rho{}^\nu x_\nu, \quad \text{covariant 4-vector.} \tag{3.26b}$$

Here, $\mathrm{L}^{-1}{}_\mu{}^\nu = [\mathbf{L}^{-1}]_\mu{}^\nu = \frac{\partial x^\nu}{\partial \widetilde{x}^\mu}$ are the components of the matrix-inverse of the matrix of Lorentz transformations $\mathrm{L}^\mu{}_\nu = [\mathbf{L}]^\mu{}_\nu = \frac{\partial \widetilde{x}^\mu}{\partial x^\nu}$. We then compute, respectively, in the new and in the old coordinates:

$$\widetilde{x}_\mu = \widetilde{\eta}_{\mu\nu}\widetilde{x}^\nu = \widetilde{\eta}_{\mu\nu}\mathrm{L}^\nu{}_\sigma x^\sigma, \quad = \mathrm{L}^{-1}{}_\mu{}^\rho x_\rho = \mathrm{L}^{-1}{}_\mu{}^\rho \eta_{\rho\sigma} x^\sigma. \tag{3.26c}$$

This implies that

$$\widetilde{\eta}_{\mu\nu}\mathrm{L}^\nu{}_\sigma = \mathrm{L}^{-1}{}_\mu{}^\rho \eta_{\rho\sigma}, \quad \text{i.e.,} \quad \widetilde{\eta}_{\mu\nu} = \mathrm{L}^{-1}{}_\mu{}^\rho \eta_{\rho\sigma}\mathrm{L}^{-1}{}_\nu{}^\sigma, \quad \text{i.e.,} \quad \widetilde{\boldsymbol{\eta}} = [\mathbf{L}^{-1}]^T\, \boldsymbol{\eta}\, \mathbf{L}^{-1}. \tag{3.26d}$$

The metric components $\eta_{\mu\nu} = [\boldsymbol{\eta}]_{\mu\nu}$ thus form a twice covariant tensor. However, as $\boldsymbol{\eta}$ and $\widetilde{\boldsymbol{\eta}}$ are numerically the same matrix (in the x- and the $\widetilde{\mathrm{x}}$-coordinate system, respectively), they are Lorentz-invariants, i.e., remain unchanged under Lorentz transformations. Then

$$\mathrm{x}\cdot\mathrm{x} \to \widetilde{\mathrm{x}}\cdot\widetilde{\mathrm{x}} = \widetilde{x}^\mu \widetilde{\eta}_{\mu\nu}\widetilde{x}^\nu = (\mathrm{L}^\mu{}_\rho x^\rho)\widetilde{\eta}_{\mu\nu}(\mathrm{L}^\nu{}_\sigma x^\sigma) \overset{(3.26d)}{=} x^\rho\, \mathrm{L}^\mu{}_\rho\, \mathrm{L}^{-1}{}_\mu{}^\nu\, \eta_{\nu\sigma} x^\sigma \tag{3.26e}$$

$$= x^\rho \delta^\nu_\rho \eta_{\nu\sigma} x^\sigma = x^\rho \eta_{\rho\sigma} x^\sigma = \mathrm{x}\cdot\mathrm{x}. \tag{3.26f}$$

Note that the result (3.26d) implies

$$\widetilde{\boldsymbol{\eta}} = [\mathbf{L}^{-1}]^T\, \boldsymbol{\eta}\, \mathbf{L}^{-1} \qquad \Leftrightarrow \qquad \boldsymbol{\eta} = [\mathbf{L}]^T\, \widetilde{\boldsymbol{\eta}}\, \mathbf{L}, \tag{3.26g}$$

so that the Lorentz transformation matrices are η-orthogonal; see equation (3.20). This provides the desired spacetime (Lorentzian) generalization of the more familiar (Euclidean) definition of orthogonal matrices $\mathbb{O}^T\mathbb{1}\mathbb{O} = \mathbb{1}$ by replacing $\mathbb{1} \to \boldsymbol{\eta}$; see Appendix A.5.

Further details on tensor calculus and with arbitrary coordinate systems may be found in Chapter 9 and many books; see Refs. [508, 62, 367, 548, 66, 96], to begin with.

The symbol $\eta^{\mu\nu}$ denotes (the components of) the matrix-inverse to $\eta_{\mu\nu}$, so that

$$\eta^{\mu\nu}\, \eta_{\nu\rho} = \eta_{\rho\nu}\, \eta^{\nu\mu} = \delta^\mu_\rho, \qquad \text{so} \qquad x_\mu := \eta_{\mu\nu}\, x^\nu, \quad x^\mu = \eta^{\mu\nu}\, x_\nu. \tag{3.27}$$

Note that $(x^\mu) = (ct, x^1, x^2, x^3)$ and $(x_\mu) = (ct, x_1, x_2, x_3) = (\eta_{\mu\nu}x^\nu) = (ct, -x^1, -x^2, -x^3)$: the value of the covariant spatial components of a 4-vector have the opposite sign from the values of the contravariant spatial components of the 4-vector.

3.1.3 Mass, energy and linear momentum

The Hamilton action of a free particle is chosen to be proportional to the length of the "worldline," so Hamilton's least action principle would minimize this length. In turn, the worldline can be parametrized by the proper time τ of the same particle:

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

$$S = -\int_{\tau_A}^{\tau_B} \mathrm{d}(c\tau)\, \alpha \overset{(3.7)}{=} -\int_{t_A}^{t_B} \mathrm{d}t\, \frac{\alpha c}{\gamma}, \tag{3.28}$$

where α is some positive constant specific for the considered particle, and the sign is negative so that the resting particle would constitute the *minimum*[7] of the function S, in agreement with Hamilton's least action principle. The expression (3.28) implies that the Lagrangian[8] of a free particle

$$L = -\alpha c\sqrt{1-\frac{v^2}{c^2}} \approx -\alpha c + \frac{1}{2}\alpha c\frac{v^2}{c^2} + \alpha c\, O\Big(\frac{v^4}{c^4}\Big), \tag{3.29}$$

where we used the non-relativistic expansion (3.9c). Since the initial constant, $-\alpha c$, is irrelevant for dynamics, comparing the v^2-term with the one in the non-relativistic expression $L_{\mathrm{NR}} = \frac{1}{2}mv^2$ fixes $\alpha = mc$, and the relativistic Lagrangian of a free particle is determined to be

$$L = -mc^2\gamma^{-1} \;=\; -mc^2\sqrt{1-\frac{\vec{v}^{\,2}}{c^2}} \;=\; -mc^2\sqrt{1-\frac{1}{c^2}|\dot{\vec{r}}|^2}. \tag{3.30}$$

Relativistic momentum and energy From equation (3.30) and using the *canonical definition*, we have

$$\vec{p} := \frac{\partial L}{\partial \dot{\vec{r}}} = \frac{\partial L}{\partial \vec{v}} = m\gamma\vec{v} \overset{(3.9a)}{\approx} m\vec{v} \,+\, \cdots, \tag{3.31}$$

where we dropped the terms that are at least $O(\frac{v^2}{c^2})$ smaller than $m\vec{v}$, and this canonical definition indeed agrees with the usual non-relativistic definitions, for velocities sufficiently smaller than c. Also, the Hamiltonian, i.e., the energy of a free particle, is, by the *canonical definition* ($H = p_i\dot{q}^i - L$),

$$E := \vec{p}\cdot\dot{\vec{r}} - L = m\gamma\,\vec{v}\cdot\vec{v} + mc^2\gamma^{-1} = m\gamma c^2, \tag{3.32a}$$

$$\overset{(3.9a)}{\approx} \underbrace{mc^2}_{\substack{\text{rest}\\ \text{energy}}} + \underbrace{\tfrac{1}{2}m\vec{v}^{\,2}}_{\substack{\text{non-relativ.}\\ \text{kin. energy}}} + \underbrace{\tfrac{1}{2}m\vec{v}^{\,2}\Big[\tfrac{3}{4}\tfrac{\vec{v}^{\,2}}{c^2} + \tfrac{5}{8}\tfrac{\vec{v}^{\,4}}{c^4} + \cdots\Big]}_{\text{relativistic corrections}}. \tag{3.32b}$$

Recall that the energy, by its definition, is a measure of the ability to do work. From the result (3.32a), it follows that a free particle has the ability to do work not only by virtue of its motion (the kinetic energy), but also owing to simply having a nonzero mass! Indeed, the expression (3.32a) clearly expresses energy as a function of velocity, one that does not vanish in the rest-frame of a particle, in which it is of course at rest:

$$E_0 := E\big|_{\vec{v}=0} \;=\; mc^2, \qquad \text{rest energy.} \tag{3.33}$$

The discovery contained in the relation (3.33) is Einstein's best known formula. This is the ideal place to cite Professor Okun's warning [393], that the relation (3.33) – and not "$E = mc^2$" – is the *real* Einstein formula [☞ Exercise 3.1.2].

Of course, the kinetic energy of a particle is then *[✎ why?]*

$$T := E - E_0 = m(\gamma-1)c^2 \approx \underbrace{\tfrac{1}{2}m\vec{v}^{\,2}}_{\substack{\text{non-relativ.}\\ \text{kin. energy}}} + \underbrace{\tfrac{1}{2}m\vec{v}^{\,2}\Big[\tfrac{3}{4}\tfrac{\vec{v}^{\,2}}{c^2} + \tfrac{5}{8}\tfrac{\vec{v}^{\,4}}{c^4} + \cdots\Big]}_{\text{relativistic corrections}}. \tag{3.34}$$

[7] The time between the events A and B is maximal in the system where A and B are in the same place, hence the worldline from A to B is entirely along the time coordinate. In all other systems, the worldline from A to B also extends partially in the spatial directions, and the time $t_B - t_A$ is shorter [☞ time dilation (3.7)].

[8] The term "Lagrangian" and its derivatives honor the French mathematician Joseph Louis Lagrange.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

The energy–momentum 4-vector On par with the spacetime 4-vector $\mathrm{x} = ((ct), x^1, x^2, x^3)$, we define also the 4-momentum [☞ Digressions 3.6 and 3.7]:

$$\mathrm{p} = (p_\mu) := (-E/c, \vec{p}) \;=\; (-m\gamma c, m\gamma\vec{v}). \tag{3.35}$$

From this, we have that

$$\mathrm{p}^2 := p_\mu\,\eta^{\mu\nu}p_\nu = E^2/c^2 - \vec{p}^{\,2} = m^2\gamma^2c^2 - \vec{p}^{\,2} = m^2\gamma^2c^2\Big(1 - \frac{v^2}{c^2}\Big) = m^2c^2. \tag{3.36}$$

As the left-hand side quantity is evidently Lorentz invariant [☞ Exercise 3.1.3], so then is the mass m. Just as proper time is the Lorentz-invariant magnitude of the position 4-vector $\mathrm{x} = (ct, \vec{r})$, (the c-multiple of) mass is the Lorentz-invariant magnitude of the 4-momentum $\mathrm{p} = (-E/c, \vec{p})$. A very useful formula follows from equation (3.36):

$$E^2 \;=\; \vec{p}^{\,2}c^2 \;+\; m^2c^4. \tag{3.37}$$

Rewriting this as $(mc^2)^2 = E^2 - (c\vec{p})^2$ exhibits the direct parallel with equation (3.15). In turn, the 4-momentum is indeed a covariant 4-vector, as defined in equation (3.26b), and its components transform under Lorentz transformations as $p'_\mu = \mathrm{L}^{-1}{}_\rho{}^\nu\, p_\nu$.

Digression 3.6 To justify the definition (3.35) – the covariance and signs of the components (3.35) – it is simplest to rely on quantum mechanics, where in coordinate representation the components of the *operator* of 4-momentum p become $p_\mu = \frac{\hbar}{i}\frac{\partial}{\partial x^\mu}$:

$$p_0 = \frac{\hbar}{i}\frac{\partial}{\partial x^0} = \frac{\hbar}{i}\frac{\partial}{\partial(ct)} = -\frac{1}{c}\,i\hbar\frac{\partial}{\partial t} = -\frac{1}{c}H, \quad \text{but} \quad \vec{p} = +\frac{\hbar}{i}\vec{\nabla}. \tag{3.38}$$

The peculiar negative sign in the identification of $p_0 = -\frac{1}{c}H$ owes to the *standard* identifications in non-relativistic quantum mechanics, $H = i\hbar\frac{\partial}{\partial t}$ *vs.* $\vec{p} = \frac{\hbar}{i}\vec{\nabla}$, and to insisting that the non-relativistic energy operator of a system should be the limit of the relativistic one, *with the same sign*.

Digression 3.7 The same conclusion may also be derived classically, i.e., non-quantum mechanically. Note first that the components of the canonical linear momentum 3-vector are naturally covariant. This is seen from the explicitly written definition (3.31):

$$p_i \;:=\; \frac{\partial L}{\partial v^i}, \quad \text{where} \quad v^i := \frac{\partial x^i}{\partial t}, \qquad i = 1,2,3. \tag{3.39a}$$

To extend this canonical definition to the relativistic 4-vector, use the earlier defined (3.10) 4-vector $(x^\mu) = ((x^0 := ct), x^1, x^2, x^3)$, so that

$$(v^\mu) := \frac{\partial x^\mu}{\partial t} = (c, \dot{x}^1, \dot{x}^2, \dot{x}^3). \tag{3.39b}$$

In turn, the Hamilton action (3.28)–(3.30) may be rewritten as

$$S = -\int_{t_A}^{t_B} \mathrm{d}t\; mc^2\sqrt{1 - \frac{\vec{v}^{\,2}}{c^2}} = \int_{x_A^0}^{x_B^0} \mathrm{d}x^0\; L_0, \qquad L_0 := (L/c) = -m\sqrt{c^2 - \vec{v}^{\,2}}\,, \tag{3.39c}$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

where we note that $[L_0] = \frac{ML}{T}$ has the physical dimensions of linear momentum and not those of energy as does mc^2/γ. From this we have (3.39b):

$$v^0 := \frac{\partial x^0}{\partial t} = \frac{\partial (ct)}{\partial t} = c, \tag{3.39d}$$

as well as that (x^1, x^2, x^3) *depend* on t and so also on $x^0 = ct$:

$$p_\mu := \frac{\partial L_0}{\partial \frac{\partial x^\mu}{\partial x^0}} = \frac{\partial L_0}{\frac{1}{c}\partial \dot{x}^\mu} = c\,\frac{\partial L_0}{\partial v^\mu} \quad \Rightarrow \quad \begin{cases} p_0 := c\,\frac{\partial \left(-m\sqrt{c^2-\vec{v}^2}\right)}{\partial c} = -m\gamma c = -E/c, \\ p_i := c\,\frac{\partial \left(-m\sqrt{c^2-\vec{v}^2}\right)}{\partial v^i} = m\gamma\,\delta_{ij}\,v^j, \end{cases} \tag{3.39e}$$

which reproduces equation (3.35).

By the way, the expression $S = \int \mathrm{d}x^0\, L_0$ of course does not seem to be Lorentz-invariant, since the coordinate x^0 is singled out. However, the spacetime Lagrangian L_0 may be expressed as a spatial integral of the Lagrangian density:

$$L_0 = \int_V \mathrm{d}^3\vec{r}\,\mathscr{L}, \qquad \text{such that} \qquad S = -\int_{(t_A,V)}^{(t_B,V)} \mathrm{d}^4x\,\mathscr{L}, \tag{3.40}$$

where $\mathscr{L}$ is a scalar density: with respect to coordinate change $x^\mu \to y^\mu$, we have that $\mathrm{d}^4x \to \left|\frac{\partial x}{\partial y}\right| \mathrm{d}^4y$, where $\left|\frac{\partial x}{\partial y}\right|$ is the determinant of the matrix of partial derivatives $\frac{\partial x^\mu}{\partial y^\nu}$. For the Hamilton action to be independent of any (invertible) choice and/or change of coordinates, it must be that $\mathscr{L}(x) \to \left|\frac{\partial x}{\partial y}\right|^{-1} \mathscr{L}(y)$, which is the defining property of scalar densities of weight -1 [☞ Section B.2].

Massless particle In non-relativistic physics, a particle with no mass is nonsense: for such a particle both the linear momentum and the kinetic energy would also have to vanish. Then, its response to the action of a force could not be computed by Newton's laws, since the formula $a = \frac{1}{m}F$ would imply that any finite force would cause its infinitely large acceleration. On the other hand, the *relativistic* formulae are self-consistent. Indeed, from the relation (3.36), it follows that

$$m = 0 \quad \Leftrightarrow \quad E^2 = \vec{p}^{\,2}c^2 \quad \Leftrightarrow \quad E = c|\vec{p}|, \tag{3.41}$$

which, when combined with results (3.31) and (3.32a), gives

$$\gamma mc^2 = \gamma mc|\vec{v}| \qquad \Rightarrow \qquad |\vec{v}| = c. \tag{3.42}$$

That is, a massless particle must move at the speed of light. So far, only the photons provide a manifest and directly observable example.

3.1.4 Exercises for Section 3.1

✎ ***3.1.1*** *Simplify the relations (3.1) for the oft-cited case* $\vec{v} = v\,\hat{\mathrm{e}}_z$.

✎ ***3.1.2*** *Without consulting Ref. [393], prove that the equality "*$E = mc^2$*" is nonsense, contradicting the provided definitions and the physical meaning of energy* E *and mass* m.

✎ ***3.1.3*** *Prove that the quantities* $\mathrm{p}^2 := p_\mu\,\eta^{\mu\nu}\,p_\nu$ *and* $\mathrm{x}\cdot\mathrm{p} = x^\mu p_\mu$ *are Lorentz-invariant.*

✎ ***3.1.4*** *Verify the transformations (3.22)–(3.23)–(3.24).*

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

3.2 Relativistic kinematics: limitations and consequences

The essential reason for defining the 4-momentum (3.31) with (3.32a) is the fact that this 4-vector physical quantity is conserved [☞ Footnote *15* on p. 54] and transforms akin to (3.13). Because of the typical application, we will consider collisions and decays.

Using the definitions (3.32a), (3.34) and (3.35), for collisions we have:

1. The sum of relativistic 4-momenta is strictly conserved.
2. The sum of relativistic kinetic energies:
   (a) is conserved in *elastic collisions*;
   (b) grows in "exo-energetic" (fissile or explosive) processes;
   (c) is diminished in "endo-energetic" (fusing, implosive or sticking) processes.

Since the mass equals $(E - T)/c^2$, it is conserved only in elastic collisions. In explosive/fissile collisions, the total mass is diminished, which supports the impression that part of the mass was "converted" into kinetic energy; in implosive/fusing/sticking processes, the total mass grows, as if part of the kinetic energy was "converted" into mass. One must keep in mind that the total mass of a composite system equals (up to the coefficient of proportionality, c^2) the rest energy, which includes various "internal forms of energy," as these are usually called in non-relativistic physics. Thus, e.g., the total mass of a hydrogen atom equals $(m_p + m_e)c^2 + E_b$, where E_b is the binding energy of the hydrogen atom in the particular state, in the first approximation given by Bohr's formula (1.31).

Example 3.1 Two equal snowballs of mass m fly with the same speed $|\vec{v}_i| = \beta c$, $0 < \beta \leqslant 1$, towards each other, then collide and fuse into one large snowball. For what speed of the colliding snowballs will the resulting snowball have a mass of $M = 3m$ (so that "$m + m \to 3m$")?

Solution Given that $\vec{p}_1 = -\vec{p}_2$, conservation of the linear momentum 3-vector gives that $\vec{p}_{1+2} = \vec{p}_1 + \vec{p}_2 = 0$. That is, the resulting snowball is at rest (which should be obvious). Conservation of p_0 now gives $E_A + E_B = E_{A+B}$, i.e.,

$$2m\gamma c^2 = Mc^2 \qquad \Rightarrow \qquad M = \frac{2m}{\sqrt{1-\beta^2}} > 2m, \quad \text{since} \quad \beta > 0. \tag{3.43}$$

Inserting $M = 3m$, solve the equation (3.43) for $\beta = \frac{v_i}{c}$ to obtain $|v_i| = \frac{\sqrt{5}}{3}c \approx 74.54\%\, c$.

— ❦ —

Part of the analysis of this process, the one that relies exclusively on applications of the 4-momentum conservation law is usually referred to as "kinematics." Sometimes, that term also implies the application of the conservation law of angular momentum. For the remainder of this chapter, angular momentum considerations are omitted, and a few simple processes are studied "kinematically" as a user's guide for application in general.

3.2.1 Decays

Two-particle decays

The simplest decay is of the form $A \to B + C$. Label the 4-momenta:

$$A \to B + C, \qquad \mathrm{p}_A = (-m_A c, \vec{0}), \quad \mathrm{p}_B = (-E_B/c, \vec{p}_B), \quad \mathrm{p}_C = (-E_C/c, \vec{p}_C), \tag{3.44}$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

where we used the fact that, before the decay, particle A (with $m_A \neq 0$) defines an inertial system where it is at rest, so that its total relativistic energy reduces to rest energy, $E_A = m_A c^2$. The 4-momentum conservation law provides

$$\mathrm{p}_A = \mathrm{p}_B + \mathrm{p}_C, \qquad \text{or} \qquad \mathrm{p}_B = \mathrm{p}_A - \mathrm{p}_C, \tag{3.45}$$

which includes the usual, 3-momentum conservation:

$$\vec{p}_B = 0 - \vec{p}_C. \tag{3.46}$$

Squaring relation (3.45) for the 4-momenta produces[9]

$$\begin{array}{ccc} \mathrm{p}_B^{\,2} = (\mathrm{p}_A - \mathrm{p}_C)^2 = \mathrm{p}_A^{\,2} + \mathrm{p}_C^{\,2} - 2\,\mathrm{p}_A \cdot \mathrm{p}_C, & & \\ \| & \| & \\ m_B^{\,2}\, c^2 & m_A^{\,2}\, c^2 + m_C^{\,2}\, c^2 - 2\,\dfrac{E_A}{c}\,\dfrac{E_C}{c} = m_A^{\,2}\, c^2 + m_C^{\,2}\, c^2 - 2\, m_A\, E_C. & \end{array} \tag{3.47}$$

From this, it follows that

$$E_C = \left(\frac{m_A^{\,2} + m_C^{\,2} - m_B^{\,2}}{2m_A}\right) c^2. \tag{3.48}$$

The magnitude of the linear momentum 3-vector is now determined from the relation (3.37), $E_C = c\sqrt{m_C^{\,2} c^2 + \vec{p}_C^{\,2}}$:

$$\begin{aligned} |\vec{p}_C| &= \sqrt{\frac{E_C^{\,2}}{c^2} - m_C^{\,2}\, c^2} = c\sqrt{\left(\frac{m_A^{\,2} + m_C^{\,2} - m_B^{\,2}}{2m_A}\right)^2 - m_C^{\,2}} \\ &= c\,\frac{\sqrt{(m_A + m_B + m_C)(m_A - m_B + m_C)(m_A + m_B - m_C)(m_A - m_B - m_C)}}{2m_A} \\ &= c\,\frac{\sqrt{m_A^{\,4} + m_B^{\,4} + m_C^{\,4} - 2m_A^{\,2}m_B^{\,2} - 2m_A^{\,2}m_C^{\,2} - 2m_B^{\,2}m_C^{\,2}}}{2m_A}. \end{aligned} \tag{3.49}$$

From the relation (3.46) it follows that $|\vec{p}_B| = |\vec{p}_C|$, which also follows from the $B \leftrightarrow C$ symmetry of the formula (3.49). The analogous derivation gives $E_B = \left(\frac{m_A^{\,2} + m_B^{\,2} - m_C^{\,2}}{2m_A}\right) c^2$. Note that both E_B and E_C are fully determined by decay kinematics. This was crucial in the discussion on p. 54, and induced Bohr to question the validity of the energy conservation law, and Pauli to predict the neutrino in order to save the energy conservation law. On the other hand, besides the relation (3.46) amongst the magnitudes, there is nothing to determine the direction of $\hat{p}_B = -\hat{p}_C$, which thus remains arbitrary. This implies that, in a large ensemble of $A \to B + C$ decays, the angular distribution of the direction of $\hat{p}_B = -\hat{p}_C$ is expected to be uniform.

Digression 3.8 The same result is obtained starting with equation (3.45), written in Cartesian components, say,

$$E_A = E_B + E_C, \tag{3.50a}$$

$$\vec{0} = \vec{p}_A = \vec{p}_B + \vec{p}_C. \tag{3.50b}$$

[9] The final result, of course, may just as well be obtained by combining the separately stated conservation laws of the linear momentum 3-vector and the relativistic energy. However, by squaring directly the 4-vector equality (3.45), the result (3.48) is obtained faster, because of the simplifying circumstance that three of the components of p_A vanish.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

From the equation (3.50b), it follows that $\vec{p}_C = -\vec{p}_B =: \vec{p}$, and in equation (3.50a), express all three energies in terms of the linear momenta and masses using the general relation (3.37):

$$m_A c^2 = c\sqrt{m_B^2 c^2 + \vec{p}^{\,2}} + c\sqrt{m_C^2 c^2 + \vec{p}^{\,2}}. \tag{3.50c}$$

From here, by squaring, rearranging terms to isolate the square-root, then by squaring again, we obtain *[✎ verify]*

$$\left[m_A^4 + (m_B^2 - m_C^2)^2\right] c^2 = 2m_A^2\left[(m_B^2 + m_C^2)c^2 + 2\vec{p}^{\,2}\right]. \tag{3.50d}$$

Solving this for $|\vec{p}\,|$ one re-derives the result (3.49). *[✎ Verify.]*

Digression 3.9 On the other hand, if we express (in the equation (3.50a)) E_A and one of E_B, E_C in terms of linear momenta and masses using the general relation (3.37), we obtain, e.g.,

$$m_A c^2 = c\sqrt{m_B^2 c^2 + \vec{p}^{\,2}} + E_C, \qquad \text{i.e.,} \qquad m_A c^2 - E_C = c\sqrt{m_B^2 c^2 + \vec{p}^{\,2}}, \tag{3.51a}$$

the square of which gives, after a little simplifying *[✎ verify]*,

$$E_C^2 - 2m_A c^2 E_C + \left[(m_A^2 - m_B^2)c^4 - \vec{p}^{\,2} c^2\right] = 0. \tag{3.51b}$$

After inserting the previous result (3.49) and simplifying, the solutions of this quadratic equations are *[✎ verify]*

$$E_C^{(\pm)} = \left[m_A \pm \frac{m_A^2 + m_B^2 - m_C^2}{2m_A}\right] c^2, \tag{3.51c}$$

where $E_C^{(-)}$ equals the result (3.48).

That the solution $E_C^{(+)}$ is not physical is quickest to see from the special case when $m_B = m_C = 0$, as is the case in the decay $\pi^0 \to 2\gamma$. For this case,

$$E_B^{(+)} = E_C^{(+)} = \frac{3}{2} m_A c^2, \quad \text{and so} \quad m_A c^2 = E_A \overset{(3.50a)}{=} E_B^{(+)} + E_C^{(+)} = 3m_A c^2, \tag{3.51d}$$

which is clearly a contradiction. This leaves $E_C^{(-)}$ in the result (3.51c) as the only consistent solution for the energy of the product in a two-particle decay, confirming the result (3.48).

The technical advantage in using the square of a suitably chosen form of the 4-momentum conservation equation (3.47) is fully understood only through filling in the derivation steps that had been omitted here (mostly, in rearranging and simplifying). The diligent Student is therefore highly recommended to complete these alternate computations.

Many-particle decays

The analysis of a decay of a particle into more than two "fragments" is of course more complicated. However, the starting point is again the 4-momentum conservation, which may be written in any of the following forms:

$$\mathrm{p} = \sum_i \mathrm{p}_i, \qquad \mathrm{p}_i = \mathrm{p} - \sum_{j\neq i} \mathrm{p}_j, \qquad \mathrm{p} - \mathrm{p}_i = \sum_{j\neq i} \mathrm{p}_j, \quad \forall i, \tag{3.52a}$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

$$\mathrm{p}_i + \mathrm{p}_j = \mathrm{p} - \sum_{k\neq i,j} \mathrm{p}_k, \qquad \mathrm{p} - \mathrm{p}_i - \mathrm{p}_j = \sum_{k\neq i,j} \mathrm{p}_k, \quad \forall i,j, \quad \text{etc.} \tag{3.52b}$$

Squaring the 4-vector equations (3.52) in the rest-frame of the decaying particle, where

$$\mathrm{p} = (-E/c, \vec{0}), \qquad \text{so that} \qquad \mathrm{p}^2 = m^2c^2 = E^2/c^2, \tag{3.53}$$

we respectively obtain the equations:

$$\tfrac{1}{2}\big(m^2 - \sum_i {m_i}^2\big)c^4 = \sum_{j>i} \big(E_iE_j - |\vec{p}_i||\vec{p}_j|c^2\cos(\phi_{ij})\big), \tag{3.54a}$$

$$\tfrac{1}{2}\big(m^2 - m_i^2 + \sum_{j\neq i} {m_j}^2\big)c^4 = mc^2\sum_{j\neq i} E_j - \sum_{\substack{j<k\\ j,k\neq i}} \big(E_jE_k - |\vec{p}_j||\vec{p}_k|c^2\cos(\phi_{jk})\big), \quad \forall i, \tag{3.54b}$$

$$\tfrac{1}{2}\big(m^2 + m_i^2 - \sum_{j\neq i} {m_j}^2\big)c^4 = mc^2E_i + \sum_{\substack{j<k\\ j,k\neq i}} \big(E_jE_k - |\vec{p}_j||\vec{p}_k|c^2\cos(\phi_{jk})\big), \quad \forall i, \tag{3.54c}$$

$$\begin{aligned}\tfrac{1}{2}\big(m^2 - m_i^2 - m_j^2 + \sum_{k\neq i,j} {m_k}^2\big)c^4 &= mc^2\sum_{k\neq i,j} E_k + \big(E_iE_j - |\vec{p}_i||\vec{p}_j|c^2\cos(\phi_{ij})\big)\\ &\quad - \sum_{\substack{k<\ell\\ k,\ell\neq i,j}} \big(E_kE_\ell - |\vec{p}_k||\vec{p}_\ell|c^2\cos(\phi_{k\ell})\big), \quad \forall i,j,\end{aligned} \tag{3.54d}$$

$$\begin{aligned}\tfrac{1}{2}\big(m^2 + m_i^2 + m_j^2 - \sum_{k\neq i,j} {m_k}^2\big)c^4 &= mc^2(E_i + E_j)\\ &\quad + \sum_{\substack{k<\ell\\ k,\ell\neq i,j}} \big(E_kE_\ell - |\vec{p}_k||\vec{p}_\ell|c^2\cos(\phi_{k\ell})\big), \quad \forall i,j,\end{aligned} \tag{3.54e}$$

etc.

$$\text{where} \quad E_i^2 = m_i^2c^4 + |\vec{p}_i|^2c^2, \qquad \text{but} \quad E = E_0 = mc^2, \tag{3.54f}$$

using the particular consequences of the general relation (3.37) and also that

$$\mathrm{p}_i{\cdot}\mathrm{p}_j = p_{i\mu}\eta^{\mu\nu}p_{j\nu} = \Big(-\frac{E_i}{c}\Big)\Big(-\frac{E_j}{c}\Big) - \vec{p}_i{\cdot}\vec{p}_j = \frac{E_iE_j}{c^2} - |\vec{p}_i||\vec{p}_j|\cos(\phi_{ij}). \tag{3.55}$$

The combinatorially growing system (3.52)–(3.54) contains more equations than unknowns, which is convenient, as we can select a subset of the equations (3.52)–(3.54) that provides the simplest way to solve for the desired quantities. (Since the relations (3.52) are all just variants of the same equation, it is clear that the system (3.54) cannot be over-determined.)

3.2.2 Scattering

Besides decays, in elementary particle physics one most often considers the scattering of two particles. The 4-momentum conservation here has the general form

$$\mathrm{p}_1 + \mathrm{p}_2 = \sum_{i>2} \mathrm{p}_i, \tag{3.56}$$

which may, of course, be rewritten in several different ways, just like equations (3.52) are different forms of $\mathrm{p} = \sum_i \mathrm{p}_i$ for a decay. Also, collisions may be analyzed either[10]

[10] The term "CM system" stands for "center-of-momentum system" and is defined by the property that the total linear momentum 3-vector vanishes in it, clearly adapting equations (3.53).

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

$$\text{CM system, before:}\quad \mathrm{p}_1+\mathrm{p}_2=\Big(-\frac{E_1}{c}-\frac{E_2}{c}\,,\,\vec{0}\Big),\qquad \text{as}\quad \vec{p}_1+\vec{p}_2=0, \tag{3.57}$$

$$\text{CM system, after:}\quad \sum_{i>2}\mathrm{p}_i=\Big(-\sum_{i>2}\frac{E_i}{c}\,,\,\vec{0}\Big),\qquad \text{as}\quad \sum_{i>2}\vec{p}_i=0, \tag{3.58}$$

or in the target system (choosing, say, target = "2," so that $\mathrm{p}_2=(-m_2c,\vec{0})$):

$$\text{target system, before:}\quad \mathrm{p}_1'+\mathrm{p}_2'=\Big(-\frac{E_1'}{c}-m_2c\,,\,\vec{p}_1'\Big),\qquad \text{as}\quad \vec{p}_2'=\vec{0}, \tag{3.59}$$

$$\text{target system, after:}\quad \sum_{i>2}\mathrm{p}_i'=\sum_{i>2}\Big(-\frac{E_i'}{c}\,,\,\vec{p}_i\Big),\qquad \text{where}\quad \sum_{i>2}\vec{p}_i'=\vec{p}_1'. \tag{3.60}$$

Here, the vertical equality between (3.57)–(3.58) and (3.59)–(3.60) respectively is, of course, the statement of the 4-momentum conservation law.

What's more, by using only Lorentz-invariant expressions (such as squares of 4-vectors), we may combine *both* systems! That is, the 4-vectors in the "vertical" equation (3.57)–(3.58) and the 4-vector on the left-hand side of (3.59)–(3.60) are, of course, not equal, $\mathrm{p}_1+\mathrm{p}_2\neq\mathrm{p}_1'+\mathrm{p}_2'$. However, the squares of these 4-vectors *are* equal – as is the square of any 4-vector – and this provides the continued equality:

$$(\mathrm{p}_1+\mathrm{p}_2)^2=\Big(\sum_{i>2}\mathrm{p}_i\Big)^2=(\mathrm{p}_1'+\mathrm{p}_2')^2=\Big(\sum_{i>2}\mathrm{p}_i'\Big)^2=\dots, \tag{3.61}$$

where "…" denotes similar equalities for the square of the same 4-momentum specified in any other coordinate system that we may choose for its need or convenience.

For two-particle collisions, $A+B\to C+D$, one defines:

Definition 3.4 *Mandelstam's Lorentz-invariant variables:*

$$s:=-(\mathrm{p}_A+\mathrm{p}_B)^2c^2,\qquad t:=-(\mathrm{p}_A-\mathrm{p}_C)^2c^2,\qquad u:=-(\mathrm{p}_A-\mathrm{p}_D)^2c^2. \tag{3.62}$$

These variables are often used in computations as they are Lorentz-invariant; keep in mind, however, that the 4-momentum conservation law and the relation (3.36) produce the linear relation

$$-(s+t+u)=\Big[3\mathrm{p}_A^{\,2}+\mathrm{p}_B^{\,2}+\mathrm{p}_C^{\,2}+\mathrm{p}_D^{\,2}+2\mathrm{p}_A\cdot\underbrace{(\mathrm{p}_B-\mathrm{p}_C-\mathrm{p}_D)}_{-\mathrm{p}_A}\Big]c^2=\sum_{i=A}^{D}\mathrm{p}_i^{\,2}c^2=\sum_{i=A}^{D}m_i^{\,2}c^4. \tag{3.63}$$

In turn, albeit not Lorentz-invariant, (lab-frame) energies and angles $\phi_{ij}:=\arccos(\hat{p}_i\cdot\hat{p}_j)$ are more convenient variables for comparison with experiments.

Fusing collisions

Generalizing Example 3.1 on p. 95, consider the collision of two particles that fuse into a single one, with a specified mass m_C. This process is evidently a time-reversed version of the two-particle decay, so that the computation (3.44)–(3.8) may be used by adapting the notation. However, in this case, instead of the inertial system (3.44), where the end-product is at rest, select the inertial system where the particle B ("target") is at rest:

$$A+B\to C,\qquad \mathrm{p}_A=(-E_A/c,\vec{p}_A),\quad \mathrm{p}_B=(-m_Bc,\vec{0}),\quad \mathrm{p}_C=(-E_C/c,\vec{p}_C). \tag{3.64}$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Conservation of 4-momentum gives

$$\mathrm{p}_C = \mathrm{p}_A + \mathrm{p}_B, \quad \text{i.e.,} \quad \left(-\frac{E_C}{c}\,,\,\vec{p}_C\right) = \left(-\frac{E_A}{c} - m_B c\,,\,\vec{p}_A\right), \tag{3.65}$$

from which it follows that $\vec{p}_C = \vec{p}_A =: \vec{p}$, as well as that $E_C = E_A + m_B c^2$. Squaring the 4-momentum version of equation (3.65) produces, straightforwardly,

$$\mathrm{p}_C{}^2 = \mathrm{p}_A{}^2 + \mathrm{p}_B{}^2 + 2\mathrm{p}_A{\cdot}\mathrm{p}_B, \tag{3.66}$$

$$m_C^2 c^2 = m_A^2 c^2 + m_B^2 c^2 + 2E_A m_B \quad \Rightarrow \quad E_A = \frac{m_C^2 - (m_A^2 + m_B^2)}{2m_B} c^2. \tag{3.67}$$

The same relation can, of course, also be obtained using the conservation of energy and 3-momentum, and the diligent Reader is invited to do so, then compare the relative ease of this computation. *[✎ Do it.]* Since $E = mc^2 + T$, we have the condition

$$T_A = \frac{m_C^2 - (m_A + m_B)^2}{2m_B} c^2. \quad \textit{[✎ Verify.]} \tag{3.68}$$

That is, the "probe" A must have the precisely specified kinetic energy (3.68) for it to fuse with the target; for any other value of T_A, the 4-momentum conservation law strictly forbids the fusing. The process $A + B \to C$ is said to be *kinematically forbidden* except when the relation (3.68) is satisfied.

Process threshold

Following a worked-out example from Ref. [243], we can determine the "threshold" of the reaction (3.56), i.e., the minimal kinetic energy with which the probe "1" must collide with the target "2" for the particles in the product of the process (3.56) to be created. For this minimal energy, the particles in the product of the process (3.56) will have no kinetic energy, and we have that

$$\mathrm{p}_i\big|_{\text{threshold}} = \left(-m_i c\,,\,\vec{0}\,\right), \qquad i > 2, \quad \text{CM system.} \tag{3.69}$$

On the other hand, before the collision, we have equation (3.59). Using the equality of the second and the third term in equation (3.61), we have for the special "threshold" (minimal energy) case

$$\begin{aligned}
\min\left[\left(\mathrm{p}_1' + \mathrm{p}_2'\right)^2\right] &= \Big(\sum_{i>2} \mathrm{p}_i\big|_{\text{threshold}}\Big)^2, \\
\min\left[\mathrm{p}_1{}^2 + \mathrm{p}_2{}^2 + 2\mathrm{p}_1'{\cdot}\mathrm{p}_2'\right] &= \Big(\sum_{i>2} \left(m_i c\,,\,\vec{0}\right)\Big)^2, \\
\min\left[m_1{}^2 c^2 + m_2{}^2 c^2 + 2E_1'\, m_2\right] &= \Big(\sum_{i>2} m_i c\Big)^2, \\
\left(m_1{}^2 + m_2{}^2\right)c^2 + 2\min(E_1')\, m_2 &= \sum_{i,j>2} m_i\, m_j\, c^2.
\end{aligned} \tag{3.70}$$

It follows that the occurrence of the process (3.56) requires

$$E_1' \geqslant \frac{1}{2m_2}\Big[\sum_{i,j>2} m_i m_j - \left(m_1{}^2 + m_2{}^2\right)\Big] c^2, \tag{3.71}$$

and thus

$$T_1' \geqslant \frac{1}{2m_2} \sum_{i,j>2} m_i m_j\, c^2 - \frac{(m_1 + m_2)^2}{2m_2} c^2. \quad \textit{[✎ Verify.]} \tag{3.72}$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Thus, e.g., for the process $X + X \to 3X + \overline{X}$, for any particle X, the first X-particle must collide with the second, stationary X-particle with at least $6\,m_X\,c^2$ of kinetic energy.

This threshold is larger than the naive expectation, whereby the kinetic energy would need to be "only" sufficient to produce the $(3X+\overline{X}) - (X+X) = X+\overline{X}$ particles, i.e., $2m_Xc^2$. The reason for this is the inefficiency of a moving probe collision with a stationary target: Before the collision, the total linear momentum of the incoming probe-X and stationary target-X system is not zero, and must equal the total linear momentum of the $3X+\overline{X}$ system of particles after the collision. Since the linear momentum of the out-coming $3X+\overline{X}$ particles differs from zero, so does the total kinetic energy, which increases the process threshold.

Head-on collisions

In the CM system, where $\vec{p}_1 = -\vec{p}_2 = \vec{p}$, so that $E_1 = E_2 =: E$ if $m_1 = m_2$, we have

$$\begin{aligned} \min\left[(\mathrm{p}_1 + \mathrm{p}_2)^2\right] &= \Big(\sum_{i>2} \mathrm{p}_i\big|_{\text{threshold}}\Big)^2, \\ \Big(-\frac{2\min(E)}{c}\,,\,\vec{0}\Big)^2 &= \Big(\sum_{i>2} m_i c\,,\,\vec{0}\Big)^2 \quad \Rightarrow \quad \min(E) = \tfrac{1}{2}\sum_{i>2} m_i c^2. \end{aligned} \tag{3.73}$$

Since both particles have the same minimal energy (as they are identical) before the collision, and $T = E - mc^2$, we have that

$$\min\left(\sum T_X\right) = \Big(\sum_{i>2} m_i - 2m_X\Big)c^2 \overset{2X\to 3X+\overline{X}}{=} (4-2)m_Xc^2 = 2m_Xc^2, \tag{3.74}$$

exactly as expected naively. Thus, for the $X + X \to 3X + \overline{X}$ process, head-on collisions of two particles that move with the same speed towards each other (as observed in the CM system) are three times as efficient as colliding a probe-X with a stationary target-X. For the head-on collision, the apparatus must provide only $2m_Xc^2$ of energy (m_Xc^2 per X-particle before the collision) to create the $3X + \overline{X}$ system, while colliding a moving X-particle with a stationary X-particle requires providing the moving X-particle $6m_Xc^2$ of energy.

The difference of $4m_Xc^2$ in energy threshold for a probe colliding with a stationary target ensures that the resulting $3X + \overline{X}$ particle system has the kinetic energy required by the conservation of linear momentum, and which is the extra "price" in the kinetic energy of the probe before the collision. No such extra energy is needed in head-on collisions, where *all* of the kinetic energy is thus available to produce new particles – providing the basic rationale for building colliders.

The relative kinetic energy

The other aspect of the efficacy of head-on collisions is the *relative* kinetic energy: The previous section showed that if two X-particles move in the lab coordinate system one against another with the kinetic energy m_Xc^2 each, in the inertial system of one of the two X-particles (wherein it itself is at rest), the other particle moves with a kinetic energy of $6m_Xc^2$. More generally, use the equality of the first and third term in equation (3.61):

$$\begin{aligned} (\mathrm{p}_1 + \mathrm{p}_2)^2 &= (\mathrm{p}_1' + \mathrm{p}_2')^2, \\ \Big(-\frac{E_1 + E_2}{c}\,,\,\vec{0}\Big)^2 &= \Big(-\frac{E_1'}{c} - m_2 c\,,\,\vec{p}_1\Big)^2. \end{aligned}$$

If $m_1 = m_2 = m$, then from $\vec{p}_1 = -\vec{p}_2$ it follows that $E_1 = E_2 =: E$; also, write $E_1' = E'$. Using the results of the previous computations, we arrive at

$$4E^2 = 2mc^2(E' + mc^2), \tag{3.75}$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

or

$$T' = 4T\Big(1 + \frac{T}{2mc^2}\Big), \quad [\text{✎ verify}] \tag{3.76}$$

where the second term is the fast-growing relativistic correction:

T/mc^2 :	1	2	5	10	20	50	100	$\cdots$
T'/mc^2 :	6	16	70	240	880	5,200	20,400	$\cdots$

(3.77)

When two particles of the same mass collide head-on with a kinetic energy of $100\,mc^2$ each, i.e., total $200\,mc^2$, the collision has the same effect as if a particle at rest was hit by another with a kinetic energy of $20{,}400\,mc^2$ – which is 102 *times* more!

3.2.3 Lessons

This is an excellent place to highlight the differences between conserved and invariant quantities:

1. Energy is *conserved but not Lorentz-invariant*:
   (a) The total energy of each particle at any point in time (before, during, after) in a process equals this same quantity at all other points in time.
   (b) Energy (its $-\frac{1}{c}$-multiple) is the 0th component of a 4-vector, and cannot be Lorentz-invariant: It changes – it mixes with the components of $\vec{p}$ – when the observer changes from one inertial coordinate system to another.
2. Mass is *Lorentz-invariant but not conserved*:
   (a) Its value does not change when the observer changes from one inertial coordinate system to another.
   (b) Mass is not conserved, as should be obvious from Example 3.1 on p. 95.

Note that *Lorentz-invariant* means "unchanged under transforming amongst inertial coordinate systems," i.e., with respect to Lorentz transformations, while *conserved* means "unchanged during any isolated process, as *time* passes." That is, the very definition of "conservation" implies a preferred choice of *time*, which cannot possibly be a notion invariant with respect to Lorentz transformations of coordinates.

The relativistic 4-momentum of a particle is *conserved but not Lorentz-invariant* – just like its 0th component, the relativistic energy, as well as its remaining 3 components known also as the "relativistic 3-momentum."

3.2.4 Exercises for Section 3.2

✎ ***3.2.1*** *Using Bohr's formula (1.31), compute the relative difference* $\frac{m(3p)-m(1s)}{m(1s)}$ *between the hydrogen atom mass when it is in a* $3p$ *state (where* $n = 3$, $\ell = 1$, *while* $|m_\ell| \leqslant \ell$ *and* $m_s = \pm\frac{1}{2}$ *are arbitrary) and when it is in a ground,* $1s$ *state (where* $n = 1$, $\ell = 0 = m_\ell$, *while* $m_s = \pm\frac{1}{2}$ *is arbitrary).*

✎ ***3.2.2*** *Compute the relative contribution of the correction (1.33) to the relative difference between the masses computed in Exercise 3.2.1.*

✎ ***3.2.3*** *If a particle of mass* M *at rest decays into two particles of equal masses,* $m_1 = m_2 = m$, *compute the speed with which the particles leave the decay locus. Compute the relative speed with which the resulting particles move away from each other.*

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

✎ ***3.2.4*** *If a particle of mass M at rest decays into two particles of different masses, $m_1 > m_2$, compute the difference between their kinetic energies as a function of only the masses M, m_1, m_2, including the special case when $m_2 = 0$.*

✎ ***3.2.5*** *Show that the system of equations (3.54) reproduces all derived and stated results for the special case of a two-particle decay.*

✎ ***3.2.6*** *Show that for the case of a two-particle decay, the results (3.48)–(3.49) and analogously for $E_B, |\vec{p}_B|$ together with equation (3.54a) produce $\phi_{BC} = 180^\circ$, in agreement with the result obtained using the linear momentum conservation.*

✎ ***3.2.7*** *Show that a free electron can neither absorb nor emit a single photon, i.e., that the simple processes $\gamma + e^- \to e^-$ as well as $e^- \to e^- + \gamma$ are kinematically forbidden [☞ Chapter 3.3 for explanation].*

✎ ***3.2.8*** *Reconsider the fusing collision computations (3.64)–(3.68) and assume that the probe A flies into the target B with total energy E_A and fuses with it. Compute the 4-momentum and mass of the resulting fused object C.*

3.3 Feynman's diagrams and calculus

In the analysis of Section 3.2, as well as in the corresponding exercises and especially in Exercise 3.2.7, it is tacitly assumed that all particles in the analyzed processes can be observed directly, i.e., that all kinematic parameters (mass, energy, linear momentum, angular momentum, etc.) can be measured.

However, that is not always the case.

Recall Conclusion 2.3 on p. 56. Of course, this has to do with the consequence of Heisenberg's indeterminacy principle [☞ Digression 2.7 on p. 73]. That is, for kinematics, this involves the specific relations (1.42) and (1.47):

$$\triangle p_0 \, \triangle x^0 = \triangle E \, \triangle \tau \geqslant \tfrac{1}{2}\hbar, \qquad \triangle p_i \, \triangle x^i \geqslant \tfrac{1}{2}\hbar, \quad i = 1,2,3. \tag{3.78}$$

The indeterminacy principle permits the two-step process

$$\Big(e^- + \gamma \xrightarrow{1+2} e^- + \gamma \Big) = \Big(e^- + \gamma \xrightarrow{1} \big(e^{*-}\big) \xrightarrow{2} e^- + \gamma \Big), \tag{3.79}$$

even if the process "1" and the process "2" were kinematically forbidden – by themselves. Indeed, if the intermediate, "excited" electron, e^{*-}, exists only during a time shorter than

$$\triangle \tau \sim \frac{\hbar}{2(E_\gamma + m_e c^2)}, \tag{3.80}$$

that is, if the time that elapses between process "1" and process "2" is shorter than the one given by the indeterminacy relation (3.80), then the particle e^{*-} cannot possibly be observed directly: It is then possible neither to measure its kinematic parameters, nor to check the 4-momentum conservation.

Thus, the 4-momentum conservation law is neither violated nor *broken*[11]; Heisenberg's indeterminacy relations (3.78) have to do with a fundamental natural limitation of measuring. That is:

[11] It is important to differentiate between these terms. "Violation" typically refers to a particular case, event or process in which a rule, law or symmetry is not satisfied, while "breaking" applies to all cases, events and processes in a particular phase of the system. Also, in the present context, "breaking" usually refers to symmetry breaking, and in cases of gauge symmetry it also refers to the breaking of the corresponding continuity equation and charge conservation, by extension.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Conclusion 3.1 *The 4-momentum conservation law is strict, applies to all processes, and down to the measurement resolution (precision, tolerance) dictated by Heisenberg's indeterminacy principle.*

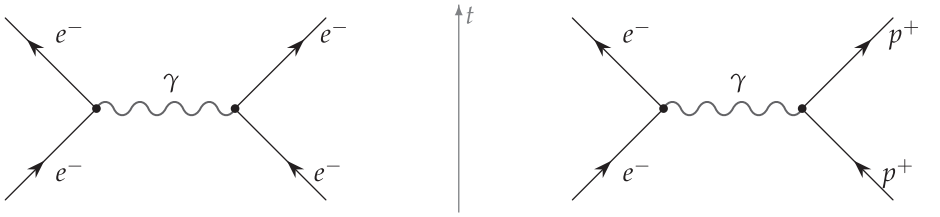
To effectively differentiate the precision of the application of the 4-momentum conservation law, we define:

Definition 3.5 *States of a (system, object, particle, etc.) that cannot be directly observed owing to Heisenberg's indeterminacy* ***principle*** *are called* ***virtual****. Processes that relate real incoming and real outgoing states are real; all others are virtual.*

Comment 3.4 *The processes (3.79) labeled "1" and "2" are virtual, but the process "1+2" is real. [✎ Why?] A virtual particle is also said to be "off-shell," i.e., off the mass shell, which is the hyperboloid* $\mathrm{p}^2 = p_\mu p^\mu = m^2c^2$ *in the 4-momentum space. That is, the 4-momentum of a particle "on-shell" satisfies the relations (3.36)–(3.37), whereas that of a particle "off shell," is not so restricted; to this end, I write* $\mathrm{p}^2 \fallingdotseq m^2c^2$ *– in distinction from "*$\mathrm{p}^2 \neq m^2c^2$*," which means that* p^2 ***must not*** *equal* m^2c^2*. [☞ Tables C.7 on p. 529 and C.8 on p. 529.]*

3.3.1 Diagrams

Processes between particles are naturally represented graphically, by so-called Feynman diagrams.[12] It is important to understand that these diagrams must not be taken as a literal rendition of a process in the "real" space, but as a schematic tool, the primary task of which is to help in the estimation and computation regarding physical processes that they represent. For example, the Feynman diagrams



(3.81)

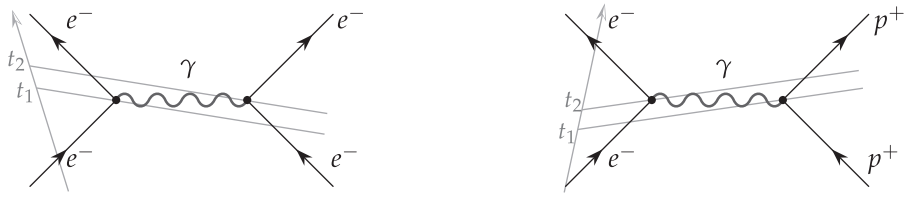
look identical although the left-hand diagram depicts the repulsive effect of (the Coulomb force due to) the exchange of one photon between two electrons, and the right-hand one depicts the attractive effect of (the Coulomb force due to) the exchange of one photon between an electron and a proton.

Except when noted differently, all Feynman diagrams herein are, by convention, drawn with time passing predominantly upward and the lines of simultaneity being oriented predominantly left–right. The tilt (angle with respect to the chosen time axis) of these lines depends on the choice of the observer,[13] which changes the interpretation of the diagram:

[12] The graphical representation of interactions is very intuitive and clear. Feynman certainly did not come up with this idea first, but he did contribute to their popularity as he worked out the technical details that make those diagrams into a useful computational tool. Ernst Stückelberg was the first to use the idea for the individual processes, before Feynman, but had no actual drawings; Freeman Dyson was the first to rigorously establish the link between these diagrams and the well-known perturbative computations. Feynman linked these diagrams to the so-called path integrals, which became a standard only years later.

[13] To be precise, the tilt of all lines changes depending on the observer. However, the tilt of virtual lines – which represent particles that are unobservable in principle and so do not satisfy any classical equation of motion – may change radically, and represent the motion of a massive, light-like, or even tachyonic particle. In distinction, the wave-functions of real particles satisfy their classical equations of motion, and so have the same character for all observers: either massive or light-like. (Or tachyonic – should they ever be experimentally detected [☞ Digression 7.1 on p. 261].)

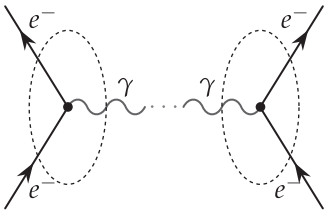
https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press



(3.82)

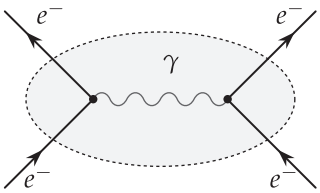
According to the interpretation on the left-hand side, the left-hand particle first emits a photon, which the right-hand particle then absorbs; according to the interpretation on the right-hand side, the right-hand particle emits a photon first, which the left-hand particle then absorbs. Thus we simply speak of an "exchanged" photon, and a diagram such as (3.82) is identified as a schematic representation of this process, and not as a literal, real depiction of the process in spacetime.

The exchanged photon must be virtual (after all, it is by definition never directly observed!), since the individual processes (the left-hand half and the right-hand one)



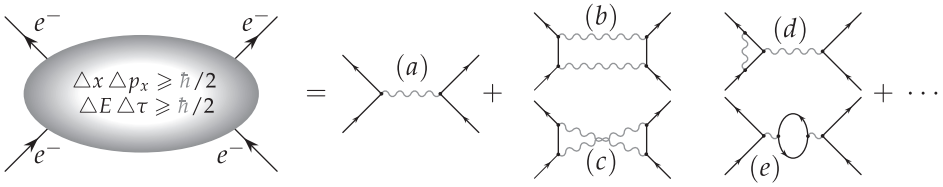
(3.83)

would be kinematically forbidden [☞ Exercise 3.2.7] – while the whole process (3.83) and those in (3.81) are real. However, this implies that processes such as either one of (3.81) must be understood as one of the contributions to the process that may be depicted as



(3.84)

where the schematic region in the shaded ellipse is the *Heisenberg zone*; particles and processes that are entirely within this region can be neither observed nor measured directly as a matter of (Heisenberg's indeterminacy) principle. On the other hand, that also means that within the shaded region of indeterminacy, *all* possible sub-processes may well occur, and in fact do occur [☞ Conclusion 2.3 on p. 56]. It remains to determine the hierarchy of their contributions to the physical quantity being computed for the considered physical process (specified by the particles *outside* the Heisenberg zone of indeterminacy!):



(3.85)

In classical physics, it makes perfect sense to ask: "In a concrete $e^- + e^- \to e^- + e^-$ scattering, which *one* of the processes happened, *either* (a) *or* (b) *or* (c) *or* (d) *or* (e) *or*...?" In quantum physics, this question makes no sense: *As a matter of principle*, not one of the processes shown in the expansion on the right-hand side of the equality (3.85) can possibly be singled out as the

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

"actual" process. *All* the possibilities that satisfy the "boundary conditions" contribute, as (virtual) *sub-processes* of the $e^- + e^- \to e^- + e^-$ scattering. In this context, "boundary conditions" are the data reliably established outside the Heisenberg zone, the region obscured by Heisenberg's indeterminacy relations.

Besides intuitively depicting by graphs the interactive processes between particles, the Feynman diagrams are also a precise instrument for computing probabilities as well as other measurable parameters of the considered process. The goal of every application of Feynman diagrams is the establishment of precise 1–1 correspondences between:

1. the fundamental theory that designs the considered process, usually in terms of a specified Lagrangian density,
2. individual Feynman diagram elements as the graphical representation of individual terms from the specified Lagrangian density,
3. the rules of linking these graphical elements into a complete diagram, as a graphical depiction of the computation with the individual terms from the specified Lagrangian density,
4. the rules of listing all possible Feynman diagrams that need to be included in a computation,
5. the final mathematical expression (usually, in terms of an algebraic sum of various multiple integrals over various 4-momenta), the final result of which is the desired physical quantity,

and finally,

6. the *computation* (or, more often, an *estimate*) of the value of the mathematical expression depicted by the Feynman diagram.

Here, we skip the derivations of the second and third steps in this listing; that would be the task of a field theory course. Instead, we consider some examples [☞ Chapters 5–7] from the Standard Model, to illustrate the application of the last three steps and will only heuristically motivate their relationship to the first step – the construction of appropriate Lagrangians, which however we will discuss at length.

A complete discussion of all aspects of this task is beyond the scope of an *introductory* text such as this. Reference [305] describes the early history of Feynman diagrams and the reasons for the variety of "styles" and conventions in their application; see, e.g., Refs. [61, 474, 537], the field-theory texts [64, 63, 48, 257, 307, 221, 159, 422, 423, 538, 250, 389, 243, 45, 580, 238, 241, 239, 240], as well as those specializing in path-integral methods [459, 165, 123, 277]. However, since the Feynman diagram technique is quite widespread – even in topics well outside elementary particle physics [☞ e.g., Refs. [357, 316]] – we first turn to non-relativistic quantum mechanics, where the well-known perturbative computations are also representable graphically.

3.3.2 Quantum-mechanical digression

As a "warm-up," recall the perturbative computations in non-relativistic quantum mechanics: the relations (1.17)–(1.19) are very often listed and derived in almost all textbooks. Most textbooks also give the basic idea behind the derivation of such oft-used results, but the derivation itself and the results are hardly ever given for corrections of higher order. However, adopting the standard derivation, we write

$$H\,|n\rangle = E_n\,|n\rangle, \qquad \text{where} \qquad H := H_0 + \lambda\, H' \tag{3.86}$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

is the "true" Hamiltonian, given as a sum of a "known" Hamiltonian H_0 and a "perturbation" H', and where λ serves to consistently count the order of perturbation. Suppose that for the "known" system (designated by the Hamiltonian H_0) the complete system of orthonormalized solutions is known:

$$H_0|n;0\rangle = E_n^{(0)}|n;0\rangle, \qquad \begin{cases} \langle n;0|n';0\rangle &= \delta_{n,n'}, \\ \sum_n |n;0\rangle\langle n;0| &= \mathbb{1}, \end{cases} \tag{3.87}$$

and the solutions of equation (3.86) may be found in the analytic form

$$E_n = \sum_{k=0}^{\infty} \lambda^k E_n^{(k)}, \qquad |n\rangle = \sum_{k=0}^{\infty} \lambda^k |n;k\rangle, \tag{3.88}$$

with the normalizations

$$\langle m;k|n;k\rangle = \delta_{mn}, \ \ \forall m,n, \qquad \text{and} \qquad \langle n;k|n;\ell\rangle = \delta_{k\ell}, \ \ \forall k,\ell. \tag{3.89}$$

The energy $E_n^{(k)}$ is the kth order perturbative correction to the original, unperturbed energy $E_n^{(0)}$, and $|n;k\rangle$ is the kth order perturbative correction to the original, unperturbed state $|n;0\rangle$. The treatment of the general situation with (partial) continuous and/or degenerate spectrum is only technically more complicated,[14] and so will not be discussed here.

Introducing the definition

$$\widehat{\Pi}_n^{\alpha} := \sum_{m\neq n} \frac{|m;0\rangle\langle m;0|}{(E_n^{(0)} - E_m^{(0)})^{\alpha}}, \qquad \text{so} \qquad \widehat{\Pi}_n^{\alpha}\,\widehat{\Pi}_n^{\beta} = \widehat{\Pi}_n^{\alpha+\beta}, \tag{3.90}$$

where the superscript in $\widehat{\Pi}_n^{\alpha}$ really behaves like an exponent, the standard recursive formulae[15] for the kth correction to the state and energy are

$$|n;k\rangle = \widehat{\Pi}_n^1 H'|n;k{-}1\rangle - \sum_{i=1}^{k-1} E_n^{(i)}\,\widehat{\Pi}_n^1\,|n;k{-}i\rangle, \qquad k>0, \tag{3.91a}$$

$$E_n^{(k)} = \langle n;0\big|H'\big|n;k{-}1\rangle. \tag{3.91b}$$

The first several iterations of these recursive formulae are:

$$E_n^{(1)} = \langle n;0|H'|n;0\rangle, \tag{3.92a}$$

$$|n;1\rangle = \widehat{\Pi}_n^1 H'|n;0\rangle, \tag{3.92b}$$

$$E_n^{(2)} = \langle n;0\big|H'\,\widehat{\Pi}_n^1\,H'\big|n;0\rangle, \tag{3.92c}$$

$$\begin{aligned} |n;2\rangle &= \widehat{\Pi}_n^1(H' - E_n^{(1)})|n;1\rangle \\ &= \widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle - \widehat{\Pi}_n^1\,\widehat{\Pi}_n^1\,H'|n;0\rangle\langle n;0\big|H'\big|n;0\rangle \\ &= \big[\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1 - \widehat{\Pi}_n^2\,H'|n;0\rangle\langle n;0|\big]H'|n;0\rangle, \end{aligned} \tag{3.92d}$$

[14] The basis of states $|n;k\rangle$ must be redefined so as to eliminate the meaningless terms such as $\frac{\langle m;k|H'|n;k\rangle}{E_m^{(0)}-E_n^{(0)}} \sim \frac{1}{0}$ for $m\neq n$ – which is always possible, by (at least a partial) diagonalization of the perturbation matrix $\langle m;k|H'|n;k\rangle$.

[15] Most quantum mechanics texts list only the results for $E_n^{(1)}$, $|n;1\rangle$ and $E_n^{(2)}$; for a more complete treatment, see e.g., Ref. [362, pp. 685–695].

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

$$
\begin{aligned}
E_n^{(3)} &= \langle n;0|H'|n;2\rangle \\
&= \langle n;0|H'\big[\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1 - \widehat{\Pi}_n^2\,H'|n;0\rangle\langle n;0|\big]\,H'|n;0\rangle \\
&= \langle n;0|H'\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle - \langle n;0|H'\,\widehat{\Pi}_n^2\,H'|n;0\rangle\langle n;0|H'|n;0\rangle
\end{aligned}
\tag{3.92e}
$$

$$
\begin{aligned}
|n;3\rangle &= \widehat{\Pi}_n^1\,\big((H'-E_n^{(1)})|n;2\rangle - E_n^{(2)}|n;1\rangle\big) \\
&= \widehat{\Pi}_n^1\,H'|n;2\rangle - \widehat{\Pi}_n^1\,|n;2\rangle\langle n;0|H'|n;0\rangle - \widehat{\Pi}_n^1\,|n;1\rangle\langle n;0|H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle \\
&= \widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle - \widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^2\,H'|n;0\rangle\langle n;0|H'|n;0\rangle \\
&\quad - \widehat{\Pi}_n^2\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle\langle n;0|H'|n;0\rangle - \widehat{\Pi}_n^2\,H'|n;0\rangle\langle n;0|H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle \\
&\quad + \widehat{\Pi}_n^3\,H'|n;0\rangle\langle n;0|H'|n;0\rangle^2,
\end{aligned}
\tag{3.92f}
$$

and so on. The expressions after (3.92c) indeed become increasingly more and more tedious, and very quickly. However, the particular ordering in these expressions uncovers a simple algorithm for finding all subtractions, in the form of "excisions" from the original expression:

$$
\begin{aligned}
|n;3\rangle &= \widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle && \leftarrow \text{original expression} \\
&\quad - \widehat{\Pi}_n^1\,\underline{[H']}\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle && \leftarrow \text{1st excision} \\
&\quad - \widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,\underline{[H']}\,\widehat{\Pi}_n^1\,H'|n;0\rangle && \leftarrow \text{2nd excision} \\
&\quad - \widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,\underline{[H']}|n;0\rangle && \vdots \\
&\quad - \widehat{\Pi}_n^1\,\underline{[H']}\,\widehat{\Pi}_n^1\,\underline{[H']}\,\widehat{\Pi}_n^1\,H'|n;0\rangle \\
&\quad - \widehat{\Pi}_n^1\,\underline{[H'\,\widehat{\Pi}_n^1\,H']}\,\widehat{\Pi}_n^1\,H'|n;0\rangle \\
&\quad - \widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,\underline{[H'\,\widehat{\Pi}_n^1\,H']}|n;0\rangle
\end{aligned}
\tag{3.93}
$$

where the $\underline{[\cdots]}$-bracketed factors in the original expression are successively "excised." For example,

$$
\widehat{\Pi}_n^1\,\underline{[H']}\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle = \widehat{\Pi}_n^1\,\underbrace{\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle}\,\langle n;0|H'|n;0\rangle,
\tag{3.94a}
$$

$$
\widehat{\Pi}_n^1\,\underline{[H'\,\widehat{\Pi}_n^1\,H']}\,\widehat{\Pi}_n^1\,H'|n;0\rangle = \widehat{\Pi}_n^1\,\underbrace{\widehat{\Pi}_n^1\,H'|n;0\rangle}\,\langle n;0|H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle, \quad \text{etc.}
\tag{3.94b}
$$

The right-most "excisions" in (3.93) vanish:

$$
\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,\underline{[H']}|n;0\rangle = \widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'\,\underbrace{\widehat{\Pi}_n^1\,|n;0\rangle}_{=0}\langle n;0|H'|n;0\rangle,
\tag{3.95a}
$$

$$
\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,\underline{[H'\,\widehat{\Pi}_n^1\,H']}|n;0\rangle = \widehat{\Pi}_n^1\,H'\,\underbrace{\widehat{\Pi}_n^1\,|n;0\rangle}_{=0}\langle n;0|H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle,
\tag{3.95b}
$$

owing to the fact that the normalization (3.89) guarantees

$$
\widehat{\Pi}_n^\alpha\,|n;0\rangle = \sum_{m\neq n}\frac{|m;0\rangle\langle m;0|}{(E_n^{(0)}-E_m^{(0)})^\alpha}|n;0\rangle = \sum_{m\neq n}\frac{1}{(E_n^{(0)}-E_m^{(0)})^\alpha}|m;0\rangle\,\underbrace{\langle m;0|n;0\rangle}_{=0\ (\because\ m\neq n)}.
\tag{3.96}
$$

Since only factors of the form $(H'\,\widehat{\Pi}_n^\alpha\cdots\widehat{\Pi}_n^\beta\,H')$ have a non-vanishing expectation value in the original, "known" state $|n;0\rangle$, only such factors may be "excised." The relations (3.91) may then be written as

$$
|n;k\rangle = (\widehat{\Pi}_n^1\,H')^k|n;0\rangle - \text{all "excisions"}, \qquad k\geqslant 0,
\tag{3.97a}
$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

$$E_n^{(k)} = \langle n;0|H'(\widehat{\Pi}_n^1 H')^{k-1}|n;0\rangle - \text{all "excisions"}. \qquad k \geqslant 1. \tag{3.97b}$$

Using this "excising" notation, e.g., the expression (3.92e) becomes

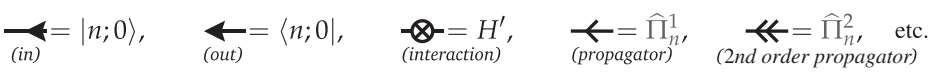
$$\begin{aligned} E_n^{(3)} &= \langle n;0|H'\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle - \langle n;0|H'\,\widehat{\Pi}_n^1\,\underline{[H']}\,\widehat{\Pi}_n^1\,H'|n;0\rangle \\ &= \langle n;0|H'\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle - \langle n;0|H'\,\widehat{\Pi}_n^2\,H'|n;0\rangle\langle n;0|H'|n;0\rangle. \end{aligned} \tag{3.98}$$

The diligent Student is expected to verify *[✎ do it]* that the formulae (3.97a)–(3.97b) reproduce at least the above results (3.92d)–(3.92f).

Digression 3.10 It is not hard to see that the expression (3.98) has no other non-vanishing "excisions." Take, for instance, the candidate

$$\begin{aligned} \langle n;0|\underline{[H']}\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle &:= \langle n;0|\underbrace{\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1}\,H'|n;0\rangle\,\langle n;0|H'|n;0\rangle \\ &= \underbrace{\langle n;0|\,\widehat{\Pi}_n^1}_{=0}\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle\,\langle n;0|H'|n;0\rangle \;=\; 0. \end{aligned} \tag{3.99}$$

The results (3.91) may be depicted graphically, drawing



(3.100)

Then we have

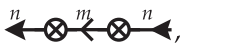
$$\langle n;0|H'|n;0\rangle \overset{(3.92a)}{\longmapsto} E_n^{(1)} =$$

(3.101a)

$$\widehat{\Pi}_n^1\,H'|n;0\rangle \overset{(3.92b)}{\longmapsto} |n;1\rangle =$$

(3.101b)

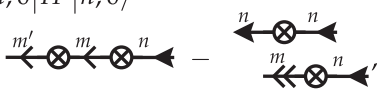
$$\langle n;0|H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle \overset{(3.92c)}{\longmapsto} E_n^{(2)} =$$



(3.101c)

$$\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle - \widehat{\Pi}_n^1\,\widehat{\Pi}_n^1\,H'|n;0\rangle\langle n;0|H'|n;0\rangle$$

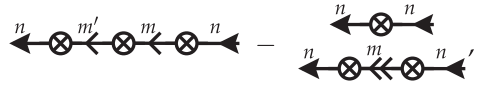
$$\overset{(3.92d)}{\longmapsto} |n;2\rangle =$$



(3.101d)

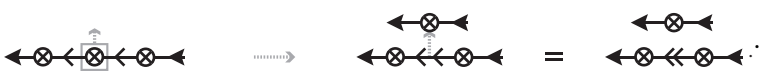
$$\langle n;0|H'\,\widehat{\Pi}_n^1\,H'\,\widehat{\Pi}_n^1\,H'|n;0\rangle - \langle n;0|H'\,\widehat{\Pi}_n^2\,H'|n;0\rangle\langle n;0|H'|n;0\rangle$$

$$\overset{(3.92e)}{\longmapsto} E_n^{(3)} =$$



(3.101e)

and so on, where the subtractions are shown as stacks of diagrams, and represent a product of the corresponding factors. The "excising" algorithm (3.92d) may be graphically depicted also as



(3.102)

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Thus, the whole stationary perturbation theory in non-relativistic quantum mechanics exemplified by (3.101a)–(3.101e) may be written unambiguously and precisely using the graphical symbols (3.100). Similarly, the whole perturbation theory in field theory may be faithfully written in terms of Feynman diagrams.

The detached portions in these "excision" diagrams, such as  in (3.101d)–(3.102), may well be thought of as the quantum-mechanical analogue of "vacuum diagrams" in field theory: These diagrams begin and end at the same state in the Hilbert space, $|n;0\rangle$; these being stationary states, they do not change in time; finally, fixing $|n;0\rangle$ to be the ground state would indeed refer to the "vacuum."

3.3.3 Decays, scattering and calculations

In elementary particle physics, one studies decays, collisions/scatterings, and bound states of these elementary particles. The analysis of bound states uses very successfully the non-relativistic quantum mechanics in Schrödinger's picture, with perturbatively added relativistic corrections [☞ Section 4.1]. On the other hand, decays and collisions/scatterings typically require relativistic analysis. Our goal here will be to estimate the lifetime for the particle A decaying as $A \to B + C$, and the differential as well as the total cross-section (probability) of the $A + B \to C \to A + B$ scattering. The relativistic computations using Feynman diagrams are convenient for this, and we follow the standard approach, adopting Griffiths's conventions [243].

Decays and the half-life

Particles (and even composite systems such as atoms and atomic nuclei) decay probabilistically: It is not possible to specify precisely when a specific particle will decay, but it is possible to determine the average lifetime τ, i.e., half-life, $t_{1/2} = \ln(2)\tau$, where

$$N(t) = N(0)\, e^{-t/\tau} = N(0)\, \left(\tfrac{1}{2}\right)^{t/t_{1/2}} \tag{3.103}$$

is the number of certain particles at time $t > 0$ within a sample where there existed $N(0)$ particles at time $t = 0$. The decay rate (a.k.a. the decay constant) is defined as

$$\Gamma := \frac{1}{\tau} = \frac{\ln(2)}{t_{1/2}}. \tag{3.104}$$

Most particles decay in several ways; in 99.80% cases, π^0 decays into two photons, but in 1.20% cases into an $e^- + e^+$-pair. Other particles have many more "modes" of decay: Each particular decay mode then has a corresponding decay rate Γ_i, and of course

$$\Gamma_{\text{tot}} = \sum_i \Gamma_i, \qquad \tau = \frac{1}{\Gamma_{\text{tot}}}. \tag{3.105}$$

The ratios $\Gamma_i/\Gamma_{\text{tot}}$ are called branching ratios; for the five most significant decay modes of the K^+ particle, $\Gamma_i/\Gamma_{\text{tot}}$ are listed, as percentages, in Table 3.1.

Table 3.1 The significant decay modes of the K^+ meson

$\mu^- + \overline{\nu}_\mu$	$63.44 \pm 0.14\,\%$	$\pi^+ + \pi^0$	$20.92 \pm 0.12\,\%$
$\pi^+ + \pi^+ + \pi^-$	$5.590 \pm 0.031\,\%$	$\pi^0 + e^+ + \nu_e$	$4.98 \pm 0.07\,\%$
$\pi^0 + \mu^+ + \nu_\mu$	$3.32 \pm 0.06\,\%$	plus a dozen or so rare modes	

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Scattering and the effective cross-section

The "effective cross-section" is used to design scattering of one particle on another. Conceptually, this is the literal geometrical description of the target, as seen by the incoming probe. For an archer, the probability of a hit is proportional to the cross-section of the target: that's why swordsmen turn sideways, for the opponent to "see" a smaller cross-section, so as to diminish the probability of a stab.

This literal geometric figure is really faithful [☞ Example 3.2] in the case of "hard" targets, as in the case of a collision of a pool ball, a marble, a cannonball, etc. Such objects have a "binary" interaction: they either collide or they miss. That is, there exists a very well determined *critical* distance, d_c, between the centers of such objects. Should the objects pass by each other so that the distance between them is always bigger than d_c, they do not interact at all. For two regular spheres, d_c equals the sum of their radii.

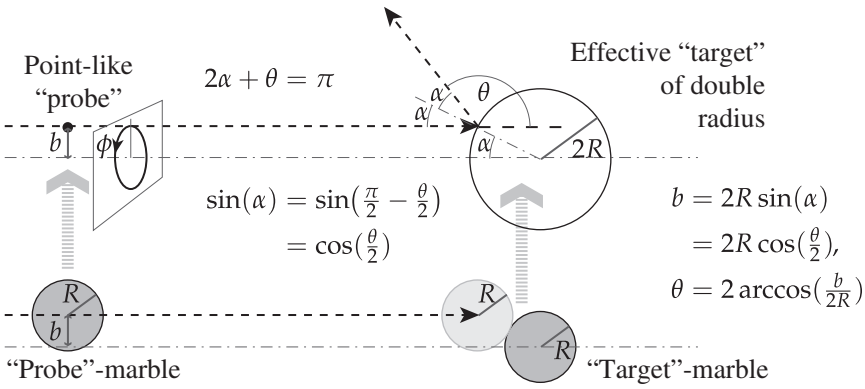


Figure 3.1 The collision of two marbles.

Example 3.2 The classical collision of "hard" marbles of radius R may be analyzed geometrically, as shown in Figure 3.1, where the left-hand marble is replaced by a material point, and the radius of the right-hand marble is doubled. The left-hand marble plays the role of a "probe," and the right-hand one that of the "target." If the orthogonal distance b from the target center is changed a little, $b \to b + \mathrm{d}b$, the scattering angle θ also changes, $\theta \to \theta + \mathrm{d}\theta$. As the collision geometry has axial symmetry, the same result holds if the "probe" approaches the target from any other angle ϕ, so that the "probe" passes through the "surface" element $\mathrm{d}\sigma = |\mathrm{d}b\, b\, \mathrm{d}\phi|$. The out-coming space-angle is then $\mathrm{d}\Omega = |\sin\theta\, \mathrm{d}\theta\, \mathrm{d}\phi|$, so that the ratio

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left|\frac{b}{\sin\theta}\left(\frac{\mathrm{d}b}{\mathrm{d}\theta}\right)\right| = \left|\frac{2R\cos\left(\frac{\theta}{2}\right)}{\sin\theta}\left((2R)\left[-\tfrac{1}{2}\sin\left(\tfrac{\theta}{2}\right)\right]\right)\right| = R^2. \tag{3.106}$$

That produces the total effective cross-section

$$\sigma = \int \mathrm{d}\sigma = \int \mathrm{d}\Omega\, \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = 4\pi R^2 = \pi(2R)^2, \tag{3.107}$$

which is the cross-section of a circle of radius $2R$: Every "probe" the center of which passes through this effective circle of double radius will collide with the "target," all other probes miss.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

The "hard" target *models* evidently cannot hold for scattering of two charged particles, since the electromagnetic interaction extends infinitely far, and the two charged particles always interact, regardless of the smallest distance between them. Of course, the *intensity*, i.e., the force of interaction, diminishes with the square of the distance. However, there is no regime in which the interaction completely vanishes. In distinction from the previous, "hard" targets, such targets are then called "soft."

Molecular forces, which decay as $\sim 1/r^n$ where $n > 2$, as well as forces of Yukawa type (which decay as $\sim e^{-r/r_0}/r^2$) evidently behave between the two limiting cases. The *effective cross-section* is then a measure of the mutual "hardness" of the target and the probe.

Besides, the collision probability may also depend on the nature of the probe as well as the target, of the interaction, and even the number and type of out-coming particles. Indeed, the elastic collision $e^- + p^+ \to e^- + p^+$ is relatively simple at sufficiently small energies. However, at growing energy collisions, we may have

$$e^- + p^+ \to e^- + p^+ + \gamma, \qquad e^- + p^+ \to e^- + p^+ + \pi^0, \qquad e^- + p^+ \to e^- + n^0 + \pi^+,$$
$$\text{and then also} \quad e^- + p^+ \to \nu_e + \Lambda^0 + K^0, \qquad \text{etc.} \tag{3.108}$$

For each of these processes, the exclusive (partial) scattering effective cross-section may be computed, and their sum is then the inclusive (total) scattering effective cross-section.

Finally, the effective cross-section is a measure of the interaction of the "probe" and the target, and must depend on the speed of the "probe": The faster it moves, the less time is available for the interaction, and the effects of the interaction should diminish. Thus, the effective cross-section should depend inversely on the speed of the "probe." In realistic scattering, this dependence of the effective cross-section as a function of speed – or, more frequently, energy – is not so simple: near certain values of speed (i.e., energy) the effective cross-section is significantly amplified. Because of the similarity with the amplification of alternating current when its frequency is near a natural frequency of the circuit, this effect is also called "resonance." In such resonant collisions, the collision energy is just right for the "probe" and the "target" to produce a virtual intermediate state that decays before it could be detected directly [☞ equations (3.67) and (3.48)], and this is the most frequent way of (indirect) observation of new particles.

Example 3.2 on p. 111 shows that the physical meaning of the effective cross-section coincides with the naive measure of interaction – the cross-section of the effective target of doubled radius. In the general case, instead of a point-like "probe" one uses a beam of "probes," of luminosity L, defined as the number of point-like "probes" in unit time and unit area. Thus, we have that

$$\mathrm{d}N = L\,\mathrm{d}\sigma \qquad \Rightarrow \qquad \frac{\mathrm{d}N}{\mathrm{d}\Omega} = L\,\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}, \qquad \text{i.e.,} \qquad \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{L}\frac{\mathrm{d}N}{\mathrm{d}\Omega}. \tag{3.109}$$

This shows that the differential cross-section may be understood as the number of point-like "probes" that reach the detector in the interval of space angles $[\Omega, \Omega + \mathrm{d}\Omega]$, per unit luminosity. The first of these relations (3.109) gives the number of scattered probe-particles expected to be observed in the detectors placed in the interval $[\Omega, \Omega + \mathrm{d}\Omega]$, if the total luminosity of the beam of probes is L; that is the theoretical (computed) result that may be compared with experimental results directly.

For dimensional analysis, and to check the results, note the following relation between Γ and σ: For a decay of a two-particle bound state, Γ must be proportional to the effective cross-section, σ, of the collision of the two particles within the bound state, to the relative speed of these particles, as well as the value of the probability distribution in the place where the particles meet:

1. If the collision effective cross-section vanishes, there is no direct interaction between them, and there can be no decay of their bound state.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

2. If the relative speed of the two particles vanishes, they will never meet, nor interact.
3. If the probability of the two particles to be in the same place vanishes, the direct interaction cannot happen, nor can the decay.

Dimensional analysis in fact even fixes the *linear* dependence on v, σ and $|\Psi(\vec{0},t)|^2$:

> ***Conclusion 3.2*** *The physical units for the decay rate are evidently T^{-1}, and for the effective cross-section (both total and differential) they are L^2. It follows that (see Exercise 3.3.2):*
>
> $$\Gamma \;\propto\; \sigma\, v\, |\Psi(\vec{0},t)|^2. \tag{3.110}$$

Fermi's golden rule

The basic idea of the so-called Fermi's golden rule is that the computation of a physical quantity such as a decay rate or a scattering effective cross-section, both total (inclusive) and exclusive (partial) may be written (up to conventional numerical factors) as a product of two factors:

1. the modulus-squared of the so-called "matrix element," i.e., "amplitude" of the process,
2. the sum/integral over the "phase space" – i.e., over aspects of the process that are *not* being measured/observed, and so do not specify the process.

This approach gives the formulae, cited here from Ref. [243] without derivation:

$$A \to C_1 + C_2 + \cdots \qquad \text{decay :} \tag{3.111}$$

$$\mathrm{d}\Gamma = |\mathfrak{M}|^2\, \frac{S}{2\hbar m_A}\, (2\pi)^4 \delta^4\big(\mathrm{p}_A - \textstyle\sum_i \mathrm{p}_i\big) \prod_j \frac{c\, \mathrm{d}^3\vec{p}_j}{2(2\pi)^3 E_j}, \tag{3.112}$$

where S is the product of "statistical factors," one $(k!)^{-1}$ factor for every group of k identical particles amongst the decay products.

$$A + B \to C_1 + C_2 + \cdots \qquad \text{collision/scattering :} \tag{3.113}$$

$$\mathrm{d}\sigma = |\mathfrak{M}|^2\, \frac{\hbar^2\, S}{4\sqrt{(\mathrm{p}_A{\cdot}\mathrm{p}_B)^2 - (m_A m_B c^2)^2}}\, (2\pi)^4 \delta^4\big(\mathrm{p}_A + \mathrm{p}_B - \textstyle\sum_i \mathrm{p}_i\big) \prod_j \frac{c\, \mathrm{d}^3\vec{p}_j}{2(2\pi)^3 E_j}, \tag{3.114}$$

where, in both results, the energy of the jth particle amongst the process products is a function of the linear momentum:

$$E_j \equiv E_j(\vec{p}_j) = c\sqrt{m_j^2 c^2 + \vec{p}_j^{\,2}}, \tag{3.115}$$

since all particles in these processes are real, i.e., they can be observed directly in detectors, and so are "on-shell," i.e., on the $E^2 = m^2c^4 + \vec{p}^{\,2}c^2$ hyperboloid. In all these formulae, the indices i, j count the process products $(C_1, C_2, \ldots)$, not the components of the linear momenta.

Example 3.3 Consider the two-particle decay $A \to C_1 + C_2$, where the products have masses m_1 and m_2, respectively, and where the linear momenta of the products are not measured, and so must be integrated over. Adapting equation (3.112), we have

$$\Gamma = \frac{S}{2\hbar m_A} \int |\mathfrak{M}|^2\, (2\pi)^4 \delta^4\big(\mathrm{p}_A - \mathrm{p}_1 - \mathrm{p}_2\big)\, \frac{c\, \mathrm{d}^3\vec{p}_1}{2(2\pi)^3 E_1(\vec{p}_1)}\, \frac{c\, \mathrm{d}^3\vec{p}_2}{2(2\pi)^3 E_2(\vec{p}_2)}. \tag{3.116}$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Translating first into the rest-frame of particle A, we have $\mathrm{p}_A = (m_A c, \vec{0})$. The 4-dimensional δ-function factorizes: $\delta^4(\mathrm{p}_A - \mathrm{p}_1 - \mathrm{p}_2) = \delta(m_A c - E_1/c - E_2/c)\delta^3(-\vec{p}_1 - \vec{p}_2)$. Using the 3-dimensional factor that imposes $\vec{p}_2 = -\vec{p}_1$ and cancels the $\mathrm{d}^3\vec{p}_2$-integration, we have

$$\Gamma = \frac{S}{2(4\pi)^2\hbar m_A} \int \mathrm{d}^3\vec{p}_1 \, |\mathfrak{M}|^2 \, \frac{\delta\Big(m_A c - \sqrt{m_1^2c^2 + \vec{p}_1^2} - \sqrt{m_2^2c^2 + (-\vec{p}_1)^2}\Big)}{\sqrt{m_1^2c^2 + \vec{p}_1^2}\sqrt{m_2^2c^2 + (-\vec{p}_1)^2}}. \tag{3.117}$$

On one hand, we know that $|\mathfrak{M}|$ is a Lorentz-invariant (scalar) function of the vectors $\vec{p}_1$ and $\vec{p}_2 = -\vec{p}_1$. So $\mathfrak{M} = \mathfrak{M}(\vec{p}_1, \vec{p}_2) = \mathfrak{M}(\vec{p}_1, -\vec{p}_1) = \mathfrak{M}(\vec{p}_1)$ may depend on the direction of $\vec{p}_1$ only if it also depends on some other (reference) vector quantity in the resulting particles, such as their spin: Then $\mathfrak{M}$ may depend also on the scalars $\vec{p}_i \cdot \vec{\mathcal{S}}_j$ and $\vec{\mathcal{S}}_i \cdot \vec{\mathcal{S}}_j$, where $\vec{\mathcal{S}}_i$ is the (operator of) spin of the ith particle, and $i, j = 1, 2$. In turn, if we may assume that the decay process does not depend on any such additional vector quantities, every scalar function of the vector $\vec{p}_1$ must in fact depend only on the modulus $\rho := |\vec{p}_1|$. It is therefore convenient to use spherical coordinates for the $\mathrm{d}^3\vec{p}_1$-integration. Angular integration gives a factor 4π, and we remain with

$$\Gamma = \frac{S}{8\pi\hbar m_A} \int_0^\infty \frac{\rho^2 \mathrm{d}\rho \, |\mathfrak{M}|^2}{\sqrt{m_1^2c^2 + \rho^2}\sqrt{m_2^2c^2 + \rho^2}} \, \delta\Big(m_A c - \sqrt{m_1^2c^2 + \rho^2} - \sqrt{m_2^2c^2 + \rho^2}\Big). \tag{3.118}$$

To simplify the integral, introduce

$$\mathcal{E} = c\Big(\sqrt{m_1^2c^2 + \rho^2} + \sqrt{m_2^2c^2 + \rho^2}\Big), \tag{3.119}$$

from which it follows that

$$\frac{\mathrm{d}\mathcal{E}}{\mathcal{E}} = \frac{\rho(\mathcal{E})\,\mathrm{d}\rho}{\sqrt{m_1^2c^2 + \rho^2}\sqrt{m_2^2c^2 + \rho^2}} \quad \text{so} \quad \frac{\rho^2 \mathrm{d}\rho}{\sqrt{m_1^2c^2 + \rho^2}\sqrt{m_2^2c^2 + \rho^2}} = \rho(\mathcal{E})\frac{\mathrm{d}\mathcal{E}}{\mathcal{E}}, \tag{3.120}$$

where we intentionally leave $\rho = \rho(\mathcal{E})$ as is, and have

$$\Gamma = \frac{S}{8\pi\hbar m_A} \int_{(m_1+m_2)c^2}^\infty \frac{\mathrm{d}\mathcal{E}}{\mathcal{E}} \, |\mathfrak{M}|^2 \, \rho(\mathcal{E}) \, \delta(m_A c - \mathcal{E}/c). \tag{3.121}$$

Since $\delta(m_A c - \mathcal{E}/c) = c\delta(\mathcal{E} - m_A c^2)$, we finally use the δ-function to cancel the $\mathrm{d}\mathcal{E}$-integral:

$$\Gamma = \begin{cases} \dfrac{S\,\rho_0}{8\pi\hbar m_A^2 c} \, |\mathfrak{M}(\rho_0)|^2, & \text{if} \quad m_A > m_1 + m_2; \\ 0, & \text{if} \quad m_A \leqslant m_1 + m_2, \end{cases} \tag{3.122}$$

where $\rho_0 = |\vec{p}_1|_0$ solves the relation (3.119) with $\mathcal{E} = m_A c^2$:

$$\rho_0 = |\vec{p}_1|_0 = \frac{c}{2m_A}\sqrt{m_A^4 + m_1^4 + m_2^4 - 2m_A^2\, m_1^2 - 2m_A^2\, m_2^2 - 2m_1^2\, m_2^2}, \quad \textit{[✎ verify]} \tag{3.123}$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

and satisfies the linear momentum conservation law. It is useful to list a few simplifications: When the two products have the same mass but are not the same particle, $S = 1$ and

$$\Gamma = \frac{\sqrt{1-(2m/m_A)^2}}{16\pi\hbar m_A}\left|\mathfrak{M}\left(\tfrac{c}{2}\sqrt{m_A^2-(2m)^2}\right)\right|^2. \tag{3.124}$$

If, furthermore, $m_1 = 0 = m_2$ but the resulting particles are still distinct (e.g., a neutrino and a photon, or two *different* neutrinos and where $m \approx 0$ is a pretty good *approximation* for neutrinos), we have

$$\Gamma = \frac{1}{16\pi\hbar m_A}\left|\mathfrak{M}\left(\tfrac{1}{2}m_A c\right)\right|^2. \tag{3.125}$$

Finally, if the products really are identical particles, $S = \frac{1}{2}$, and the decay rate is one half of the previously listed results (3.122), and (3.124)–(3.125).

Example 3.4 Consider the inelastic scattering $A + B \to C_1 + C_2$, where the particles have masses m_A, m_B, m_1 and m_2, respectively, and where the products' linear momenta are not measured, and so must be integrated over. The expression (3.114) must be integrated over $\mathrm{d}^3\vec{p}_1\mathrm{d}^3\vec{p}_2$, and the procedure is similar to that in Example 3.3. However, this time note that $\mathfrak{M}$ in principle depends on all four linear momenta, $\vec{p}_A, \vec{p}_B, \vec{p}_1$ and $\vec{p}_2$. Since $\mathfrak{M}$ is a scalar function, it may depend only on the scalar quantities constructed from these four 3-vectors.

However, if these 3-vectors are expressed in the coordinate system where $\vec{p}_A + \vec{p}_B = \vec{0}$, it follows that $\vec{p}_A = -\vec{p}_B = \vec{p}_i$ (initial) and $\vec{p}_1 = -\vec{p}_2 = \vec{p}_f$ (final). Scalar functions of these two 3-vectors can only depend on $|\vec{p}_i|$, $|\vec{p}_f|$ and $\vec{p}_i\cdot\vec{p}_f = |\vec{p}_i||\vec{p}_f|\cos\vartheta$, where ϑ is the angle between the initial and the final linear momentum, $\vec{p}_A$ and $\vec{p}_1$.

Since the initial linear momentum $\vec{p}_A$ is known, amongst the integration variables in the integral of the expression (3.114), $\mathfrak{M}$ may depend only on $|\vec{p}_1|$ and ϑ. Choosing the spherical coordinate system where $\hat{\mathrm{e}}_z \| \vec{p}_A$, we have

$$\mathrm{d}^3\vec{p}_1 = \rho^2\mathrm{d}\rho\,\sin(\vartheta)\mathrm{d}\vartheta\,\mathrm{d}\varphi = \rho^2\mathrm{d}\rho\,\mathrm{d}\Omega. \tag{3.126}$$

Repeating the simplification of the integral just as done in Example 3.3 and using the result (3.163), we obtain

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{\hbar c}{8\pi}\right)^2 \frac{S\,|\mathfrak{M}|^2}{(E_A+E_B)^2}\frac{|\vec{p}_f|}{|\vec{p}_i|}. \tag{3.127}$$

To compute the final expression for the total effective cross-section, σ, by angular integration over $\mathrm{d}\Omega = \sin(\vartheta)\mathrm{d}\vartheta\,\mathrm{d}\varphi$, the angular dependence of $\mathfrak{M}$ on ϑ, φ must be known.

3.3.4 A simple toy-model example

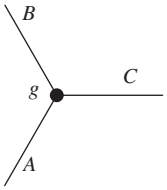
In this section we consider the Feynman calculus in a very simple toy-model. In Chapters 5 and 6, this procedure will be applied to concrete and realistic processes in the Standard Model. This

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

toy-model is "borrowed" from Ref. [243], where it is attributed to Max Dresden; ultimately of course, the number of very simple but nontrivial models is very limited.

— ❦ —

There are only three types of particles in this model, A, B, C, with $m_A > m_B + m_C$, and such that there exists only one elementary process:



(3.128)

We assume that the constant (charge/strength) of interaction g is sufficiently small to serve as a perturbation parameter, at least formally. The computation of any physical quantity is thus organized as a power series in g, and we compute all contributions of order g^n, ranging from the lowest possible value of $n \geqslant 0$, towards increasingly higher values of n.

Procedure 3.1 *The contribution to the matrix element (amplitude) $\mathfrak{M}$ corresponding to a Feynman diagram in the ABC-model is computed following the algorithm:*

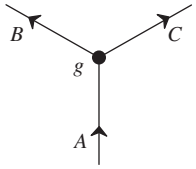
1. ***Notation:*** *Denote the incoming and outgoing 4-momenta by $\mathrm{p}_1, \mathrm{p}_2, \ldots$ and the "internal" 4-momenta (assigned to lines that connect two vertices within the graph) $\mathrm{q}_1, \mathrm{q}_2, \ldots$ Orient each line, selecting the positive sense of the corresponding 4-momentum.*
2. ***Vertices:*** *Assign to each vertex the factor $-ig$.*
3. ***Lines:*** *Assign to the jth internal line the factor $\frac{i}{\mathrm{q}_j^2 - m_j^2 c^2}$, the so-called **propagator**. As this depicts a virtual particle,[16] $\mathrm{q}_j^2 \neq m_j^2 c^2$.*
4. ***4-momentum conservation:*** *Assign to each vertex the factor $(2\pi)^4 \delta^4(\sum_j \mathrm{k}_j)$, where k_j $(-\mathrm{k}_j)$ are the 4-momenta entering (leaving) the vertex.*
5. ***4-momentum integration:*** *Assign to the jth internal line the $\int \frac{\mathrm{d}^4 \mathrm{q}_j}{(2\pi)^4}$-integral.*
6. ***Reading off the amplitude:*** *The above procedure produces*

$$\left(-i\,\mathfrak{M}\right)(2\pi)^4 \delta^4\Big(\sum_j \mathrm{p}_j\Big), \tag{3.129}$$

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

The $A \longrightarrow B + C$ decay

The lowest order contribution is of the order g^1:

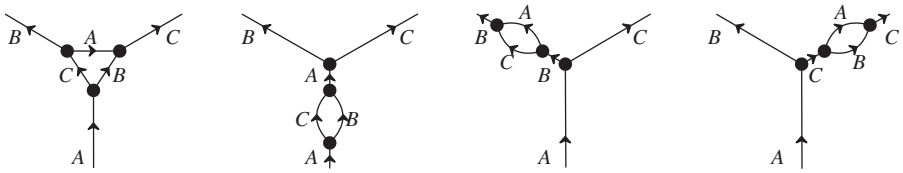


(3.130)

[16] For virtual particles, it is not that q^2 is required to not equal m^2c^2 (i.e., $\mathrm{q}^2 \neq m^2c^2$), but rather that q^2 is not required to equal m^2c^2. In distinction from "does not equal," the relation "not required to equal" will herein be denoted by the non-standard symbol "$\neq$" [☞ Tables C.7 on p. 529 and C.8 on p. 529].

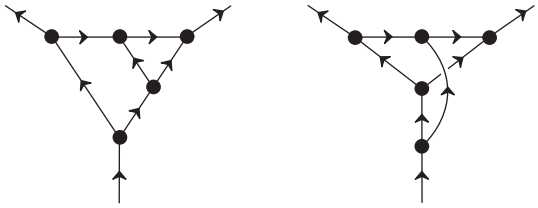
https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

The time axis is directed vertically, upward. The next contributions are of the order g^3:



(3.131)

and so on. The lowest order contribution (3.130) is depicted by a tree-graph (with no closed loop). The subsequent contributions (3.131) all have precisely one closed loop and are of the order g^3; there are no contributions of even order g^{2k}. However, starting with the next (g^5) order, a novelty appears, which can be seen by comparing the following two graphs:



(3.132)

The left-hand graph is planar, but the right-hand graph is not. This property of planarity may be used for a finer classification of graphs, and proves to be very useful in computations for the strong nuclear interaction [511]; for a recent review, see Ref. [349]. Also, only *connected* diagrams contribute: diagrams such as any one of the above but with a disconnected component (e.g., ) added do not contribute; this recalls the "excisions" in Section 3.3.2, the contribution of which had to be subtracted in non-relativistic stationary state perturbation theory.

Return to the contribution of the lowest order (3.130), where there are no internal lines. Procedure 3.1 reduces to:

1. Let the "external" 4-momenta be p_A, p_B and p_C.
2. Assign $-ig$ to the vertex.
3. There are no internal lines, and so no propagators either.
4. Assign $(2\pi)^4\delta^4(\mathrm{p}_A - \mathrm{p}_B - \mathrm{p}_C)$ to the vertex.
5. There are no internal lines, and so no integration either.
6. We've obtained

$$-ig\,(2\pi)^4\delta^4(\mathrm{p}_A - \mathrm{p}_B - \mathrm{p}_C) = (-i\,\mathfrak{M})\,(2\pi)^4\delta^4(\mathrm{p}_A - \mathrm{p}_B - \mathrm{p}_C), \tag{3.133}$$

from where $\mathfrak{M}$ is read off – here, to order g^1. Thus, $\mathfrak{M}^{(1)} = g$.

Inserting this result into the expression (3.122) we obtain

$$\Gamma^{(1)} = \frac{g^2|\vec{p}_B|_0}{8\pi\hbar m_A^2 c}, \qquad \text{so that} \qquad \tau^{(1)} = \frac{1}{\Gamma^{(1)}} = \frac{8\pi\hbar m_A^2 c}{g^2|\vec{p}_B|_0}, \tag{3.134}$$

where

$$|\vec{p}_B|_0 = \frac{c}{2m_A}\sqrt{m_A^4 + m_B^4 + m_C^4 - 2m_A^2m_B^2 - 2m_A^2m_C^2 - 2m_B^2m_C^2} = |\vec{p}_C|_0. \tag{3.135}$$

The result $\mathfrak{M}^{(1)} = g$, and so also (3.134), is analogous to the result (3.92a): $\mathfrak{M}^{(1)} = g$ is the first-order result in $\mathfrak{M}$ expanded in a power series over g, as is $E_n^{(1)}$ the first-order result in a power-series expansion of the energy over λ. In this sense, the constant of interaction g serves

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

as the perturbation parameter, and its numerical value must be sufficiently "small" so that such a power series would make sense,[17] so that the interaction parameter g has the same formal role as the perturbation parameter λ in non-relativistic quantum mechanics, in Section 3.3.2.

However, unlike this formal parameter, the interaction constant g has a physical meaning and its physical value can be measured. In this toy-model, it would suffice to measure the lifetime of the A-particle, then use the relation (3.134) to compute g – to the lowest perturbative approximation. An experiment would, of course, follow an *ensemble* of a large number of A-particles, and the diminishing of their number during time would determine the average value of the half-life $t^{(1)}_{1/2} = \tau^{(1)}\ln(2)$.

Note, however, that there would occur an additional "correction": Processes (3.130), (3.131), (3.132), and others of the same order in g (and then also the higher-order ones) are contributions only to the exclusive (partial) decay $A \to B + C$. If the mass m_A is sufficiently bigger than m_B, m_C, the A-particle may also decay into more-particle modes:

$$A \to 3B + C, \qquad A \to B + 3C, \qquad \dots \qquad A \to pB + qC, \tag{3.136a}$$

which are limited by the relation

$$\sum_{p,q} p\, m_B + q\, m_C < m_A, \tag{3.136b}$$

as well as the nature of the decay graphs, from which it follows, e.g.,

$$(p,q) \neq (1,2), (2,2), (2,3), \dots \tag{3.136c}$$

The A+A⟶ B+B scattering

The constant g may also be measured – in a thought experiment since this is a toy-model – more directly, by measuring the intensity of the interaction during scattering. Griffiths [243] analyzes the inelastic decay $A + A \to B + B$, where "incoming" A-particles are assigned the 4-momenta p_1 and p_2, and the outgoing B-particles p_3 and p_4.

By definition and using the relation (3.37), we have that $\frac{E_i}{c} = \sqrt{m_i^2 c^2 + \vec{p}_i^{\,2}}$ for $i = 1,2,3,4$. In the CM system, where

$$\left(-\frac{E_1}{c}, \vec{p}_1\right) + \left(-\frac{E_2}{c}, \vec{p}_2\right) = (p_0, \vec{0}) = \left(-\frac{E_3}{c}, \vec{p}_3\right) + \left(-\frac{E_4}{c}, \vec{p}_4\right), \tag{3.137}$$

the total linear momentum vanishes so $\vec{p}_1 = -\vec{p}_2$ and $\vec{p}_3 = -\vec{p}_4$. Denote $\theta := \measuredangle(\vec{p}_1, \vec{p}_3) = \measuredangle(\vec{p}_2, \vec{p}_4)$, so that $\measuredangle(\vec{p}_1, \vec{p}_4) = \measuredangle(\vec{p}_2, \vec{p}_3) = (\pi - \theta)$. We also have that

$$E_1 = c\sqrt{m_A^2 c^2 + \vec{p}_1^{\,2}} = c\sqrt{m_A^2 c^2 + (-\vec{p}_1)^2} = E_2, \tag{3.138}$$

$$E_3 = c\sqrt{m_B^2 c^2 + \vec{p}_3^{\,2}} = c\sqrt{m_B^2 c^2 + (-\vec{p}_3)^2} = E_4. \tag{3.139}$$

From conservation of energy, i.e., the energy component (3.137), it follows that

$$E := E_1 = E_2 = E_3 = E_4. \tag{3.140}$$

[17] It would be ideal if this power series would converge. Within field theory, in practice – if this can be determined at all – one mostly obtains asymptotic or even formally divergent sums, for which one must independently establish if the sum may be unambiguously assigned a particular value [☞ [259] for "summability"] for the given value of the constant g as its argument. Not infrequently, one only knows that the first several orders of perturbative computations are ever smaller "corrections," but with no formal proof about the nature of the whole infinite series. On the other hand, practical computations in quantum electrodynamics show ***unprecedented precision***: both perturbative computations and experimental measurements are found to agree with a relative error $< O(10^{-10})$ [293], better than anywhere else in natural sciences!

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Equating the squares of these energies, we obtain

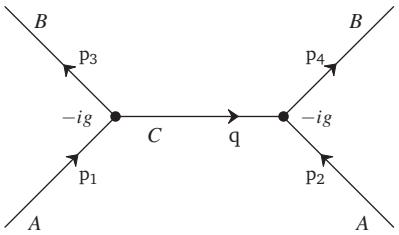
$$m_A^2c^2 - \vec{p}_1^{\,2} = m_A^2c^2 - \vec{p}_2^{\,2} = m_B^2c^2 - \vec{p}_3^{\,2} = m_B^2c^2 - \vec{p}_4^{\,2}. \tag{3.141}$$

In the limiting case of this toy-model, when $m_A = m_B = m$, but $m_C = 0$, the relation (3.141) gives

$$|\vec{p}| \;:=\; |\vec{p}_1| = |\vec{p}_2| = |\vec{p}_3| = |\vec{p}_4|. \tag{3.142}$$

Scattering In this case, following Procedure 3.1 we have:

1. Denote the "incoming" 4-momenta by $\mathrm{p}_1, \mathrm{p}_2$, and the "outgoing" ones by $\mathrm{p}_3, \mathrm{p}_4$:

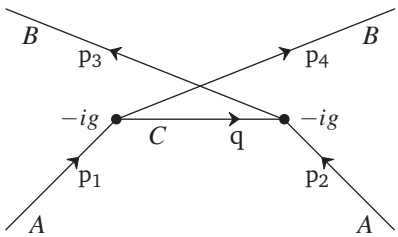


(3.143)

2. Assign to both vertices a factor of $-ig$.
3. Assign to the internal line the 4-momentum q, and the propagator $\frac{i}{\mathrm{q}^2 - m_C^2c^2}$.
4. Assign to the vertices the factors $(2\pi)^4\delta^4(\mathrm{p}_1 - \mathrm{q} - \mathrm{p}_3)$ and $(2\pi)^4\delta^4(\mathrm{p}_2 + \mathrm{q} - \mathrm{p}_4)$.
5. Integrate over $\frac{\mathrm{d}^4\mathrm{q}}{(2\pi)^4}$.
6. We have thus obtained:

$$\begin{aligned} -i\,\mathfrak{M}\,(2\pi)^4\delta^4(\mathrm{p}_1 + \mathrm{p}_2 - \mathrm{p}_3 - \mathrm{p}_4) \\ &= \int \frac{\mathrm{d}^4\mathrm{q}}{(2\pi)^4}\,(-ig)^2\frac{i}{\mathrm{q}^2 - m_C^2c^2}\,(2\pi)^4\delta^4(\mathrm{p}_1 - \mathrm{q} - \mathrm{p}_3)\,(2\pi)^4\delta^4(\mathrm{p}_2 + \mathrm{q} - \mathrm{p}_4) \\ &= -ig^2(2\pi)^4\int \frac{\mathrm{d}^4\mathrm{q}}{\mathrm{q}^2 - m_C^2c^2}\,\delta^4(\mathrm{p}_1 - \mathrm{q} - \mathrm{p}_3)\,\delta^4(\mathrm{p}_2 + \mathrm{q} - \mathrm{p}_4) \\ &= -i\,\frac{g^2}{(\mathrm{p}_4 - \mathrm{p}_2)^2 - m_C^2c^2}\,(2\pi)^4\delta^4(\mathrm{p}_1 + \mathrm{p}_2 - \mathrm{p}_3 - \mathrm{p}_4). \end{aligned} \tag{3.144}$$

However, this is not the only Feynman diagram that produces a g^2 contribution; holding the positions of the outgoing lines and their assigned 4-momenta, it is clear that a "topologically" distinct diagram is obtained by swapping the vertices to which the outgoing lines connect:



(3.145)

and which clearly produces a contribution of the same form (3.144), however, with the exchange $\mathrm{p}_3 \leftrightarrow \mathrm{p}_4$. As there are no other Feynman diagrams, the amplitude $\mathfrak{M}$ is read off from the sum of these two contributions:

$$\mathfrak{M} = \frac{g^2}{(\mathrm{p}_4 - \mathrm{p}_2)^2 - m_C^2c^2} + \frac{g^2}{(\mathrm{p}_3 - \mathrm{p}_2)^2 - m_C^2c^2}. \tag{3.146}$$

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

This amplitude is to be substituted into the expression (3.127):

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \Big(\frac{\hbar c}{8\pi}\Big)^2 \frac{S\,|\mathfrak{M}|^2}{(E_1+E_2)^2}\,\frac{|\vec{p}_3|}{|\vec{p}_1|}, \tag{3.147}$$

where $S = \frac{1}{2}$, we used the result (3.163), and chose to use a spherical coordinate system in which the angle θ equals the angle $\measuredangle(\vec{p}_1, \vec{p}_3)$.

The denominators in the ratios (3.146) are

$$\begin{aligned}(\mathrm{p}_4-\mathrm{p}_2)^2 - m_C^2c^2 &= (m_A^2+m_B^2-m_C^2)c^2 - 2\Big(\frac{E_4}{c}\frac{E_2}{c} - \vec{p}_4\cdot\vec{p}_2\Big)\\ &= (m_A^2+m_B^2-m_C^2)c^2 - 2\Big(\sqrt{(m_B^2c^2+\vec{p}_4{}^2)(m_A^2c^2+\vec{p}_2{}^2)} - \vec{p}_4\cdot\vec{p}_2\Big),\\ (\mathrm{p}_3-\mathrm{p}_2)^2 - m_C^2c^2 &= (m_A^2+m_B^2-m_C^2)c^2 - 2\Big(\sqrt{(m_B^2c^2+\vec{p}_3{}^2)(m_A^2c^2+\vec{p}_2{}^2)} - \vec{p}_3\cdot\vec{p}_2\Big),\end{aligned}$$

which significantly simplifies in the limiting case when $m_A = m_B = m$ and $m_C = 0$:

$$(\mathrm{p}_4-\mathrm{p}_2)^2 - m_C^2c^2 = 2m^2c^2 - 2\Big((m^2c^2+\vec{p}^2) - \vec{p}^2\cos\theta\Big) = -2\vec{p}^2(1-\cos\theta), \tag{3.148}$$

$$(\mathrm{p}_3-\mathrm{p}_2)^2 - m_C^2c^2 = 2m^2c^2 - 2\Big((m^2c^2+\vec{p}^2) - \vec{p}^2\cos(\pi-\theta)\Big) = -2\vec{p}^2(1+\cos\theta). \tag{3.149}$$

Thus, in the limiting case $m_A = m_B = m$, and $m_C = 0$:

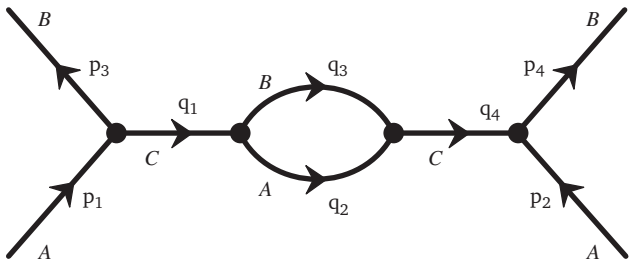
$$\mathfrak{M} = -\frac{g^2}{\vec{p}^2\sin^2\theta}, \qquad \text{so} \qquad \Big(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\Big) = \frac{1}{2}\Big(\frac{\hbar c}{16\pi}\frac{g^2}{E\,\vec{p}^2\sin^2\theta}\Big)^2, \tag{3.150}$$

which may be used to measure – in a thought experiment for this toy-model – the interaction constant g by measuring the differential effective cross-section as a function of the deflection angle $\theta := \measuredangle(\vec{p}_1, \vec{p}_2)$, and the energy and linear momentum of the "incoming" particles, $\vec{p} = \vec{p}_1$ and $E = E_1$, respectively.

Of course, the diagrams (3.143) and (3.145) and so also the result (3.150) are all only the contributions of the lowest order. In the next, $O(g^4)$, order, we have the diagrams listed in Figure 3.2.

The number of Feynman diagrams shown in Figure 3.2[18] indicates the volume of the task in computing physical quantities, such as the differential effective cross-section, order by order in perturbation theory. It is fairly obvious that the task of computing even just the first few order contributions (in the expansion organized into growing powers of the interaction constant) to a physical quantity is already a very demanding exercise, so that discussions and analyses of the convergence of the whole perturbative sum must limit to general properties.

Besides, diagrams such as the 15th in Figure 3.2 uncover a new property: divergences and renormalization. Consider this diagram, redrawn here as



(3.151)

[18] The counting given in Ref. [243, 1st edn.] was imprecise: one-third of the diagrams counted there are either impossible in the A-B-C toy-model (diagrams 4 and 7, in Figure 3.2), or are counted twice (diagrams 11, 12 and 14). However, this is seen only when the lines are assigned particles and one verifies all vertices to be of the form (3.128).

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

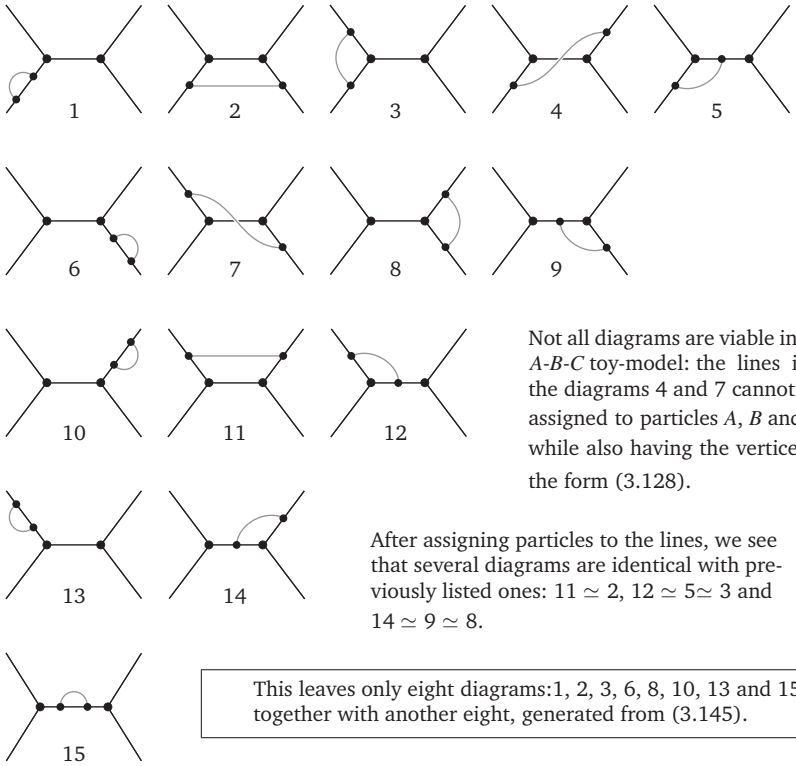


Figure 3.2 Fifteen possible $O(g^4)$ diagrams, before assigning particles to lines. All diagrams are generated from the $O(g^2)$ diagram (3.143), by adding one internal line. When assigning particles to lines, maintain the "external conditions": two incoming A-type particles (initial, lower in the diagrams) and two outgoing B-type particles (final, upper in the diagrams). These conditions reduce the total number to eight.

Following the Procedure 3.1, we obtain

$$(-ig)^4 \int \frac{i\,\mathrm{d}^4\mathrm{q}_1}{\mathrm{q}_1^2 - m_C^2c^2}\frac{i\,\mathrm{d}^4\mathrm{q}_2}{\mathrm{q}_2^2 - m_A^2c^2}\frac{i\,\mathrm{d}^4\mathrm{q}_3}{\mathrm{q}_3^2 - m_B^2c^2}\frac{i\,\mathrm{d}^4\mathrm{q}_4}{\mathrm{q}_4^2 - m_C^2c^2} \times \delta^4(\mathrm{p}_1 - \mathrm{p}_3 - \mathrm{q}_1)\,\delta^4(\mathrm{q}_1 - \mathrm{q}_2 - \mathrm{q}_3)\,\delta^4(\mathrm{q}_2 + \mathrm{q}_3 - \mathrm{q}_4)\,\delta^4(\mathrm{p}_2 + \mathrm{q}_4 - \mathrm{p}_4), \tag{3.152}$$

where the first δ-function cancels the $\mathrm{d}^4\mathrm{q}_1$-integral and replaces $\mathrm{q}_1 \to (\mathrm{p}_1 - \mathrm{p}_3)$, while the last δ-function cancels the $\mathrm{d}^4\mathrm{q}_4$-integral and replaces $\mathrm{q}_4 \to (\mathrm{p}_4 - \mathrm{p}_2)$. Thereafter, the second δ-function cancels the $\mathrm{d}^4\mathrm{q}_2$-integral and replaces $\mathrm{q}_2 \to (\mathrm{q}_1 - \mathrm{q}_3) = (\mathrm{p}_1 - \mathrm{p}_3 - \mathrm{q}_3)$ and turns the remaining, third δ-function into the expected factor $\delta^4(\mathrm{p}_1 + \mathrm{p}_2 - \mathrm{p}_3 - \mathrm{p}_4)$. This then leaves

$$\mathfrak{M} = \frac{(g/2\pi)^4}{\left[(\mathrm{p}_1 - \mathrm{p}_3)^2 - m_C^2c^2\right]^2}\int \frac{\mathrm{d}^4\mathrm{q}_3}{[(\mathrm{p}_1 - \mathrm{p}_3 - \mathrm{q}_3)^2 - m_A^2c^2](\mathrm{q}_3^2 - m_Bc^2)}. \tag{3.153a}$$

This $\mathrm{d}^4\mathrm{q}_3$-integral necessarily diverges: in 4-dimensional spherical coordinates, we have that $\mathrm{d}^4\mathrm{q}_3 = \rho^3\mathrm{d}\rho\,\mathrm{d}\Omega_{(3)}$, so the "radial" integral becomes, near the upper (infinite) limit:

$$\sim \int^{\infty}\frac{\rho^3\mathrm{d}\rho}{\rho^4} = \lim_{R\to\infty}\int^R\frac{\mathrm{d}\rho}{\rho} = \lim_{R\to\infty}\ln(R), \tag{3.153b}$$

which diverges logarithmically.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Comment 3.5 *Note two of the properties of this divergent result:*

1. *The divergence may occur only in the integration over a 4-momentum associated with a closed loop in the Feynman diagram, since only such a 4-momentum is not determined from external data by the 4-momentum conservation law.*
2. *This divergence does not emerge in attempting to sum an infinite series as is the case in a so-called asymptotic series,*[19] *but occurs in a single, concrete contribution to a summand in such a series.*

Digression 3.11 Oppenheimer and Waller seem to have been the first to notice, independently and in 1930, the appearance of divergences in perturbative calculations in field theory; this discovery was such a shock that Pauli at first did not want to believe in its correctness [552].

The appearance of divergences in contributions such as (3.153) reminds us a little of the situation in non-relativistic stationary state perturbation theory in systems with degeneration. There, the formula (1.18), which provides the first-order correction to the state $|n\rangle$, becomes meaningless if there exists

$$|m\rangle \neq |n\rangle \ : \qquad E_m^{(0)} = E_n^{(0)}, \quad \langle m|H'|n\rangle \neq 0. \tag{3.154}$$

Such contributions in the sum (1.18) are of the form $\frac{1}{0}$ and literally make no sense. In that simpler case, the problem is solved by changing the basis of states so that the problematic combinations (3.154) and terms of the type $\frac{1}{0}$ in the sum (1.18) no longer occur.

Digression 3.12 To eliminate the offending situation (3.154), one defines

$$|m'\rangle = c_{mm}|m\rangle + c_{mn}|n\rangle, \quad |n'\rangle = c_{nm}|m\rangle + c_{nn}|n\rangle, \tag{3.155a}$$

and requires that $\langle m'|H'|n'\rangle = 0$. This implies that H_0 and H' have been simultaneously diagonalized over the $\{|m\rangle, |n\rangle\} \in \mathscr{H}$ subspace of the Hilbert space. In turn, this implies that

$$[H_0, H'] = 0 \quad \text{over} \quad \{|m\rangle, |n\rangle\} \in \mathscr{H}. \tag{3.155b}$$

In this sense the "$\frac{1}{0}$-divergence" is "removed." It is also clear from the structure of the sums in the results (1.18), (1.19), and also the entire algorithm given in Section 3.3.2, that the re-diagonalization of the basis $|n\rangle$ that removes the $\frac{1}{0}$-divergences from the sum (1.18) to the first perturbative order also removes all divergences of this type in the whole perturbative procedure.

However, divergences of the form (3.153) are harder to "remove," and their treatment has halted the first physically significant field theory – quantum electrodynamics – for almost two decades. The work of many physicists on this problem culminated in independent and equivalent methods by Richard Feynman, Julian Schwinger and Shin-Ichiro Tomonaga, which were systematized by Freeman Dyson.

[19] H. Poincaré defined the series $\sum_k c_k x^k$ where $\lim_{x\to\infty} x^k R_k(x) = 0$ for any fixed k, but $\lim_{k\to\infty} x^k R_k(x) = \infty$ to be *asymptotic* (semi-convergent). Here $R_k(x) := [f(x) - \sum_{i=0}^{k} c_i x^{-i}]$, and $f(x)$ asymptotically agrees with the sum for large $x \sim \infty$.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Digression 3.13 The systematic procedure and *idea* of renormalization stems from an older and unrelated idea: In 1902, Max Abraham proposed the model [4] in which the electron was a sphere of finite radius on the surface of which the electron charge is uniformly distributed. In 1904, Hendrik Lorentz [346] developed this idea, so that the model is now called the Abraham–Lorentz model. In this model, the work required to assemble the electron charge by bringing it from infinity into the Gaussian sphere of radius r_e contributes to the mass of the electron. If this is the only contribution, that work may be equated with the (electromagnetic) rest energy of the electron:

$$m_{\mathrm{em}}c^2 = \tfrac{1}{2}\int \mathrm{d}^3\vec{r}\,\vec{E}^2 = 2\pi\int_{r_e}^{\infty} r^2\mathrm{d}r\,\left(\frac{e}{4\pi\epsilon_0\, r^2}\right)^2 = \frac{1}{2}\frac{e^2}{4\pi\epsilon_0}\frac{1}{r_e}, \tag{3.156a}$$

if the electron charge is distributed uniformly on the surface of the sphere of radius r_e. If the charge is distributed uniformly throughout the entire sphere, the factor $\frac{1}{2}$ is replaced by $\frac{3}{5}$. Equating m_{em} with the measured electron mass (as if the electron mass stems entirely from the electric field that this electron produces, and the charged shell has no mass of its own) and neglecting the numerical factor $\frac{1}{2} - \frac{3}{5}$, we obtain the classical electron radius:

$$r_e = \frac{e^2}{4\pi\epsilon_0}\frac{1}{m_e\, c^2} = \alpha_e\frac{\hbar}{m_e c} = 2.817\,940\,325\times 10^{-15}\,\mathrm{m}. \tag{3.156b}$$

Because of the dependence $m_{\mathrm{em}} \propto r_e^{-1}$, it follows that the electron cannot be ideally point-like: if it were, its mass would be infinitely large. Since the total effective mass must also include (realistically, a non-vanishing) mass of the spherical shell, the "true" electron radius may differ from the result (3.156b). However, for the radius to be *smaller* than r_e, it would be necessary for the mass of the spherical shell to be *negative*. In the limiting case of the point-like electron, this infinitely small shell would have to have an infinitely negative mass, which is an evidently meaningless value.

However, from these considerations about the Abraham–Lorentz model of the electron stems the idea that the measured values of a physical quantity may consist of several contributions, which – in the limiting case – may each diverge, as long as their sum (which is what is compared with the experimental data!) is a finite quantity.

The basic idea, schematically, is that for each parameter there exists

$$m_{\mathrm{physical}} = m_{\mathrm{bare}} + \delta m, \qquad g_{\mathrm{physical}} = g_{\mathrm{bare}} + \delta g, \qquad \text{etc.} \tag{3.157}$$

where the "bare" version of the parameter is the one showing up in the classical Lagrangian theory, and the "quantum correction" δm often diverges. However, m_{bare} is defined so as to also diverge,[20] and precisely so that the physical value of the parameter remains finite and comparable with the experiments. Besides, the systematic procedure of renormalization guarantees that the so-defined finite part and divergent part of the result may be consistently separated order by order in perturbation theory for quantum electrodynamics, which then serves as a template for all other existing field theory models.

[20] One of the methods of "canceling" divergences requires that the integrals are computed in finite limits, $\pm\Lambda$, so that one can isolate the portion of the contributions that are independent of Λ in the sums of the form (3.157). In the $\Lambda \to \infty$ limit, that Λ-independent portion represents the desired physical quantity. There exist several other methods for isolating the "finite part" from divergent integrals, but there is no general formal proof that the finite result does not depend on the method of its isolation.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

Most nontrivial models in field theory are in practice *defined* by means of the perturbative computations, including some variant of the renormalization procedure. Certain results in a growing class of models can be computed by non-perturbative means – or by using an essentially different perturbative method where individual terms of "lower order" may represent a sum of a large number of (contributions from) standard Feynman diagrams. A formal and rigorous proof of finiteness of all possible observables is not known in general, and in this sense field theory is in general not formally rigorously defined, nor does one know if field theory in general – or even a certain concrete model, such as quantum electrodynamics – is formally self-consistent![21] Nevertheless, perturbative and other concrete results offer enough useful data to compare with experiments, which suffices for a pragmatic acceptance of this theoretical system – all the more so, since (**1**) the various renormalization prescriptions invariably produce final results agreeing for observables, and (**2**) no contradiction has been detected in anomaly-free theoretical models [☞ Section 7.2.3].

The appearance and the conspicuous cancellation of divergent contributions in perturbative computations of evidently measurable (and finite!) physical quantities is still cited as the cause for a principled disagreement with the entire renormalization procedure [☞ e.g., Ref. [29]]. However, the number of living physicists who openly oppose this procedure is decreasing.[22] The Reader with a piqued interest in the subject should turn to the texts on quantum field theory [63, 48, 441, 459, 154, 474, 249, 240, 425, 554, 555, 484, 588, 496, 446, 589, 316, 7, 586, 277, 590], texts on renormalization itself [113, 212], and research articles, such as [44, 431, 343, 146, 253].

3.3.5 Exercises for Section 3.3

✎ ***3.3.1*** *In the special case when $\vec{v} = v\,\hat{e}_z$, show that the transformations (3.1) acquire the well-known form:*

$$x' = x, \quad y' = y, \quad z' = \gamma(z - v\,t), \qquad x = x', \quad y = y' \quad z = \gamma(z' + vt'), \tag{3.158a}$$

$$t' = \gamma\Big(t - \frac{v\,z}{c^2}\Big), \qquad\qquad t = \gamma\Big(t' + \frac{v\,z'}{c^2}\Big), \tag{3.158b}$$

with the usual $\gamma = (1 - \vec{v}^2/c^2)^{-1/2}$.

✎ ***3.3.2*** *Using that $[\Gamma] = T^{-1}$, $[\sigma] = L^2$, $[v] = LT^{-1}$ and $\big[|\Psi|^2\big] = L^{-3}$, prove equation (3.110) and Conclusion 3.2.*

✎ ***3.3.3*** *For the elastic collision $A + B \to A' + B'$, in a system where B is originally at rest (and is the target), derive*

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \approx S\Big(\frac{\hbar}{8\pi}\Big)^2 \frac{|\mathfrak{M}|^2}{m_B} \frac{\vec{p}_{A'}^{\,2}}{|\vec{p}_A|\,\big|(|\vec{p}_{A'}|(E_A + m_B c^2) - |\vec{p}_A|E_{A'}\cos\theta)\big|}. \tag{3.159}$$

Here A and A' denote the incoming and outgoing, but otherwise identical particles, just as do B and B'.

✎ ***3.3.4*** *Show that the result of the previous problem simplifies when $(m_A/m_B) \ll 1$:*

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \approx S\Big(\frac{\hbar\,E_{A'}}{8\pi\,E_A}\Big)^2 \frac{|\mathfrak{M}|^2}{m_B^2}. \tag{3.160}$$

[21] In view of Gödel's incompleteness theorem, a formally rigorous proof of self-consistency of field theory may turn out to be a pipe dream, since theoretical axiomatic systems that are sufficiently strong (e.g., include standard arithmetics) turn out to also be incapable of proving their own consistency [☞ Appendix B.3].

[22] To paraphrase Max Planck [428, pp. 33–34], new scientific truths do not convince their opponents, they outlive them.

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press

✎ ***3.3.5*** *For the elastic collision in Exercise 3.3.3 but in the case when the recoil of the target after the collision may be neglected since $m_B c^2 \gg E_A$, derive*

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \approx \left(\frac{\hbar}{8\pi m_B c}\right)^2 |\mathfrak{M}|^2. \tag{3.161}$$

✎ ***3.3.6*** *For the inelastic collision $A + B \rightarrow C_1 + C_2$, in a system where B (the target) is originally at rest, and $(m_{C_i}/m_A) \ll 1$ and $(m_{C_i}/m_B) \ll 1$, derive*

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \approx \left(\frac{\hbar}{8\pi}\right)^2 \frac{S\,|\mathfrak{M}|^2}{m_B(E_A + m_B c^2 - |\vec{p}_A| c\,\cos\theta)} \frac{|\vec{p}_{C_1}|}{|\vec{p}_A|}, \tag{3.162}$$

where θ is the angle between $\vec{p}_1$ and $\vec{p}_3$.

✎ ***3.3.7*** *Why is the cascade decay $A \rightarrow B + C \rightarrow 2B + A$ in the toy-model of Section 3.3.4 forbidden, but $A \rightarrow B + C \rightarrow 2B + A \rightarrow 3B + C$ may be allowed? What is the condition for the latter process to be viable?*

✎ ***3.3.8*** *Using only Feynman diagrams to analyze the possible decay modes of particle A, show that p and q in relation (3.136) must both be odd integers.*

✎ ***3.3.9*** *For a head-on collision of two particles of masses m_1 and m_2, we have $\mathrm{p}_1 = (E_1/c, \vec{p})$ and $\mathrm{p}_2 = (E_2/c, -\vec{p})$ in the CM system. Show that*

$$c\sqrt{(\mathrm{p}_1 \cdot \mathrm{p}_2)^2 - (m_1 m_2 c^2)^2} = (E_1 + E_2)|\vec{p}|. \tag{3.163}$$

✎ ***3.3.10*** *Prove that, in the A-B-C toy-model and with $g < 1$, the elastic collisions are $O(g^4)$ times less probable than the inelastic collisions such as (3.143).*

✎ ***3.3.11*** *Prove that equation (3.155b) is satisfied for all of the finitely many degenerate states (3.154), so that the standard procedure described in Digression 3.12 on p. 122 is always possible.*

https://doi.org/10.1017/9781009291507.006 Published online by Cambridge University Press