2

Relativistic kinematics, electromagnetic fields and the method of virtual quanta

The dynamics of the massless relativistic string (which we will meet at very many different places in this book) is a delightful theoretical laboratory to study the properties of the theory of special relativity. To make the book self-contained and also to define our notation we will briefly review in this chapter some properties of special relativity, in particular with respect to its implications for high-energy particle kinematics.

We will also review some properties of electromagnetic fields with particular emphasis on the features we are going to make use of later in the book. We will end with a description of the *interaction ability of an electrically charged particle.*

This is the first but not the last example in this book of the *law of the conservation of useful dynamics.* This says that every new generation of theoretical physicists tends to reinvent, reuse (and usually also rename) the most useful results of earlier generations. One reason is evidently that there are few situations where it is possible to find a closed mathematical expression for the solution to a dynamical problem.

Here our basic aim is to describe the interactions between charged particles which are moving with very large velocities (as they do in high-energy physics). As a charged particle interacts via its field the question can be reformulated into finding a way to describe the field of a charged particle which is moving very fast. To account for quantum mechanics we need a way to describe the quantum properties of the charged particle's field and this problem can be solved even at a semi-classical level. It is possible to obtain a closed formula for the *flux of the field quanta* in this case.

Fermi addressed the problem in the 1920s, Weizsäcker and Williams found the method independently of Fermi and each other in the 1930s. After that it became a standard tool in connection with QED in terms of the *method of virtual quanta*, the MVQ. Later again Feynman made use

https://doi.org/10.1017/9781009401296.002 Published online by Cambridge University Press

of it in order to introduce the *parton model*. We will discuss that model repeatedly in this book, but it is useful to see how 'partons' emerge even at the semi-classical level in electromagnetism.

2.1 The Lorentz boost

Michelson and Morley demonstrated that the velocity of light, c, is independent of the direction of a light beam. Einstein interpreted this finding to imply that the velocity c is independent of the relative motion of the light source and the detector.

We are not going to dwell upon the many basic questions that are raised by this interpretation but simply accept that it has profound implications with respect to measurements of events in space and time. The resulting predictions have been tested repeatedly and always been found to be true. In this section we will briefly consider some of these predictions.

I *The Lorentz boost*. Consider two observers A and B, moving with respect to each other. We will suppose that they have calibrated their watches and decided upon a common origin in space and time as well as the directions of the coordinate axes in space. The arrangement will be that they move along their common x-axis so that B has the velocity v with respect to A. We will for simplicity use units such that the velocity of light $c = 1$. Then an event (1) which for A occurs at the space-time coordinates

$$(1) \equiv (t_{1A}, x_{1A}, y_{1A}, z_{1A}) \tag{2.1}$$

will for B, in his system, seem to occur at the time and space coordinates (with the corresponding index B):

$$\begin{aligned} t_{1B} &= \gamma(v)(t_{1A} - v x_{1A}) \\ x_{1B} &= \gamma(v)(x_{1A} - v t_{1A}) \\ y_{1B} &= y_{1A} \\ z_{1B} &= z_{1A} \end{aligned} \tag{2.2}$$

This transformation is termed a *boost along the x-axis* and $\gamma(v) = 1/\sqrt{1-v^2}$. The time- and the (*longitudinal*) x-coordinates get mixed by the transformation but the *transverse* coordinates, i.e. the y- and z-coordinates, are unaffected. Several boosts may be performed one after the other. It is easy to see that the final result does not depend upon the order and therefore the boosts along a single direction constitute a commutative (abelian) group.

More complex transformations also include rotations of the coordinate systems. Note that such rotations in general do not commute with each

https://doi.org/10.1017/9781009401296.002 Published online by Cambridge University Press

other or with the boost transformations. This means that the outcome of the total transformation depends upon the order in which each one of the rotations and boosts is done.

II *The proper time.* The coordinate and time values are all differences between the commonly agreed origin and the space-time point at which event (1) occurs. They are all *relative coordinates*. A and B will have different values for their measured t, x values for the event but there is one combination which they will agree upon,

$$t_{1A}^2 - x_{1A}^2 = t_{1B}^2 - x_{1B}^2 \equiv \tau_1^2 \tag{2.3}$$

The proper time of the event, τ_1, is evidently an *invariant* with respect to all boosts along the x-axis. This means that it does not contain any reference to the relative velocity of the observers along the x-axis.

The proper time is the value a watch would show if it started out from the origin (i.e. at $t = 0, x = 0$) in A's system and moved away with velocity $v_A = x_{1A}/t_{1A}$. Then it will arrive at x_{1A} at time t_{1A}, just when the event (1) occurs. To see this imagine that observer B had chosen the velocity $v = v_A$. It is therefore the time obtained in the *rest frame* of the watch. This is the frame in which both events occur at the same place, the space origin (make use of the second line in Eq. (2.2)!).

IIIA *Time dilation.* The observer A will conclude that the time difference in his system that corresponds to the proper time τ_1 would be (make use of the first line of Eq. (2.2)!)

$$t_{1A} = \frac{\tau_1}{\sqrt{1 - v_A^2}} \tag{2.4}$$

This means that to A it will seem that the time difference is larger, i.e. it will seem as if time is passing more slowly in the watch rest system. This effect is called time dilation.

This is a noticeable effect for the fast-moving fragments of a collision between cosmic ray elements and the atoms of the upper atmosphere. There are e.g. the μ-particles, very short-lived when we produce them basically at rest, in the laboratory on earth. The lifetime of a μ-particle is around 2×10^{-6} seconds. Therefore even if it was moving with the velocity of light it would only be able to cover about 600 metres!

Nevertheless the produced μ-particles survive a sufficiently long time to be able to go all the way from the top of the atmosphere down to earth, where we can find them in abundance.

To understand this effect we note that the decay time is related to the properties of the particle in its rest frame while the 'survival time' we observe is the time it will take a fast-moving particle (with velocity close to c) to move the distance δ from the top of the atmosphere (at a height of around 2×10^4 meters) to the observation point on earth. According to Eq. (2.4) this survival time is much longer and therefore many of the μ-particles survive to reach the ground.

IIIB *Lorentz contraction.* There is a corresponding effect for distances, which is called Lorentz contraction. For the surviving μ-particles, the distance δ, which to us is about 2×10^4 meters, will seem to be at most the 600 metres mentioned above. Considered from the rest system of the μ-particle the distance δ_{rest} is the length that the earth and its atmosphere moves towards it during its lifetime! From the Eq. (2.4) we conclude for the Lorentz contraction effect

$$\delta_{rest} = \delta\sqrt{1 - v^2} \tag{2.5}$$

IV *Covariance.* The scalar product of two ordinary vectors $\mathbf{a} \cdot \mathbf{b}$, written in terms of the coordinates as $a_x b_x + a_y b_y + a_z b_z$, is an invariant with respect to rotations. It is possible to write the invariant τ_1^2 as a (generalised) scalar product. The quantity

$$(1)(2) = t_1 t_2 - x_1 x_2 - y_1 y_2 - z_1 z_2 \tag{2.6}$$

will be invariant with respect to the general Lorentz transformations (i.e. boosts and rotations in any order) if the coordinates and times of the events (1) and (2) transform with respect to Lorentz boosts as in Eq. (2.2) (and $(\mathbf{1}) \equiv (x_1, y_1, z_1)$ and similarly $(\mathbf{2})$ transform as ordinary vectors under rotation).

Such quantities as (1) in Eq. (2.1) are called *four-vectors*. They transform as vectors with respect to the Lorentz transformations, in particular as in Eq. (2.2) for boosts along an axis. Besides the invariants, in the same way called *scalars* under the Lorentz transformations, and the four-vectors it is possible to define *four-tensors* (the electromagnetic field tensor is an example of such a quantity).

All these quantities are said to be *covariant*: they transform in a linear way with respect to the Lorentz transformations, i.e. the corresponding quantities in different Lorentz frames are related by means of linear equations.

V *The transformation of the velocity.* As an example of a quantity with more complex properties with respect to the Lorentz transformations we consider the velocity. We have already mentioned the velocity v_A measured in A's system. From B's point of view the corresponding

https://doi.org/10.1017/9781009401296.002 Published online by Cambridge University Press

velocity will be (use both the first and the second line of Eq. (2.2)!)

$$v_B = \frac{v_A - v}{1 - v_A v} \tag{2.7}$$

It is not difficult to show that if the velocities v_A, v do not exceed $c = 1$ then the velocity v_B will have the same property.

VI The *energy-momentum four-vector*. The classical (Newtonian) definition of momentum is the mass (m) times the velocity (v_p) of the particle. But from Eq. (2.7) it is obvious that the transformation properties of the velocity are complex under a Lorentz boost. In order to generalise the definition of momentum Einstein made use of the proper time of the particle motion in the following way.

The velocity of the particle is defined in terms of its trajectory $\mathbf{r}(t)$ (i.e. its space position $\mathbf{r}$ labelled by means of the time t) as

$$\mathbf{v}_p = \frac{d\mathbf{r}}{dt} \tag{2.8}$$

For every (massive) particle it is possible to imagine a rest frame in which the particle is always at the (space) origin. In this way it is possible to define the proper time τ for the particle's motion; it is the time in this, the particle's rest system.

Considered from any other Lorentz frame the proper time τ will be related to the 'ordinary' time t by means of the differential equation

$$d\tau = dt\sqrt{1 - \mathbf{v}_p^2} \tag{2.9}$$

according to Eqs. (2.3), (2.4).

The proper time $\tau(t)$ defined in this way is unique as soon as proper boundary conditions are given for the differential equation. (Its functional dependence upon the time t will in general be different in different Lorentz frames, however.)

We conclude that the corresponding *four-velocity* u defined by

$$u \equiv \left(\frac{dt}{d\tau}, \frac{d\mathbf{r}}{d\tau}\right) = \gamma(\mathbf{v}_p)(1, \mathbf{v}_p) \tag{2.10}$$

will transform covariantly as a vector under the Lorentz transformations. (The third line of Eq. (2.10) is obtained from the differential equation (2.9).) Note that the corresponding invariant $uu = u^2$ has the value $u^2 = 1$. Einstein defined the *four-momentum* p of a particle as

$$p = (e, \mathbf{p}) = mu = m\gamma(\mathbf{v}_p)(1, \mathbf{v}_p) \tag{2.11}$$

The space components $\mathbf{p}$ (from now on the *momentum*) of this four-momentum (which we sometimes will call the *energy-momentum vector*) have the property that for small velocities $|\mathbf{v}_p| \equiv |v_p|$ (which should be interpreted to mean $|v_p| \ll c$, of course) they coincide with the classical momentum components.

The 'extra' component $e = m\gamma(\mathbf{v}_p)$ can be identified with the energy of the particle because for small velocities we obtain by expanding the square root

$$m\gamma(\mathbf{v}_p) \simeq m + m\mathbf{v}_p^2/2 \tag{2.12}$$

The second term corresponds to the well-known expression for the kinetic energy of a (nonrelativistic) particle. The first term, the rest energy, corresponds to the famous Einstein conclusion that the mass content of a particle is related to a stored energy, e_s

$$e_s = mc^2 \tag{2.13}$$

The ordinary vector velocity $\mathbf{v}_p$ can according to Eq. (2.11) be expressed as

$$\mathbf{v}_p = \frac{\mathbf{p}}{e} \tag{2.14}$$

2.2 Particle kinematics

The invariance equation for the energy-momentum vector $p = (e, \mathbf{p})$, if we consider a particle moving along a fixed direction $\mathbf{p} = p\mathbf{n}$, described by the unit vector n is

$$e^2 - p^2 = m^2 \tag{2.15}$$

This means that the energy (which always is positive for a particle) can be expressed as $e = \sqrt{p^2 + m^2}$.

VII *The rapidity variable.* According to Eq. (2.15) a particle with a fixed mass has a four-momentum which lies on a hyperbola in the ep-plane. It is possible to introduce a hyperbolic angle y_p to describe any particular point on the hyperbola:

$$\begin{aligned} e &= m\cosh y_p \\ p &= m\sinh y_p \end{aligned} \tag{2.16}$$

This hyperbolic angle is called the rapidity, and we note from the relationship between (e, p) and the ordinary velocity v_p in Eq. (2.11) that

$$v_p = \tanh y_p \simeq y_p \tag{2.17}$$

with the last line valid for small values of v_p and y_p. We also note that $\gamma(v_p) = \cosh y_p$.

https://doi.org/10.1017/9781009401296.002 Published online by Cambridge University Press

For a Lorentz boost along the direction **n** we obtain, using the first two lines of Eq. (2.2), with a boost velocity $v = \tanh y$ and using the notation (e_B, p_B) for the energy-momentum components in the new frame,

$$\begin{aligned} e_B &= \gamma(v)(e - vp) \\ &= m(\cosh y_p \cosh y - \sinh y_p \sinh y) = m\cosh(y_p - y) \\ p_B &= \gamma(v)(p - vE) \\ &= m(\sinh y_p \cosh y - \cosh y_p \sinh y) = m\sinh(y_p - y) \end{aligned} \tag{2.18}$$

This means that Lorentz boosts along **n** will move us along the hyperbola of Eq. (2.15). In particular any value of the energy-momentum can be obtained by a suitable boost from the rest system $y_p = 0$. In other words the *rapidity variable is additive.*

This also comes out of the relation for adding ordinary velocities, Eq. (2.7), if we express the velocities in terms of rapidities:

$$v_B \equiv \tanh y_B = \frac{v_A - v}{1 - v_A v} = \tanh(y_A - y) \tag{2.19}$$

If the rapidity is expressed in terms of the corresponding velocity v we obtain

$$y = \frac{1}{2}\ln\left(\frac{1+v}{1-v}\right) = \frac{1}{2}\ln\left(\frac{e+p}{e-p}\right) \tag{2.20}$$

It often occurs that in a given dynamical situation there may be a direction which is of particular importance. It is then useful to describe the particles under investigation in terms of their rapidities defined with respect to that direction (even if some or all of the particles move in somewhat different directions). This corresponds to using the velocity component, v_ℓ, along that (longitudinal) direction; we then obtain

$$y_\ell \equiv \frac{1}{2}\ln\left(\frac{1+v_\ell}{1-v_\ell}\right) = \frac{1}{2}\ln\left(\frac{e+p_\ell}{e-p_\ell}\right) \tag{2.21}$$

with p_ℓ the corresponding momentum component.

VIII *The lightcone components.* It is often useful to describe the energy-momentum vector with respect to the direction **n** in terms of the components

$$p_+ = e + p = m\exp y_p, \quad p_- = e - p = m\exp(-y_p) \tag{2.22}$$

For a boost with rapidity y along n these quantities transform as

$$p_+ \to p_+\exp(-y), \quad p_- \to p_-\exp y \tag{2.23}$$

It is of course natural that their product is a constant, equal to the invariant in Eq. (2.15). For the case in Eq. (2.21) one defines the lightcone components $(e \pm p_\ell)$. They can then be described with respect to the rapidity y_ℓ in the same way as in Eq. (2.22) except that the mass m is exchanged for the *transverse mass* m_t. This quantity is defined by

$$m_t = \sqrt{m^2 + \mathbf{p}_t^2} \tag{2.24}$$

in terms of the *transverse momentum vector* $\mathbf{p}_t$, corresponding to the two components of the momentum that are transverse to the chosen longitudinal direction.

We will at this point briefly consider Heisenberg's indeterminacy relations and indicate that although the position and the conjugate momentum of a particle cannot be determined simultaneously *it is possible to determine the rapidity and the position for a high-energy particle simultaneously with any degree of exactness*, [66].

The indeterminacy relations mean that owing to the commutation relation

$$[p, x] = -i \tag{2.25}$$

it is necessary that the width of a wave-packet in position x, Δx, is related to the corresponding width in momentum p, Δp by

$$\Delta x \Delta p \geq 1/2 \tag{2.26}$$

Merzbacher shows, by defining the mean and the width in the state with the wave function ψ as

$$\begin{aligned} \langle x \rangle &= \int dx \psi^*(x) x \psi(x) \\ (\Delta x)^2 &= \left\langle (x - \langle x \rangle)^2 \right\rangle = \int dx \psi^*(x)(x - \langle x \rangle)^2 \psi(x) \end{aligned} \tag{2.27}$$

with a similar relationship for p that there is a single kind of state, the Gaussian wave packet, for which Eq. (2.26) is an equality.

We can rewrite Eq. (2.26) in the following way for a particle with energy-momentum (e, p) with rapidity according to Eq. (2.16):

$$\Delta x \frac{\Delta p}{e} \equiv \Delta x \Delta y \geq \frac{1}{2e} \to 0 \tag{2.28}$$

when e is very large. Note that Eq. (2.16) implies that $dp/e = dy$.

Relation (2.28) is shown for a free particle, in [66], by actual construction of the necessary wave-packets. It implies that, although you can never fool Heisenberg, you are allowed to choose your variables in such a way that quantum mechanical effects can be small or negligible.

As you will find in connection with the Lund model, when we are concerned with the longitudinal dynamics we shall use the freedom to

https://doi.org/10.1017/9781009401296.002 Published online by Cambridge University Press

present semi-classical pictures, in which we go between coordinate- and rapidity-space descriptions. This cannot be done in the same cavalier way in connection with the transverse dynamics, because transverse momenta are in general very limited in size in high-energy physics.

2.3 Timelike, lightlike and spacelike vectors in Minkowski space

Up to now we have neglected the fact that the invariant size of a four-vector, like the squared proper time in Eq. (2.3), is not positive definite as is the corresponding length of an ordinary vector. This means that it is possible to find space-time points for which the proper time squared is vanishing or negative.

In both these cases the interpretation of proper time discussed above is no longer valid. There is no (proper) Lorentz frame that is a rest frame for an observer, in which both the start (at the origin) and the event itself occur at the same point in space.

Those points for which the proper-time interpretation is valid are called *timelike* and we note that they fulfil

$$|t_{1A}| > |\mathbf{r}_{1A}| \equiv \sqrt{\mathbf{r}_{1A}^2} \tag{2.29}$$

This is evidently a Lorentz-covariant definition.

All energy-momentum vectors for massive particles are also in the same way called timelike.

1 *Lightlike four-vectors*

In the case when the proper time squared vanishes it is possible to send a light signal directly from the origin to the event point and we therefore refer to this situation as a *lightlike* space-time vector difference.

There are other cases for which we will meet such lightlike vectors, e.g. when we want to describe massless particles such as the quanta of the electromagnetic field, *photons*. For them the energy (cf. Eq. (2.15)) is equal to the total momentum, i.e. $e = |\mathbf{k}| = |k|$. The corresponding rapidity y_ℓ as defined in Eq. (2.21) is directly expressible in terms of the angle, θ, between a given axis and the photon direction:

$$\begin{aligned} k_\ell &= |k| \cos\theta \\ y_\ell &= \frac{1}{2} \ln\left(\frac{1+\cos\theta}{1-\cos\theta}\right) = \ln\cot\left(\frac{\theta}{2}\right) \simeq -\ln\left(\frac{\theta}{2}\right) \end{aligned} \tag{2.30}$$

The last statement is an approximation valid for small angles.

Although Eq. (2.30) is strictly valid only for massless particles it is often a very good approximation (and then the variable is called the *pseudo-rapidity*) for other particles, those whose mass is small compared to their energy. In this way we obtain another intuitive way to look at the rapidity; *it is directly related to the angle with respect to the chosen longitudinal direction.*

While both the individual masses of two lightlike particles vanish, the sum of their energy-momenta is in general no longer lightlike but timelike:

$$\begin{aligned} k_j k_j &\equiv k_j^2 = e_j^2 - (\mathbf{k}_j)^2 = 0 \\ s_{12} &\equiv (k_1 + k_2)^2 = 2k_1 k_2 = 2e_1 e_2 (1 - \cos\theta_{12}) = 4e_1 e_2 \sin^2 \theta_{12}/2 > 0 \end{aligned} \tag{2.31}$$

unless the two lightlike vectors are parallel, which means that the angle between them $\theta_{12} = 0$.

It is always possible by means of a Lorentz boost to go to the *centre-of-mass system* (from now on the cms) of two lightlike or timelike vectors. This system is defined so that the total momentum vector vanishes. If the mass of the four-vector sum $\sqrt{s_{12}}$ from Eq. (2.31) is nonvanishing, the size of the velocity of the sum is less than c:

$$\mathbf{v}_{12} = \frac{\mathbf{k}_1 + \mathbf{k}_2}{|\mathbf{k}_1| + |\mathbf{k}_2|} \tag{2.32}$$

It is a useful exercise to prove to oneself that by a boost of $\mathbf{v}_{12}$ one reaches a Lorentz frame in which the two vectors in Eq. (2.31) have after the boost, the components

$$k'_{+1} = k'_{-2} = \sqrt{s_{12}};\ k'_{-1} = k'_{+2} = 0;\ \mathbf{k}'_{t1} = \mathbf{k}'_{t2} = \mathbf{0} \tag{2.33}$$

Thus they have 'oppositely' directed lightcone components in the cms. Another way to formulate this is to note that *a timelike vector may be uniquely partitioned into two lightlike vectors* (oppositely directed in space in the restframe of the timelike vector).

2 *Spacelike four-vectors*

If the invariant length in Eq. (2.3) (generalised possibly by means of Eq. (2.6)) is negative then the four-vector is called *spacelike*. An example of a spacelike vector in space-time is the difference vector between two points in space measured at the same time.

Actually, it is always possible for a spacelike vector in space-time, to find a frame such that the time component vanishes. To see this let us assume that in the situation described above involving the two observers A and B event (1) has a spacelike difference vector with respect to the origin, e.g.

$$0 < t_{1A} < x_{1A} \quad \text{and} \quad y_{1A} = z_{1A} = 0 \tag{2.34}$$

https://doi.org/10.1017/9781009401296.002 Published online by Cambridge University Press

(the sign choice of (t_{1A}, x_{1A}) being made for convenience). Then if the observer B moves at a velocity of size $v = t_{1A}/x_{1A}$ (although it appears to be a rather peculiar 'velocity' it is evidently smaller than $c = 1$) we obtain directly from (the first line of) Eq. (2.2) that event (1) will occur for B at the same time as he starts out from the origin.

For the observer B there is, however, a (space) distance between the origin and (1), that can be obtained from (the second line of) Eq. (2.2),

$$x_{1B} = \sqrt{x_{1A}^2 - t_{1A}^2} \tag{2.35}$$

i.e. the invariant length, as expected.

When the difference vector between two space-time points is spacelike then it is impossible to send any kind of signal between them. Therefore, it is impossible for two physical events occurring at the two points to be *causally connected*. The occurrence of one of the events cannot affect the occurrence of the other. We will in the course of this book have many occasions to come back to such situations.

The typical spacelike vectors in energy-momentum space correspond to *momentum transfers*. If two particles with rest masses m_1 and m_2 are scattered elastically from each other then in general there is a momentum transfer between them. Elastic scattering means that the same kinds of particle occur in the initial state and in the final state.

The energy-momentum vectors in the initial state, p_{ji}, and in the final state, p_{jf}, of the particles indexed $j = 1, 2$ are, however, in general different. Energy-momentum conservation means that

$$\sum_{j=1}^{2} p_{ji} = \sum_{j=1}^{2} p_{jf} \tag{2.36}$$

This implies that the difference vector, q, i.e. the momentum transfer between the two particles during the scattering, fulfils

$$q = p_{1f} - p_{1i} = -(p_{2f} - p_{2i}) \tag{2.37}$$

If we analyse the situation in the cms, with the two particles approaching each other along the x-axis with $\mathbf{p}_{1i} = p\mathbf{n}_x = -\mathbf{p}_{2i}$ (see Fig. 2.1) we conclude that

I The absolute sizes of the momenta of the final-state particles are the same as for the initial-state particles. To see this we note that

1 The total momentum in the cms vanishes also in the final state. Therefore the two final-state particles must have oppositely directed momentum vectors of equal size also.

https://doi.org/10.1017/9781009401296.002 Published online by Cambridge University Press

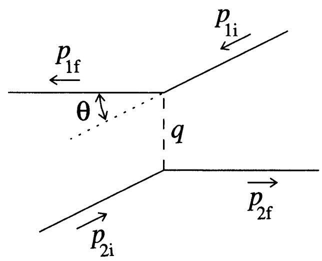

Fig. 2.1. Two particles experience elastic scattering against each other with notation described in the text.

2 Each of the particle energies is given by the momentum size, e.g. $e_{ji} = \sqrt{p^2 + m_j^2}$ and in order to conserve the total energy, cf. (2.36), the final-state momentum sizes therefore must be p, too.

II In the cms the momentum transfer four-vector, q, has no energy component, and we obtain for the invariant momentum transfer (conventionally called t or $-Q^2$)

$$-Q^2 = t \equiv q^2 = -4p^2 \sin^2(\theta/2) \simeq -p_t^2 \tag{2.38}$$

in terms of the scattering angle θ (see Fig. 2.1) and in the small angle limit, $\sin(\theta/2) \simeq (\sin\theta)/2$, in the last line with the transverse momentum $p_t = p\sin(\theta)$.

3 Minkowski space

The vector space endowed with the *metric* defined by the Lorentz-invariant four-vector product in Eqs. (2.3), (2.6) is called Minkowski space. Although ordinary space-time contains three space dimensions, it frequently occurs that physical models are formulated in lower-dimensional regions, corresponding to one- or two-dimensional space. (It is, of course, sometimes useful to make use of larger dimensions both for time and space but we shall not need to do so in this book.)

Minkowski space can be subdivided into the three different parts, considered above, i.e. into timelike, lightlike and spacelike points with respect to the origin (or for that matter with respect to any other point).

The lightlike vectors form three-dimensional regions, called *lightcones*, in between the other two classes, which are both four-dimensional. It is possible to further classify a lightcone into a *positive (forward)* part and a *negative (backward)* part, according to the sign of the time component, i.e.

$$t = \pm\sqrt{\mathbf{r}^2} \tag{2.39}$$

https://doi.org/10.1017/9781009401296.002 Published online by Cambridge University Press

In the same way timelike points can be inside the *forward* or the *backward* lightcones.

The significance of these notions is that it is always possible to reach a point inside, or on, the forward lightcone by means of a signal from the origin. In a similar way the origin can be reached from all the points inside, or on, the backward lightcone by means of a signal. All the spacelike points are, however, *non-causal* with respect to the origin, i.e. as mentioned above, events in the two points can have no dynamical influence on each other.

2.4 The electromagnetic field equations and some of their consequences

We will start with the notion of gauge invariance and after that turn to the properties of dielectrics. The rationale for introducing dielectrics is the following. The vacuum in a quantum theory, which intuitively corresponds to the no-particle state, behaves owing to quantum fluctuations in a way effectively similar to a dielectric medium.

1 Gauge invariance

The two Maxwell equations corresponding to Faraday's induction law and the absence of magnetic charges connect the electric field $\mathscr{E}$ and the magnetic field $\mathscr{B}$ in the following ways:

$$\nabla \times \mathscr{E} + \frac{\partial \mathscr{B}}{\partial t} = 0, \quad \nabla \cdot \mathscr{B} = 0 \tag{2.40}$$

These equations can be solved by introducing the four-vector potential $A_\mu \equiv (A_0, \mathbf{A})$:

$$\mathscr{B} = \nabla \times \mathbf{A}, \quad \mathscr{E} = -\nabla A_0 - \frac{\partial \mathbf{A}}{\partial t} \tag{2.41}$$

It is well known that these relations do not completely determine A_μ from a knowledge of $\mathscr{E}, \mathscr{B}$. It is always possible to introduce the change

$$\mathbf{A} \to \mathbf{A} + \nabla \Lambda, \quad A_0 \to A_0 - \frac{\partial \Lambda}{\partial t} \tag{2.42}$$

and still obtain the same electric and magnetic fields.

The transformation in Eq. (2.42) is a *local gauge transformation.* The word *local* means that it is possible to choose the function Λ so that it varies from point to point in space and time.

In somewhat loose language this means that the vector field A_μ contains redundant, non-observable, degrees of freedom and that one must by convention fix these degrees of freedom in order to be able to discuss its quantum properties.

Such gauge-fixing conventions of a more or less 'physical' kind have been suggested and used but it is essential to understand that one convention is, from a dynamical point of view, just as good as another. *Any observable result of a calculation must be gauge-independent.*

One should always remember when considering the emission of the quanta of A_μ that, with a certain gauge-fixing condition, the quanta may seem to be emitted from some particular part of the emitting current. It may well be the case, however, that the same observable quanta would seem to be emitted from a completely different part of the current if one were to use a different gauge condition (or as a matter of fact the same gauge condition but a different Lorentz frame). We will discuss these matters in more detail when we come to matter fields in Chapter 11 and to gluon radiation in Chapter 16.

If we introduce the energy-momentum-space quantities (we use the notation $\tilde{A}(q)$ or $\mathscr{A}(q)$ for the Fourier transform of a space-time quantity $A(x)$, with q the Fourier transform variable) a gauge transformation is

$$\tilde{A}(q) \to \tilde{A}(q) + iq\tilde{\Lambda}(q) \tag{2.43}$$

This means that, for a radiation field, when the vector potential $A = \epsilon \exp(ikx)$ describes a photonic quantum with energy-momentum vector k ($k^2 = 0$ for real photons) and polarisation vector ϵ, the physics results should be independent of the change

$$\epsilon \to \epsilon + ik\tilde{\Lambda}(k) \tag{2.44}$$

for any Λ.

In order to understand the relation in Eq. (2.44) we consider a boost along the direction of motion of the quantum, i.e. along the direction of $\mathbf{k}$. In the new frame the size of the momentum $|\mathbf{k}|$ and therefore also the energy are changed. For the polarisation vector ϵ this change can be compensated by a gauge transformation according to Eq. (2.44). Therefore in a charge-free region only the polarisation-vector components transverse to the direction of motion (that are invariant with respect to such boosts, i.e. those with $k\epsilon = 0$) are physically important (cf. the (brief) discussion of *helicity* in Chapter 5).

2 The notion of dielectrics

Besides the two equations mentioned above there are in Maxwell's treatment also Coulomb's and Ampère's laws, which tell us how to construct the fields from a knowledge of the charges and currents. They are expected to be precise in the microscopic sense (we use small letters to denote the microscopic fields and large letters for the corresponding macroscopic

https://doi.org/10.1017/9781009401296.002 Published online by Cambridge University Press

ones):

$$\nabla \cdot \mathbf{e} = \boldsymbol{\eta}, \quad \nabla \times \mathbf{b} - \frac{\partial \mathbf{e}}{\partial t} = \boldsymbol{\iota} \tag{2.45}$$

Here $\boldsymbol{\eta}$ and $\boldsymbol{\iota}$ are the 'local' charge and vector current densities, stemming from e.g. individual atomic charges. A quantum field does not really make sense as an operator acting as a single point (although with suitable care it is often possible to write quantum field operators in that way) because it is *distribution valued*. It should be smoothed out over a region by means of a 'test-function' f, [31]:

$$\mathbf{e}(f) \equiv \mathscr{E}(f) = \int dx f(x) \mathbf{e}(x) \tag{2.46}$$

We have here assumed that the test function f is nonvanishing (mathematically 'has support in') a region of suitable size around the point x. The typical atomic dimension is of the order of 10^{-8} cm (about twice the Bohr radius for hydrogen), and depending upon the system under consideration we may need this or other length units when we consider this averaging procedure. Jackson gives a lucid description, to which we refer the interested reader.

The result of the averaging procedure is, however, that not only 'the true' charges will affect the fields; there are also induced dipole moments, $\mathscr{P}$ and $\mathscr{M}$, stemming from the polarisation and magnetisation of the medium. The effective values of charge and current vector densities are thus changed; it is necessary to take into account also the polarisation charge, the polarisation current and the magnetic moment current. We then arrive at the macroscopic equations containing the free charge ($\boldsymbol{\rho}$) and current ($\mathbf{j}$) densities (the difference from the rapidly changing local $\boldsymbol{\eta}$ and $\boldsymbol{\iota}$ densities in Eq. (2.45), which describe individual atomic charges in motion, is that these microscopic fluctuations are averaged out, giving relatively smooth and slowly varying macroscopic quantities):

$$\begin{gathered} \nabla \cdot \mathbf{D} = \rho, \quad \nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{j} \\ \mathbf{D} = \mathscr{E} + \mathscr{P} = \epsilon \star \mathscr{E}, \quad \mathbf{H} = \mathscr{B} - \mathscr{M} = \frac{1}{\mu} \star \mathscr{B} \end{gathered} \tag{2.47}$$

Here $\mathbf{D}$ is the *electric displacement vector* and $\mathbf{H}$ is the *magnetic field*; ϵ and μ are the *dielectricity* and the *magnetic permeability*, of the material under investigation. The symbol $\star$ is used in order to indicate the possibility that, e.g.

$$\mathbf{D}(x) \equiv \epsilon \star \mathscr{E} = \int dx' \epsilon(x - x') \mathscr{E}(x') \tag{2.48}$$

https://doi.org/10.1017/9781009401296.002 Published online by Cambridge University Press

This would correspond to an energy-momentum-dependent displacement

$$\mathscr{D}(q) = [1 + \tilde{\xi}(q)]\mathscr{E}(q) \tag{2.49}$$

where we have introduced the *index of refraction* $\tilde{\xi} = \tilde{\epsilon} - 1$.

If we consider plane-wave solutions to the electromagnetic equations, (2.47), in a (true) charge- and current-free medium we may write (with the convention in classical physics that we are supposed to take the real part of all complex quantities)

$$\mathscr{E} = \mathscr{E}_0 \exp i(k\mathbf{n} \cdot \mathbf{x} - \omega t), \quad \mathscr{B} = \mathscr{B}_0 \exp i(k\mathbf{n} \cdot \mathbf{x} - \omega t) \tag{2.50}$$

We then obtain the following requirements:

$$k^2\mathbf{n}^2 - \tilde{\mu}\tilde{\epsilon}\omega^2 = 0, \quad \mathbf{n} \cdot \mathscr{E}_0 = \mathbf{n} \cdot \mathscr{B}_0 = 0, \quad \mathscr{B}_0 = \sqrt{\tilde{\mu}\tilde{\epsilon}}\mathbf{n} \times \mathscr{E}_0 \tag{2.51}$$

At this point we may consider a few limiting situations. Suppose firstly that $\tilde{\epsilon}$ is a constant and (for simplicity) $\tilde{\mu} = 1$. This means that **D** and **H** = **B** are completely local fields. We may in particular consider the vector **n** to be a unit vector. Then we will according to the last two equations of (2.51) have transverse waves in the medium. According to the first equation in (2.51) there is also a relation, usually referred to as a *dispersion relation*, between the wavenumber $k \simeq 1/\lambda$, with λ the wavelength, and the frequency ω.

To see what this relation implies we note that the transport velocity of the field energy-momentum is given by the ratio of the (space-time averaged) Poynting vector **S** ($|\mathbf{S}| \equiv S$) and the (space-time averaged) energy density u:

$$S = \frac{1}{2}|\mathscr{E} \times \mathscr{H}| = \frac{\sqrt{\tilde{\epsilon}}}{2}|\mathscr{E}|^2, \quad u = \frac{1}{4}(\tilde{\epsilon}\mathscr{E} \cdot \mathscr{E}^* + \mathscr{B} \cdot \mathscr{B}^*) = \frac{\tilde{\epsilon}}{2}|\mathscr{E}|^2 \tag{2.52}$$

The factor $\frac{1}{2}$ results from averaging the squared harmonic waves and we find in this way that the velocity has changed from $c = 1$ to $v = 1/\sqrt{\tilde{\epsilon}}$. Thus we require $\tilde{\epsilon} > 1$ in order that the transport velocity of the energy should not exceed the velocity of light in the vacuum. We note that the *phase velocity* of the waves, which is ω/k, then coincides with v.

Another case of interest is an electron plasma in the limit $\omega \gg \omega_p$, where ω_p is the plasma frequency. Then (cf. Jackson) $\tilde{\epsilon} = 1 - (\omega_p/\omega)^2$ and we obtain the same relation between k and ω as for a particle with mass ω_p (this is the only true Higgs-phenomenon we know of at present, i.e. the velocity of the electromagnetic waves in a medium is smaller than the vacuum velocity; this is tantamount to give a mass, corresponding to the plasma frequency, to the field quanta):

$$\omega^2 = k^2 + \omega_p^2 \tag{2.53}$$

In this case the phase velocity of the waves, ω/k, is greater than the

https://doi.org/10.1017/9781009401296.002 Published online by Cambridge University Press

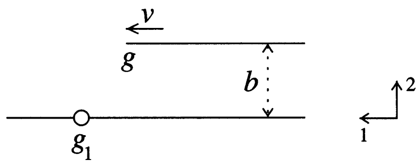

Fig. 2.2. A charged particle, g approaches a charged observer at the origin with velocity v along a direction with impact parameter b.

velocity of light. The true velocity, called the *group velocity*, is then instead the variation of ω with respect to k, $d\omega/dk = k/\omega < 1$, as we find by the well-known construction of local wave-packets from the waves in Eq. (2.50), cf. Jackson and Merzbacher. Consequently the index of refraction in Eq. (2.49) may be both positive and negative in real life situations.

We finally note that the index of refraction, $\tilde{\xi}$, may have an imaginary part. This corresponds to an absorption of the waves, i.e. to an interaction between the medium and the waves. There is a general set of relations, the Kramers-Kronig relations, [89], [88], between the real and the imaginary parts of the index of refraction. They stem from the causality requirement that there can be no effect until the waves have reached the medium. This leads to analyticity properties for $\tilde{\xi}$. We will meet the same properties in connection with the vacuum polarisation functions in quantum field theory in Chapter 4.

2.5 The method of virtual quanta

In this section we consider the electromagnetic field of a fast-moving charge and show how to express it in terms of its field quanta. The problem will be phrased as follows:

- Describe the field of an electric charge (size g), moving with velocity v along a direction (the 1-direction) having impact parameter b (for definiteness in the 12-plane) with respect to an observation point at the origin $x_1 = x_2 = 0$.

We assume that there is an observer, i.e. a detector carrying charge g_1, at the origin (Fig. 2.2). We expect that the approach of g will be noticeable as a pulse of radiation energy for this charged observer. This pulse will now be described in a semi-classical framework.

The Lorentz rest frames of the charges g and g_1 will be assumed to coincide at time $t_1 = t = 0$. Then we may calculate the Coulomb force field in the rest system of the charge g (where it is the usual spherically symmetric field falling off with distance R as $\propto 1/R^2$).

After that we may use the rules of special relativity and translate this field by a boost (with velocity $-v$) along the 1-axis to obtain the field components in the rest system of the charge g_1 (Jackson does it for us so we will not dwell upon the details):

$$\mathcal{E}_1 = -\frac{g\,v\,t\,\gamma}{r^3}, \quad \mathcal{E}_2 = -\frac{g\,b\,\gamma}{r^3}, \quad \mathcal{B}_3 = v\,\mathcal{E}_2 \tag{2.54}$$

(with r defined below in Eq. (2.55)). Note that the components in the 2- and 3-directions basically constitute a 'radiation field', i.e. $\mathcal{B} = \mathbf{v} \times \mathcal{E}$, when $v \sim c = 1$. We are now going to investigate that field.

The γ-factor is as usual $1/\sqrt{1-v^2}$ and the space extensions of the field components are Lorentz-contracted. Therefore, apart from the times $t \sim 0$, when the charges are close to each other, the distance r is a large number:

$$r = \sqrt{b^2 + (vt\gamma)^2} \tag{2.55}$$

The field components in Eq. (2.54) provide two Poynting-vector pulses, one along the 1-axis and one along the 2-axis. The latter is small and we will neglect it from now on. The main 1-axis radiation pulse is strongly Lorentz-contracted and looks like a bell-shaped curve in the time variable with a width (noticeable from Eq. (2.55)) around $t = 0$ of δt, where

$$\delta t = \frac{b}{v\gamma} \tag{2.56}$$

Note that this *typical passage time*, δt, can be written as

$$\delta t = \frac{mb}{p} \tag{2.57}$$

where m is the rest mass and $p \simeq e$ (for large $v \simeq c = 1$) are the momentum and energy of the charge g.

We can describe these results in terms of frequency (Jackson provides the exact formulas but we do not need the details). The differential intensity of the 1-axis pulse, $dI(\omega)$, where ω is the frequency will be essentially constant from a low-frequency value ω_{min} (where the wavelength becomes so long that there is nothing to observe) up to a maximum (determined by Eqs. (2.56), (2.57)):

$$\omega_{max} \simeq \frac{1}{\delta t} = \frac{p}{mb}. \tag{2.58}$$

This follows from the properties of the Fourier transform and also comes out of Jackson's formulas in terms of combinations of Bessel functions. We obtain approximately (note that the Poynting vector corresponds to the surface density of the field momentum)

$$dI(\omega, b) \simeq d\omega dA_t \frac{\hbar\alpha}{\pi^2 b^2} \tag{2.59}$$

For values of $\omega > 1/\delta t$ the distribution contains an exponential tail, with fast falloff. Here $dA_t = 2\pi b db$, i.e. the increase in the transverse area per unit impact parameter b. We have also defined the fine structure constant $\alpha = g^2/(4\pi\hbar c)$ under the assumption that g is a unit, i.e. electron, charge. We have been careful to keep Planck's constant in the expression (although we usually put $\hbar = 1$ according to the conventions in the Introduction) because up to now there has been no reference to quantum mechanics.

We may, however, now make the time-honoured transition to quantum mechanics by noting that for a fixed frequency ω the number of quanta, dn (in this case photons) in the pulse dI is given by

$$dI = \hbar\omega dn \tag{2.60}$$

This means that the whole field energy is carried by individual field quanta, each with an energy proportional to its frequency according to Einstein's proposal.

Therefore we have found an (approximate) expression for the number of field quanta which will be available for an interaction with the charge g_1 at the origin:

$$dn = \left(\frac{\alpha}{\pi}\right)\left(\frac{dA_t}{\pi b^2}\right)\frac{d\omega}{\omega} \tag{2.61}$$

This is basically a classical formula (but with quantum mechanics sneaked in through Eq. (2.60)). It describes the *flux factor* in connection with the interaction of the charged-particle field quanta. If the scattering cross section for the individual quanta is known then we simply multiply by this flux in order to get the cross section for the whole charged field.

Before the flux factor can be used we note, however, that it is singular in two different ways. The first way corresponds to the singularity for large wavelengths, $\omega \to 0$, to which reference already has been made. (The Lund model is everywhere infrared stable and we will therefore not consider the problems corresponding to infrared singularities. The main point is that when the number of quanta increases indefinitely at small frequencies then the dynamical behaviour is not given by their number but instead by their 'combined action', which corresponds to the action of a classical field.)

The second singularity is the logarithmic divergence for small values of b. This is a typical problem in all situations involving a charged particle. *It is necessary to define what is meant by the energy of the particle itself and what should be attributed to the field.* This is called *mass renormalisation*, i.e. it is necessary to provide the particle with a given rest energy equal to its mass, independently of the field surrounding it.

Classically the field energy from a point particle is always infinite and therefore after the discovery of the electron it was described not as a 'point' but as a small charged sphere with a radius $r_0 > 0$ such that its (Coulomb) field energy was exactly equal to the mass, m_e:

$$\frac{e^2}{4\pi r_0} = m_e \tag{2.62}$$

This quantity r_0, the classical electron radius, is approximately 3×10^{-15} m (using the conventions of $c = \hbar = 1$ to convert to metres) and occurs in the cross section for the interaction between an electron and low-frequency radiation, $\omega \to 0$:

$$\frac{d\sigma}{d\Omega} = \frac{r_0^2}{2}(1 + \cos^2\theta) \tag{2.63}$$

This is the Thompson cross section in the solid angle $d\Omega = \sin\theta\, d\theta d\phi$, where θ is the scattering angle and ϕ the azimuthal angle around the beam direction. It should, however, be understood that as far as we know (and this is at least down to 10^{-17} m because of the results of the LEP experiments at CERN) there is no extended space structure of the electron. The Thompson cross section therefore corresponds to the size of the Coulomb field around the particle rather than to some 'solid-sphere' behaviour.

The necessary cutoff in impact parameter depends upon the problem one is considering. It is either the Compton wavelength of the particle that is used or the characteristic size of the quantity that is probed by the field (but it is always the largest of the parameters). The Compton wavelength is $\lambda_C = \hbar/m$ and this b-cutoff therefore means that ω_{max} as defined in Eq. (2.58) will be given by

$$\omega < \omega_{max} = \frac{p}{m\lambda_C} = p \simeq e \tag{2.64}$$

This is not an unreasonable requirement. After all you cannot radiate away more energy than you have got!

The above representation is not normally used in connection with quantum field theory, where one usually describes the field not in terms of the energy and the impact parameter of the field quanta but instead in terms of their energy and transverse momentum.

The impact parameter vector $\mathbf{b}$ is, as we will see later in Chapter 10, the canonically conjugate variable to the transverse momentum $\mathbf{k}_t$ in a high-energy scattering event. Therefore one obtains the distribution of one from the other by means of a Fourier transform of the transition amplitude.

We note that the formulas above contain (as always for observables in quantum mechanics) the square of the amplitude (in this case $|\mathscr{E}|^2$) but from the scaling behaviour (no dimensional constants) we may guess that the relation between the distribution in impact parameter and the transverse momentum will be

$$\frac{2\pi b db}{\pi b^2} = \frac{db^2}{b^2} \rightarrow \frac{dk_t^2}{k_t^2} \tag{2.65}$$

and this turns out to be the right answer.

It is also conventional to rearrange the ω-dependence into a dependence upon the scaled variable $x = \omega/e$, e being the moving charged particle's energy. In that way we may write

$$dn = \left(\frac{\alpha}{\pi}\right)\left(\frac{dk_t^2}{k_t^2}\right)\frac{dx}{x} \tag{2.66}$$

which we will later meet as the spectrum for dipole bremsstrahlung radiation. The scaled variable x evidently has a range $x < 1$ according to Eq. (2.64).

Thus the method of virtual quanta (MVQ) redefines the interaction ability of a charged particle in terms of a flux of available (but virtual) field quanta, with precise properties with respect to interactions. Note that the word 'virtual' is appropriate: the field quanta are available but do not do anything until they find something to interact with.