

# Chapter 1

## Introduction

**Abstract** Sine-Gordon field as an effective description of a system of coupled pendulums in a constant gravitational field. Sine-Gordon solitons. The electromagnetic field, gauge potentials and gauge transformations. The Klein–Gordon equation and its solutions.

---

By definition, any physical system which has infinitely many degrees of freedom can be called a field. Systems with a finite number of degrees of freedom are called particles or sets of particles. The kinematics and dynamics of particles is the subject of classical and quantum mechanics. In parallel with these theories of particles there exist classical and quantum theories of fields. In this chapter we present two important examples of classical fields: the sine-Gordon effective field and the electromagnetic field.

Statistical mechanics deals with large ensembles of particles interacting with a thermal bath. If the particles are replaced by a field or a set of fields, the corresponding theory is called statistical field theory. This branch of field theory is not presented in our lecture notes.

### 1.1 Example A: Sine-Gordon Effective Field

Let us take a rectilinear, horizontal wire with  $M + N + 1$  pendulums hanging from it at equally spaced points labeled by  $x_i$ . Here  $i = -M, \dots, N$ , where  $M, N$  are natural numbers. The points  $x_i$  are separated by a constant distance  $a$ . The length of that part of the wire from which the pendulums are hanging is equal to  $(M + N)a$ . Each pendulum has a very light arm of length  $R$ , and a point mass  $m$  at the free end. It can swing only in the plane perpendicular to the wire. All pendulums are fastened to the wire stiffly, hence their swinging twists the wire (accordingly). The wire is elastic with respect to such twists. Each pendulum has one degree of freedom which may be represented by the angle  $\phi(x_i)$  between the vertical direction and the arm of the pendulum. All pendulums are subject to the constant gravitational force. In the configuration with the least energy all pendulums point downward and the wire is not

twisted. We adopt the convention that in this case the angles  $\phi(x_i)$  are equal to zero. Because of the presence of the wire  $\phi(x_i) = 0$  is not the same as  $\phi(x_i) = 2\pi k$ , where  $k = \pm 1, \pm 2, \dots$ —in the latter case the pendulum points downward but the wire is twisted, hence there is a non vanishing elastic energy. Therefore, the physically relevant range of  $\phi(x_i)$  is from minus to plus infinity.

The equation of motion for each pendulum, except for the first and the last ones, has the following form

$$mR^2 \frac{d^2 \phi(x_i, t)}{dt^2} = -mgR \sin \phi(x_i, t) + \kappa \frac{\phi(x_i - a, t) + \phi(x_i + a, t) - 2\phi(x_i, t)}{a}, \quad (1.1)$$

where  $\kappa$  is a constant which characterizes the elasticity of the wire with respect to twisting. The l.h.s. of this equation is the rate of change of the angular momentum of the  $i$ -th pendulum. The r.h.s. is the sum of all torques acting on the  $i$ -th pendulum: the first term is related to the gravitational force acting on the mass  $m$ , the second term represents the elastic torque due to the twist of the wire.

The equations of motion for the two outermost pendulums differ from (1.1) in a rather obvious way. In the following we shall assume that these two pendulums are kept motionless by some external force in the downward position, that is that

$$\phi(x_{-M}, t) = 0, \quad \phi(x_N, t) = 2\pi n, \quad (1.2)$$

where  $n$  is an integer. If we had put  $\phi(x_{-M}) = 2\pi l$  with integer  $l$  we could stiffly rotate the wire and all of the pendulums  $l$  times by the angle  $-2\pi$  in order to obtain  $l = 0$ . Therefore, the conditions (1.2) are the most general ones in the case of motionless, downward pointing outermost pendulums. In fact, these two pendulums can be removed altogether—we may imagine that the ends of the wire are tightly held in vices.

In order to predict the evolution of the system we have to solve (1.1) assuming certain initial data for the angles  $\phi(x_i, t)$ ,  $i = -M + 1, \dots, N - 1$ , and for the corresponding velocities  $\dot{\phi}(x_i, t)$ . This is a rather difficult task. Practical tools to be used here are numerical methods and computers. Numerical computations are useful if we ask for the solution of the equations of motion for a finite, and not too large, time interval. If we let the number of pendulums increase, sooner or later we will be incapable of predicting the evolution of the system except for very short time intervals, unless we restrict initial data in a special way. One such special case is in the limit of small oscillations around the least energy configuration,  $\phi(x_i) = 0$ . In this case we can linearize the equations of motion (1.1) using the approximation  $\sin \phi \approx \phi$ . The resulting equations are of the same type as those obtained for a system of coupled harmonic oscillators, treatments of which can be found in textbooks on classical mechanics.

It turns out that there is another special case which can be treated analytically. We call it the field theoretical limit because, as is explained below, we pass to an auxiliary system with an infinite number of degrees of freedom. Let us introduce a function  $\phi(x, t)$ , where  $x$  is a new real continuous variable (a coordinate along the

wire). By assumption, this function is at least twice differentiable with respect to  $x$ , and is such that its values at the points  $x = x_i$  are equal to the angles  $\phi(x_i, t)$  introduced earlier. Hence,  $\phi(x, t)$  smoothly interpolates between  $\phi(x_i, t)$  for each  $i$ . Of course, for a given set of the angles one can find infinitely many such functions. For any such function the following identity holds

$$\phi(x_i - a, t) + \phi(x_i + a, t) - 2\phi(x_i, t) = \int_0^a ds_1 \int_{-a}^0 ds_2 \frac{\partial^2 \phi(s_1 + s_2 + x, t)}{\partial x^2} \Big|_{x=x_i} .$$

Now comes the crucial assumption: we restrict our considerations to those motions of the pendulums for which there exists an interpolating function  $\phi(x, t)$  of continuous variables  $x$  and  $t$  which satisfies

$$\int_0^a ds_1 \int_{-a}^0 ds_2 \frac{\partial^2 \phi(s_1 + s_2 + x, t)}{\partial x^2} \Big|_{x=x_i} \approx a^2 \frac{\partial^2 \phi(x, t)}{\partial x^2} \Big|_{x=x_i} \quad (1.3)$$

for all times  $t$  and at all points  $x_i$ . For example, this is the case when the second derivative of  $\phi$  with respect to  $x$  is almost constant as  $x$  runs through the interval  $[x_i - a, x_i + a]$ , for all times  $t$ . With approximation (1.3) the identity written above can be replaced by the following approximate one

$$\phi(x_i - a, t) + \phi(x_i + a, t) - 2\phi(x_i, t) \approx a^2 \frac{\partial^2 \phi(x, t)}{\partial x^2} \Big|_{x=x_i} .$$

Using this formula in (1.1) we obtain

$$mR^2 \frac{d^2 \phi(x_i, t)}{dt^2} \approx -mgR \sin \phi(x_i, t) + \kappa a \frac{\partial^2 \phi(x, t)}{\partial x^2} \Big|_{x=x_i} . \quad (1.4)$$

Let us now suppose that our function  $\phi(x, t)$  obeys the following partial differential equation,

$$mR^2 \frac{\partial^2 \phi(x, t)}{\partial t^2} = -mgR \sin \phi(x, t) + \kappa a \frac{\partial^2 \phi(x, t)}{\partial x^2} , \quad (1.5)$$

where  $x \in [-Ma, Na]$ , and

$$\phi(-Ma, t) = 0, \quad \phi(Na, t) = 2\pi n, \quad (1.6)$$

where  $n$  is the same integer as in (1.2). Then, it is clear that  $\phi(x_i, t)$ ,  $i = -M + 1, \dots, N - 1$ , obey (1.4). Also the boundary conditions (1.2) are satisfied. Hence, if condition (1.3) is satisfied we obtain an approximate solution of the initial Newtonian equations (1.1).

The nonlinear partial differential equation (1.5) is well-known in mathematical physics by the jocular name ‘sine-Gordon equation’ which alludes to the Klein–Gordon equation. This latter equation is a cornerstone of relativistic field theory—we shall discuss it in Sect. 1.3. The sine-Gordon equation can be transformed into its standard form by dividing by  $mgR$ , and by rewriting it with the new, dimensionless variables

$$\tau = \sqrt{\frac{g}{R}}t, \quad \xi = \sqrt{\frac{mgR}{\kappa a}}x, \quad \Phi(\xi, \tau) = \phi(x, t).$$

The resulting standard form of the sine-Gordon equation reads

$$\frac{\partial^2 \Phi(\xi, \tau)}{\partial \tau^2} - \frac{\partial^2 \Phi(\xi, \tau)}{\partial \xi^2} + \sin \Phi(\xi, \tau) = 0. \quad (1.7)$$

There are many mathematical theorems about (1.7) and its solutions. One of them says that in order to determine a unique solution uniquely, one must specify the initial data, that is, one must fix the values of  $\Phi(\xi, \tau)$ , and  $\partial\Phi(\xi, \tau)/\partial\tau$  for a chosen instant of the rescaled time  $\tau = \tau_0$  and for all  $\xi$  in the interval  $[\xi_{-M}, \xi_N]$  (which corresponds to the interval  $[x_{-M}, x_N]$ ). One must also specify the so-called boundary conditions, that is the values of  $\Phi$  at the boundaries  $\xi = \xi_{-M}$  and  $\xi = \xi_N$  of the allowed range of  $\xi$  for all values of  $\tau$ . In our case their form follows from conditions (1.2),

$$\Phi(\xi_{-M}, \tau) = 0, \quad \Phi(\xi_N, \tau) = 2\pi n. \quad (1.8)$$

In order to specify the initial data we have to provide an infinite set of real numbers (to define the values of  $\Phi(\xi, \tau_0)$ ,  $\partial\Phi(\xi, \tau)/\partial\tau|_{\tau=\tau_0}$ ) because  $\xi$  is a continuous variable. For this reason the dynamical system defined by the sine-Gordon equation has an infinite number of degrees of freedom. This system, called the sine-Gordon field, is mathematically represented by the function  $\Phi$ , and the sine-Gordon equation is its equation of motion. The sine-Gordon field is said to be the effective field for the system of pendulums described above. Let us emphasize that the sine-Gordon effective field gives an accurate description of the dynamics of the original system only if condition (1.3) is satisfied. Such a reduction of the original problem to the dynamics of an effective field, or to a set of effective fields in other cases, has become an extremely efficient tool in theoretical investigations of many physical systems considered in condensed matter physics or particle physics.

Let us end this section with a few examples of nontrivial solutions of the sine-Gordon equation in its standard form (1.7). Let us assume that  $\Phi$  does not depend on the rescaled time  $\tau$ , that is, that  $\Phi = \Phi(\xi)$ —such solutions are referred to as static. Then, (1.7) reduces to the following ordinary differential equation

$$\Phi''(\xi) = \sin \Phi(\xi), \quad (1.9)$$

where  $'$  denotes differentiation with respect to  $\xi$ . Multiplying this equation by  $\Phi'$  we obtain

$$\frac{1}{2}(\Phi'^2)' = -(\cos \Phi)',$$

and after integration,

$$\frac{1}{2}\Phi'^2 = c_0 - \cos \Phi, \quad (1.10)$$

where  $c_0$  is a constant. The boundary conditions (1.8) imply that

$$c_0 = 1 + \frac{1}{2}\Phi'^2(\xi_{-M}) = 1 + \frac{1}{2}\Phi'^2(\xi_N). \quad (1.11)$$

It follows that  $c_0 \geq 1$ , and that  $\Phi'(\xi_{-M}) = \pm\Phi'(\xi_N)$ .

Let us first consider the case  $c_0 = 1$ . The square root of (1.10) with  $c_0 = 1$  gives

$$\Phi' = 2 \sin\left(\frac{\Phi}{2}\right), \quad (1.12)$$

or

$$\Phi' = -2 \sin\left(\frac{\Phi}{2}\right), \quad (1.13)$$

which can be easily integrated. Apart from the trivial solution  $\Phi = 0$ , there exist nontrivial solutions, denoted below by  $\Phi_+$  and  $\Phi_-$ . Integrating (1.12) and (1.13) we find that

$$\ln \left| \tan\left(\frac{\Phi}{4}\right) \right| = \pm(\xi - \xi_0),$$

where  $\xi_0$  is an arbitrary constant, and the signs  $+$  and  $-$  correspond to (1.12) and (1.13), respectively. It follows that

$$\Phi_{\pm}(\xi) = \pm 4 \arctan[\exp(\pm(\xi - \xi_0))] \bmod 4\pi. \quad (1.14)$$

Formula (1.14) implies that  $\Phi'_{\pm}(\xi) \neq 0$  for all finite  $\xi$ , and  $\Phi'_{\pm}(\xi) \rightarrow 0$  if  $\xi \rightarrow \infty$  or  $\xi \rightarrow -\infty$ . Therefore, conditions (1.11) can only be satisfied if

$$\xi_{-M} = -\infty, \quad \xi_N = +\infty.$$

With the help of the identity

$$\arctan(1/x) = \pi/2 - \arctan x,$$

one can show that formula (1.14) in fact gives two solutions which obey the conditions (1.8):

$$\Phi_{\pm}(\xi) = \pm 4 \arctan(\exp(\xi - \xi_0)). \quad (1.15)$$

It is clear that

$$\lim_{\xi \rightarrow +\infty} \Phi_{\pm}(\xi) = \pm 2\pi.$$

Hence, the integer  $n$  in (1.8) can be equal to 0 or  $\pm 1$  ( $n = 0$  corresponds to the trivial solution  $\Phi = 0$ ).

Let us summarize the case for which  $c_0 = 1$ . Static solutions obeying the boundary conditions (1.8) exist only if the range of  $\xi$  is infinite, from  $-\infty$  to  $+\infty$ , and then the nontrivial solutions have the form (1.15). The solution  $\Phi_+$  is called the soliton, and  $\Phi_-$  the antisoliton.  $\xi_0$  is called the location of the (anti-)soliton. There are no static solutions with  $|n| > 1$ .

Coming back to our system of pendulums, the solitonic solutions (1.15) are relevant if the condition (1.3) is satisfied. The two integrals on the l.h.s. of condition (1.3) can be rewritten as integrals of  $\Phi''_{\pm}$  with respect to the dimensionless variables

$$\xi_{1,2} = \sqrt{\frac{mgR}{\kappa a}} x_{1,2}.$$

Then, the integration limits are given by 0 and  $\pm\alpha$ , where

$$\alpha = \sqrt{\frac{mgRa}{\kappa}}.$$

We see that condition (1.3) is certainly satisfied if

$$\alpha \rightarrow 0,$$

because in this limit the range of integration shrinks to a point. The value of the dimensionless parameter  $\alpha$  can be made small by, e.g., choosing a wire with large  $\kappa$  or by putting the pendulums close to each other (small  $a$ ). Furthermore, note that  $\xi_N = \alpha x_N/a$ ,  $\xi_{-M} = \alpha x_{-M}/a$ ,  $x_{-M} = -Ma$  and;  $x_n = Na$ . It follows that  $\xi_N, \xi_{-M}$  can tend to  $\pm\infty$ , respectively, in the limit  $\alpha \rightarrow 0$  only if  $N, M \rightarrow \infty$ . Thus, the number of pendulums has to be very large.

The case  $c_0 > 1$  is a little bit more complicated. Equation (1.10) is equivalent to the following equations

$$\Phi' = \pm\sqrt{2c_0 - 2\cos\Phi}, \quad (1.16)$$

which give the following relations

$$\int_0^{\Phi(\xi)} ds \frac{1}{\sqrt{1 - c_0^{-1} \cos s}} = \pm\sqrt{2c_0}(\xi - \xi_{-M}). \quad (1.17)$$

These relations implicitly define the functions  $\Phi(\xi)$  which obey (1.9). The integral on the l.h.s. of (1.17) can be related to an elliptic integral of the first kind (see, e.g.

[1]), and  $\Phi(\xi)$  is then given by the inverse of the elliptic function. The constant  $c_0$  is determined from the following equation, obtained by inserting the second of the boundary conditions (1.8) into formula (1.17):

$$\int_0^{2\pi n} ds \frac{1}{\sqrt{1 - c_0^{-1} \cos s}} = \pm \sqrt{2c_0} (\xi_N - \xi_{-M}). \quad (1.18)$$

Note that now  $\xi_{-M}$  and  $\xi_N$  have to be finite, otherwise the r.h.s. of this equation would be meaningless.

One may also solve (1.16) numerically. These equations are rather simple and can easily be tackled by computer algebra systems like *Maple*® or *Mathematica*®. Equations (1.16) are considered on the interval  $(\xi_{-M}, \xi_N)$ . They are formally regarded as evolution equations with  $\xi$  playing the role of time. The boundary condition  $\Phi(\xi_{-M}) = 0$  is now regarded as the initial condition for  $\Phi(\xi)$ . The constant  $c_0$  is adjusted by trial and error until the calculation gives  $\Phi(\xi_N) \approx 2\pi n$  to the desired accuracy. For example, choosing  $\xi_{-M} = -10$  and  $\xi_N = 10$  we have obtained  $c_0 \approx 1.00000008$  for  $n = \pm 1$ ,  $c_0 \approx 1.0014$  for  $n = \pm 2$ , and  $c_0 \approx 1.0398$  for  $n = \pm 3$ .

These solutions of the sine-Gordon equation with  $c_0 > 1$  are pertinent to the physics of the system of pendulums when the parameter  $\alpha$  has a sufficiently small value, as in the case  $c_0 = 1$ . For given natural numbers  $N$  and  $M$ , the values of  $\xi_{-M}$  and  $\xi_N$  are calculated from formulas  $\xi_{-M} = -\alpha M$  and  $\xi_N = \alpha N$ . In the limit  $\alpha \rightarrow 0$  with  $\xi_{-M}$  and  $\xi_N$  kept non vanishing and constant, the number of pendulums has to increase indefinitely.

## 1.2 Example B: The Electromagnetic Field

We have just seen an example of an effective field—the sine-Gordon field  $\phi(x, t)$ —introduced in order to provide an approximate description of our original physical system: the set of coupled pendulums. Now we shall see an example from another class of fields, called fundamental fields. Such fields are regarded as elementary dynamical systems—according to present day physics there are no experimental indications that they are effective fields for an underlying system. The fundamental fields appear in particular in particle physics and cosmology. Later on we shall see several such fields. Here we briefly recall the classical electromagnetic field. It should be stressed that we regard this field as a physical entity, a part of the material world. Our main goal is to show that the Maxwell equations can be reduced to a set of uncoupled wave equations.

According to 19th century physics, the electromagnetic field is represented by two functions  $\vec{E}(t, \vec{x})$ ,  $\vec{B}(t, \vec{x})$ , the electric and magnetic fields respectively. Here  $\vec{x}$  is a position vector in the three dimensional space  $R^3$ , and  $t$  is time. The fields obey the Maxwell equations of the form (we use the rationalized Gaussian units)

$$\begin{aligned}
 \text{(a)} \quad \operatorname{div} \vec{E} &= \rho, & \text{(c)} \quad \operatorname{div} \vec{B} &= 0, \\
 \text{(b)} \quad \operatorname{rot} \vec{B} - \frac{1}{c} \frac{\partial \vec{E}}{\partial t} &= \frac{1}{c} \vec{j}, & \text{(d)} \quad \operatorname{rot} \vec{E} + \frac{1}{c} \frac{\partial \vec{B}}{\partial t} &= 0,
 \end{aligned}
 \tag{1.19}$$

where  $\rho$  is the electric charge density, and  $\vec{j}$  is the electric current density.  $\rho$  and  $\vec{j}$  are functions of  $t$  and  $\vec{x}$ , and  $c$  is the speed of light in the vacuum.

Suppose that there exist fields  $\vec{E}(t, \vec{x})$ ,  $\vec{B}(t, \vec{x})$  obeying Maxwell equations (1.19). Acting with the div operator on (1.19b), then using the identity  $\operatorname{div}(\operatorname{rot}) \equiv 0$  and (1.19a), we obtain the following condition on the charge and current density

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \vec{j} = 0.
 \tag{1.20}$$

This is a well-known continuity equation. It is equivalent to conservation of electric charge. From the mathematical viewpoint, it should be regarded as a consistency condition for the Maxwell equations—if it is not satisfied they do not have any solutions.

Equation (1.19c) is satisfied by any field  $\vec{B}$  of the form

$$\vec{B} = \operatorname{rot} \vec{A},
 \tag{1.21}$$

where  $\vec{A}(t, \vec{x})$  is a (sufficiently smooth) function of  $\vec{x}$ . Conversely, one can prove that any field  $\vec{B}$  which obeys (1.19c) has the form (1.21). From (1.21) and (1.19d) follows the identity

$$\operatorname{rot} \left( \vec{E} + \frac{1}{c} \frac{\partial \vec{A}}{\partial t} \right) = 0.$$

There is a mathematical theorem (the Poincaré lemma) which says that an identity of the form  $\operatorname{rot} \vec{X} = 0$  implies that the vector function  $\vec{X}$  is the gradient of a scalar function  $\sigma$ , i.e.  $\vec{X} = \nabla \sigma$ . Therefore, there exists a function  $A_0$  such that

$$\vec{E} + \frac{1}{c} \frac{\partial \vec{A}}{\partial t} = -\nabla A_0$$

(the minus sign is dictated by tradition). Thus,

$$\vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} - \nabla A_0.
 \tag{1.22}$$

The functions  $A_0$  and  $\vec{A}$  are called gauge potentials for the electromagnetic field. Note that the choice of  $A_0$  and  $\vec{A}$  for a given electric and magnetic fields is not unique—instead of  $A_0$  and  $\vec{A}$  one may just as well take

$$\vec{A}'(t, \vec{x}) = \vec{A}(t, \vec{x}) - \nabla \chi(t, \vec{x}), \quad A'_0(t, \vec{x}) = A_0(t, \vec{x}) + \frac{1}{c} \frac{\partial \chi(t, \vec{x})}{\partial t},
 \tag{1.23}$$



where  $\chi(t, \vec{x})$  is a sufficiently smooth but otherwise arbitrary function of the indicated variables. This freedom of choosing the gauge potentials is called the gauge symmetry. Formulas (1.23) can be regarded as transformations of the gauge potentials, and are called the gauge transformations. Often they are called local gauge transformations in order to emphasize the fact that the function  $\chi$  is space and time dependent. One should keep in mind the fact that the gauge transformations appear because we adopt the mathematical description of the electromagnetic field in terms of the potentials. The fields  $\vec{E}$  and  $\vec{B}$  do not change under these transformations. The potentials  $A_0, \vec{A}$  and  $A'_0, \vec{A}'$  from formulas (1.23) describe the same physical situation. The freedom of performing the gauge transformations means that the potentials form a larger than necessary set of functions for describing a given physical configuration of the electromagnetic field. Nevertheless, it turns out that the description in terms of the potentials is a most economical one, especially in quantum theories of particles or fields interacting with the electromagnetic field. In fact, it has been commonly accepted that the best mathematical representation of the electromagnetic field—one of the basic components of the material world—is given by the gauge potentials  $A_0$  and  $\vec{A}$ .

Expressing  $\vec{E}$  and  $\vec{B}$  by the gauge potentials we have explicitly solved (1.19c, d). Now let us turn to (1.19a, b). First, we use the gauge transformations to adjust the vector potential  $\vec{A}$  in such a way that

$$\operatorname{div} \vec{A} = 0. \quad (1.24)$$

This condition is known as the Coulomb gauge condition. One can easily check that for any given  $\vec{A}$  one can find a gauge function  $\chi$  such that  $\vec{A}'$  obeys the Coulomb condition, provided that  $\operatorname{div} \vec{A}$  vanishes sufficiently quickly at the spatial infinity. For that matter, let us note that from a physical viewpoint it is sufficient to consider electric and magnetic fields which smoothly<sup>1</sup> vanish at the spatial infinity. For such fields there exist potentials  $A_0$  and  $\vec{A}$  which also smoothly vanish as  $|\vec{x}| \rightarrow \infty$ . It is quite natural to assume that the gauge transformations leave the potentials within this class. Therefore, we assume that the gauge function  $\chi$  also smoothly vanishes at the spatial infinity. We might have assumed that it could approach a non vanishing constant in that limit. However, such a constant leads to a trivial gauge transformation because then the derivatives present in formulas (1.23) vanish. For this reason it is natural to choose this constant equal to zero. Note that now the Coulomb gauge condition determines the gauge completely. By this we mean that if both  $\vec{A}$  and  $\vec{A}'$ , which are related by the local gauge transformation (1.23), obey the Coulomb gauge condition, then  $\chi = 0$ , that is the two potentials coincide. This follows from the facts that if (1.24) is satisfied by  $\vec{A}$  and  $\vec{A}'$  then  $\chi$  obeys the Laplace equation,  $\Delta \chi = 0$ , and the only nonsingular solution of this equation which vanishes at the spatial infinity is  $\chi = 0$ .

---

<sup>1</sup>Here this means that all derivatives of the fields with respect to the Cartesian coordinates  $x^i$  also vanish at the spatial infinity.

The condition that  $\chi$  vanishes at the spatial infinity is also welcome for another reason—it makes a clear distinction between (local) gauge transformations and global transformations. Global transformations will be introduced in Chap. 3. They are given by  $\chi$  which are constant in time and space. Such transformations can act nontrivially on fields other than the electromagnetic field. With the definitions we have adopted, the global transformations are not contained in the set of gauge transformations.

Equations (1.19a, b) are reduced in the Coulomb gauge to the following equations

$$\Delta A_0 = -\rho, \quad \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \Delta \vec{A} + \frac{1}{c} \nabla \frac{\partial A_0}{\partial t} = \frac{1}{c} \vec{j}. \quad (1.25)$$

The solution of the first equation has the form

$$A_0(t, \vec{x}) = \frac{1}{4\pi} \int d^3x' \frac{\rho(t, \vec{x}')}{|\vec{x} - \vec{x}'|}, \quad (1.26)$$

provided that  $\rho$  vanishes sufficiently quickly at the spatial infinity to ensure that the integral is convergent. The r.h.s. of formula (1.26) is often denoted by  $-\Delta^{-1}\rho$ . Because the potential  $A_0$  is just given by integral (1.26)—there is not any evolution equation for it to be solved—it is not a dynamical variable. In the final step, formula (1.26) is used to eliminate  $A_0$  from the second of the equations (1.25). We also eliminate  $\partial\rho/\partial t$  with the help of continuity equation (1.20). The resulting equation for  $\vec{A}$  can be written in the form

$$\frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \Delta \vec{A} = \frac{1}{c} \vec{j}_T, \quad (1.27)$$

where

$$\vec{j}_T = \vec{j} - \nabla(\Delta^{-1}\text{div}\vec{j}), \quad (1.28)$$

and

$$\Delta^{-1}\text{div}\vec{j}(t, \vec{x}) = -\frac{1}{4\pi} \int d^3x' \frac{\text{div}\vec{j}(t, \vec{x}')}{|\vec{x} - \vec{x}'|}.$$

Of course, we assume that  $\text{div}\vec{j}$  vanishes sufficiently quickly at the spatial infinity.  $\vec{j}_T$  is called the transverse part of the external current  $\vec{j}$ . The reason for such a name is that

$$\text{div}\vec{j}_T \equiv 0, \quad (1.29)$$

as it immediately follows from the definition of  $\vec{j}_T$ . For the same reason, the potential  $\vec{A}$  which obeys the Coulomb gauge condition is called the transverse vector potential.

Note that identity (1.29) is a necessary condition for the existence of the solutions of (1.27)—applying the div operator to both sides of (1.27) and using the Coulomb condition we would obtain a contradiction if (1.29) were not true.

To summarize, the set of Maxwell equations (1.19) has been reduced to (1.27) together with the Coulomb gauge condition (1.24). Equation (1.27) determines the time evolution of the electromagnetic field. It plays the same role as Newton's equation in classical mechanics. From a mathematical viewpoint, equation (1.27) is a set of three linear, inhomogeneous, partial differential equations: one equation for each component  $A^i$  of the vector potential.<sup>2</sup> These equations are decoupled, that is they can be solved independently from each other. They are called wave equations.

As in the case of the sine-Gordon equation (1.7), in order to uniquely determine a solution of (1.27) we have to specify the initial data at the time  $t_0$ :

$$\vec{A}(t_0, \vec{x}) = \vec{f}_1(\vec{x}), \quad \left. \frac{\partial \vec{A}(t, \vec{x})}{\partial t} \right|_{t=t_0} = \vec{f}_2(\vec{x}), \quad (1.30)$$

where  $\vec{f}_1$  and  $\vec{f}_2$  are given vector fields, vanishing at the spatial infinity. Moreover, in order to ensure that the Coulomb gauge condition is satisfied at the time  $t = t_0$ , we assume that

$$\operatorname{div} \vec{f}_1 = 0, \quad \operatorname{div} \vec{f}_2 = 0. \quad (1.31)$$

It turns out that conditions (1.31) and equation (1.27) imply that  $\operatorname{div} \vec{A} = 0$  for all times  $t$ . The point is that equation (1.27) implies that  $\operatorname{div} \vec{A}$  obeys the homogeneous equation

$$\frac{1}{c^2} \frac{\partial^2 \operatorname{div} \vec{A}}{\partial t^2} - \Delta(\operatorname{div} \vec{A}) = 0.$$

Due to the assumptions (1.31) the initial data for this equation are homogeneous ones, that is

$$\operatorname{div} \vec{A}|_{t=t_0} = 0, \quad \partial_t \operatorname{div} \vec{A}|_{t=t_0} = 0,$$

where  $\partial_t$  is a short notation for the partial derivative  $\partial/\partial t$ . We shall see in the next section that all this implies that

$$\operatorname{div} \vec{A} = 0$$

for all times. In consequence, we do not have to worry about the Coulomb gauge condition provided that the initial data (1.30) obey the conditions (1.31)—the Coulomb gauge condition has been reduced to a constraint on the initial data.

---

<sup>2</sup>We adhere to the convention that vectors denoted by the arrow have components with upper indices.

### 1.3 Solutions of the Klein–Gordon Equation

The considerations of the electromagnetic field have led us to an evolution equation of the form

$$\square \phi = \eta(t, \vec{x}), \quad (1.32)$$

where

$$\square \equiv \Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2},$$

$\phi$  is a function of  $(t, \vec{x})$ , and  $\eta$  is an a priori given function, called the source. The wave equation (1.32) is a particular case of the more general Klein–Gordon equation

$$\square \phi - m^2 \phi = \eta(t, \vec{x}), \quad (1.33)$$

where  $m^2$  is a real, non-negative constant of the dimension  $\text{cm}^{-2}$ , and  $\phi$  is a real or complex function. The Klein–Gordon equation is the basic evolution equation in relativistic field theory. It also appears in non relativistic settings. For example, sine-Gordon equation (1.7) reduces to the Klein–Gordon equation with just one spatial variable  $\xi$  if we consider  $\Phi$  close to 0, because in this case  $\sin \Phi$  can be approximated by  $\Phi$ . Therefore, one should be acquainted with solutions of the Klein–Gordon equation.

Let us introduce concise, four-dimensional relativistic notation:

$$x = (ct, \vec{x}), \quad k = (k_0, \vec{k}), \quad kx = ck_0t - \vec{k}\vec{x}, \quad d^4x = cd^3xdt, \quad d^4k = d^3kdk_0.$$

Here  $k_0$  is a real variable, and  $\vec{k}$  is a real 3-dimensional vector called the wave vector.  $k_0$  and  $\vec{k}$  have the dimension  $\text{cm}^{-1}$ .  $\omega = ck_0$  is a frequency. Furthermore, we shall often use  $x^0 = ct$  instead of the time variable  $t$  and call it time too. This notation reflects the Lorentz invariant structure of space-time. In particular, the form of  $kx$  corresponds to the diagonal metric tensor of the space-time  $(\eta_{\mu\nu}) = \text{diag}(1, -1, -1, -1)$ , where  $\text{diag}$  denotes the diagonal matrix with the listed elements on its diagonal. Note that  $kx$  is dimensionless.

Because the Klein–Gordon equation is linear with respect to  $\phi$  and has constant coefficients, we may use the Fourier transform technique for solving it. We denote by  $\tilde{\phi}(k)$  the Fourier transform of  $\phi(x)$ . It is defined as follows:

$$\tilde{\phi}(k) = \int d^4x e^{ikx} \phi(x). \quad (1.34)$$

The inverse Fourier formula has the form

$$\phi(x) = \frac{1}{(2\pi)^4} \int d^4k e^{-ikx} \tilde{\phi}(k). \quad (1.35)$$

Analogously,

$$\tilde{\eta}(k) = \int d^4x e^{ikx} \eta(x).$$

The Klein–Gordon equation is equivalent to the following algebraic (not differential!) equation for  $\tilde{\phi}$

$$(k_0^2 - \vec{k}^2 - m^2) \tilde{\phi}(k) = \tilde{\eta}(k). \quad (1.36)$$

Its solutions should be sought in a space of generalized functions. An excellent introduction to the theory of generalized functions with its applications to linear partial differential equations can be found in, e.g., [2]. Some pertinent facts can be found in Appendix A.

One can prove that the most general solution of (1.36) has the form

$$\tilde{\phi}(k) = \text{“} \frac{\tilde{\eta}(k)}{k_0^2 - \vec{k}^2 - m^2} \text{”} + C(k_0, \vec{k}) \delta(k_0^2 - \vec{k}^2 - m^2), \quad (1.37)$$

where  $C(k_0, \vec{k})$  is an arbitrary smooth function of the indicated variables. The first term on the r.h.s. denotes a particular solution of the inhomogeneous equation (1.36). We have put the quotation marks around it because in fact that term written as it stands is not correct. We explain and solve this problem shortly. The second term on the r.h.s. gives the general solution of the homogeneous equation

$$(k_0^2 - \vec{k}^2 - m^2) \tilde{\phi}(k) = 0.$$

Formula (1.37) is in accordance with the well-known fact that the general solution of an inhomogeneous linear equation can always be written as the sum of a particular solution of that equation and of a general solution to the corresponding homogeneous equation.

The problem with the term in quotation marks is that it is not a generalized function. In consequence, its Fourier transform, formula (1.35), does not have to exist, and indeed, it does not exist. One can see this easily by looking at the integral over  $k_0$ —there are non integrable singularities of the integrand at  $k_0 = \pm\omega(\vec{k})/c$ , where

$$\omega(\vec{k}) = c \sqrt{\vec{k}^2 + m^2}. \quad (1.38)$$

In order to obtain the correct formula for the solution we first find a generalized function  $\tilde{G}(k)$  which obeys the equation

$$(k_0^2 - \vec{k}^2 - m^2) \tilde{G}(k) = 1. \quad (1.39)$$

The corresponding  $G(x)$  is calculated from a formula analogous to (1.35). It obeys the following equation

$$(\square - m^2) G(x) = \delta(x), \quad (1.40)$$

and is called the Green's function of the Klein–Gordon equation. Knowing  $\tilde{G}(k)$ , we may replace the “ ” term by the mathematically correct expression

$$\text{“ } \frac{\tilde{\eta}(k)}{k_0^2 - \vec{k}^2 - m^2} \text{”} \rightarrow \tilde{\eta}(k)\tilde{G}(k),$$

provided that  $\tilde{\eta}$  is a smooth function of  $k^0$  and  $\vec{k}$ .

Important Green's functions for the Klein–Gordon equation have Fourier transforms of the form

$$\tilde{G}(k) = \frac{c^2}{2\omega(\vec{k})} \left( \frac{1}{ck_0 - \omega(\vec{k}) \pm i0_+} - \frac{1}{ck_0 + \omega(\vec{k}) \pm i0_+} \right). \quad (1.41)$$

The meaning of the symbol  $\pm i0_+$  is explained in the Appendix. The choice  $+i0_+$  in both terms of formula (1.41) gives the so called retarded Green's function

$$G_R(x - y) = \frac{c^2}{(2\pi)^4} \int d^4k \frac{e^{-ik(x-y)}}{2\omega(\vec{k})} \left( \frac{1}{ck_0 - \omega(\vec{k}) + i0_+} - \frac{1}{ck_0 + \omega(\vec{k}) + i0_+} \right). \quad (1.42)$$

The integral over  $k_0$  can be performed with the help of contour integration in the plane of complex  $k_0$ . The trick consists of completing the line of real  $k_0$  to a closed contour by adding upper (lower) semicircle with the center at  $k_0 = 0$  and infinite radius when  $x^0 - y^0 < 0$  ( $x^0 - y^0 > 0$ ). We obtain

$$G_R(x - y) = \frac{-ic}{2(2\pi)^3} \Theta(x^0 - y^0) \int \frac{d^3k}{\omega(\vec{k})} \left( e^{-ik(x-y)} - e^{ik(x-y)} \right) \Bigg|_{k_0=\omega(\vec{k})/c}, \quad (1.43)$$

where  $\Theta(x^0 - y^0)$  denotes the Heaviside step function.<sup>3</sup>

The Green's function  $G_R$  is used in order to obtain a particular solution of the inhomogeneous Klein–Gordon equation, denoted below by  $\phi_\eta$ . Namely,

$$\phi_\eta(x) = \int d^4y G_R(x - y)\eta(y). \quad (1.44)$$

This solution is causal in the classical sense: the values of  $\phi_\eta(x^0, \vec{x})$  at a certain fixed instant  $x^0$  are determined by values of the external source  $\eta(y^0, \vec{y})$  at earlier times, i.e.,  $y^0 \leq x^0$ . More detailed analysis shows that the contributions come only from the interior and boundaries of the past light-cone with its tip at the point  $x$ , that is,

---

<sup>3</sup> $\Theta(x) = 1$  for  $x > 0$ ,  $\Theta(x) = 0$  for  $x < 0$ . The value of  $\Theta(0)$  does not have to be specified because the step function is used under the integral. Formally, the step function is a generalized function, and for such functions their values at a given single point are not defined. Therefore, the question, “what is the value of  $\Theta(0)$ ?” is meaningless.

from  $y$  such that  $(x - y)^2 \geq 0$  and  $x^0 - y^0 \geq 0$ . This can be seen from the following formula, see Appendix 2 in [3],

$$G_R(x) = -\frac{1}{2\pi} \Theta(x^0) \left[ \delta(x^2) - \Theta(x^2) \frac{m}{2\sqrt{x^2}} J_1(m\sqrt{x^2}) \right],$$

where  $x^2 = (x^0)^2 - \vec{x}^2$ , and  $J_1$  is a Bessel function. Therefore, waves of the field emitted from a spatially localized source  $\eta$  travel with velocity not greater than the velocity of light in the vacuum  $c$ . Choosing the  $-i0_+$  in formula (1.41) we would obtain the so called advanced Green's function, which is anti-causal—in this case  $\phi_\eta(x)$  is determined by values of  $\eta(y)$  in the future light cone,  $y^0 \geq x^0$  and  $(x - y)^2 \geq 0$ . In general, the choice of Green's function is motivated by the underlying physical problem. On purely mathematical grounds there are infinitely many Green's functions. All have the form  $G_R(x) + \phi_0(x)$ , where  $\phi_0(x)$  is a particular solution of the homogeneous Klein-Gordon equation.

Now that we have found a particular solution for the inhomogeneous Klein-Gordon equation, let us turn our attention to finding the general solution of the homogeneous Klein-Gordon equation. The second term in formula (1.37) gives

$$\phi_0(x) = \frac{1}{(2\pi)^4} \int d^4k e^{-ikx} C(k^0, \vec{k}) \delta(k_0^2 - \vec{k}^2 - m^2). \quad (1.45)$$

With the help of formula

$$\delta(k_0^2 - \vec{k}^2 - m^2) = \frac{\delta(k_0 - \omega(\vec{k})/c)}{2\omega(\vec{k})/c} + \frac{\delta(k_0 + \omega(\vec{k})/c)}{2\omega(\vec{k})/c},$$

$\phi_0$  can be written in the form

$$\phi_0(x) = \int \frac{d^3k}{\sqrt{2(2\pi)^3\omega(\vec{k})}} \left( a_+(\vec{k}) e^{-ikx} + a_-(\vec{k}) e^{ikx} \right) \Big|_{k_0=\omega(\vec{k})/c}, \quad (1.46)$$

where

$$a_\pm(\vec{k}) = \frac{C(\pm\omega(\vec{k}), \pm\vec{k})}{(2\pi)^2 \sqrt{4\pi\omega(\vec{k})}}.$$

The functions  $a_\pm(\vec{k})$  are called the momentum space amplitudes of the field  $\phi_0$ . The part of  $\phi_0(x)$  with  $a_+$  ( $a_-$ ) is called the positive (negative) frequency part of the Klein–Gordon field. If we require that all values of  $\phi(x)$  are real, we have to restrict the amplitudes  $a_\pm$  by the condition

$$a_+^*(\vec{k}) = a_-(\vec{k}), \quad (1.47)$$

where  $*$  denotes the complex conjugation.

Formula (1.46), regarded as a relation between the amplitudes and the field  $\phi_0$ , can be inverted. It is convenient first to introduce the operator  $\hat{P}_{\vec{k}}(y_0)$ ,

$$\hat{P}_{\vec{k}}(y^0)\phi(y^0, \vec{y}) = i \int d^3y \left( f_{\vec{k}}^*(y^0, \vec{y}) \frac{\partial \phi(y^0, \vec{y})}{\partial y^0} - \frac{\partial f_{\vec{k}}^*(y^0, \vec{y})}{\partial y^0} \phi(y^0, \vec{y}) \right), \quad (1.48)$$

where  $f_{\vec{k}}$  is a normalized plane wave

$$f_{\vec{k}}(y^0, \vec{y}) = \frac{e^{-iky}}{\sqrt{2(2\pi)^3\omega(\vec{k})}} \quad (1.49)$$

with  $k_0 = \omega(\vec{k})/c$ . Simple calculations show that

$$\hat{P}_{\vec{k}}(y^0)f_{\vec{k}'}(y^0, \vec{y}) = \delta(\vec{k} - \vec{k}'), \quad \hat{P}_{\vec{k}}(y^0)f_{\vec{k}'}^*(y^0, \vec{y}) = 0, \quad (1.50)$$

for any choice of  $y^0$ . It follows that

$$\hat{P}_{\vec{k}}(y^0)\phi_0(y^0, \vec{y}) = a_+(\vec{k}). \quad (1.51)$$

Note that there is no restriction on the choice of  $y^0$  present on the l.h.s. of this formula.

Formulas (1.51) and (1.47) inserted in formula (1.46) give the following identity

$$\phi_0(x) = \int d^3k \left( f_{\vec{k}}(x) \hat{P}_{\vec{k}}(y^0)\phi_0(y^0, \vec{y}) + \text{c.c.} \right). \quad (1.52)$$

Here c.c. stands for the complex conjugate of the preceding term. At this point it is convenient to define several new generalized functions:

$$\Delta^{(+)}(x) = -\frac{ic}{2(2\pi)^3} \int \frac{d^3k}{\omega(\vec{k})} e^{-ikx} \Big|_{k_0=\omega(\vec{k})/c}, \quad (1.53)$$

$$\Delta^{(-)}(x) = (\Delta^{(+)}(x))^*, \quad \Delta(x) = \Delta^{(+)}(x) + \Delta^{(-)}(x),$$

called the Pauli–Jordan functions. They obey the homogeneous Klein–Gordon equation. After simple manipulations, identity (1.52) can be rewritten in the following form

$$\phi_0(x) = - \int d^3y \left[ \Delta(x-y) \frac{\partial \phi_0(y)}{\partial y^0} + \frac{\partial \Delta(x-y)}{\partial x^0} \phi_0(y) \right]. \quad (1.54)$$

This very important formula gives an explicit solution to the homogeneous Klein–Gordon equation in terms of the initial data. We just take  $y^0 = ct_0$ , where  $t_0$  is the time at which  $\phi_0(y^0, \vec{y})$  and  $\partial \phi_0(y^0, \vec{y})/\partial y^0|_{y^0=ct_0}$  are explicitly specified as the



initial data. In particular, we see from formula (1.54) that vanishing initial data imply that  $\phi_0(x) = 0$ . This result was used at the end of the previous section.

The explicit formula for the Pauli–Jordan function  $\Delta(x)$  has the form (Appendix 2 in [3])

$$\Delta(x) = -\frac{1}{2\pi} \text{sign}(x^0) \left[ \delta(x^2) - \Theta(x^2) \frac{m}{2\sqrt{x^2}} J_1(m\sqrt{x^2}) \right],$$

where

$$\text{sign}(x^0) = +1 \text{ if } x^0 > 0, \quad \text{sign}(x^0) = -1 \text{ if } x^0 < 0.$$

One can see from this formula that the initial data are propagated in space with the velocity not greater than  $c$ . In particular, if the initial data taken at the time  $t_0$  vanish outside a certain bounded region  $V$  in space, then  $\phi_0(x)$  at later times  $t > t_0$  certainly vanishes at all points  $\vec{x}$  which cannot be reached by a light signal emitted from  $V$ . Another implication of formula (1.54) is the Huygens principle: the value of  $\phi_0$  at the point  $\vec{x}$  at the time  $t$  is a linear superposition of contributions from all points in space for which the initial data do not vanish (and which do not lie too far from  $\vec{x}$ ). This principle reflects the linearity of the Klein–Gordon equation.

## Exercises

1.1 (a) Check that the functions

$$\Phi_{+,v}(\xi, \tau) = 4 \arctan(\exp[\gamma(\xi - v\tau)]),$$

$$\Phi_{+,+}(\xi, \tau) = 4 \arctan\left(\frac{v \sinh(\gamma\xi)}{\cosh(v\gamma\tau)}\right), \quad \Phi_{+,-}(\xi, \tau) = 4 \arctan\left(\frac{\sinh(v\gamma\tau)}{v \cosh(\gamma\xi)}\right),$$

where  $\gamma = 1/\sqrt{1-v^2}$  and  $v$  is a real parameter such that  $0 \leq |v| < 1$ , are solutions of the sine-Gordon equation (1.7). Justify their interpretation:  $\Phi_{+,v}$  represents the soliton moving with constant velocity  $v$ ,  $\Phi_{+,+}$ —two solitons,  $\Phi_{+,-}$ —soliton + antisoliton pair.

(b) Comparing the asymptotic forms of solutions at  $\tau \rightarrow -\infty$  and  $\tau \rightarrow +\infty$  show that there is a repulsive force between the two solitons, and an attractive one in the case of the soliton + antisoliton pair.

(c) Check that the substitution  $v = iu$ ,  $u$ -real, in the  $\Phi_{+,-}$  solution gives a real-valued solution of the sine-Gordon equation which is periodic in time. Interpret this solution as a bound state of the soliton with the antisoliton (called the breather).

*Hints:* In the cases of  $\Phi_{+,+}$ ,  $\Phi_{+,-}$  consider the limits  $\tau \rightarrow \pm\infty$ . Use the identity

$$\arctan \frac{x-y}{1+xy} = \arctan x - \arctan y.$$

In order to show the presence of the forces, analyze shifts of the position of the soliton and the antisoliton with respect to the trajectory of the single (anti-)soliton.

**1.2** (a) The advanced Green's function  $G_A$  for the Klein–Gordon equation is obtained by choosing  $-i0_+$  in both terms in formula (1.41). Obtain formula analogous to (1.43) in this case.

(b) Prove also that  $G_F(x)$  defined as

$$G_F(x) = \frac{1}{(2\pi)^4} \int d^4k \frac{e^{-ikx}}{k^2 - m^2 + i0_+},$$

where  $k^2 = k_0^2 - \vec{k}^2$ , is another Green's function for the Klein–Gordon equation.  $G_F$  is related to the free propagator of the scalar field, and it plays an important role in the quantum theory of such fields. What is the choice of the signs  $\pm$  in formula (1.41) in this case?

**1.3** Using  $G_R$ , prove that

$$\vec{A}(t, \vec{x}) = \frac{1}{4\pi c} \int d^3y \frac{\vec{j}_T(\underline{t}, \vec{y})}{|\vec{x} - \vec{y}|},$$

where  $\underline{t} = t - |\vec{x} - \vec{y}|/c$ , is a solution of the wave equation (1.27).