Physics – Uspekhi **40** (3) 291 – 303 (1997)

METHODOLOGICAL NOTES PACS numbers: 03.65.Bz, **12.20.–m**, **42.50.–p**

Photons in a waveguide (some thought experiments)

L A Rivlin

Contents

Abstract. A set of thought experiments with guided waves (as the simplest example of spatially localized fields) shows that photons occupying a waveguide mode possess all the characteristics of nonzero inertial and gravitational rest mass. The corresponding quantity originates from the standing-wave component of the field and is merely an equivalent of the real energy from the 'raking up' of zero-point vacuum fluctuations from all unbounded space. It is impossible to distinguish this quantity from the standard concept of mass. This conclusion is valid for photons of any real spatially bounded fields. Two different classes of resonances with boson- and fermion-like features arise in waveguide ring structures depending on their field topology. The heuristic prospects for these observations are assessed.

"*... and God divided the light from the darkness...*"
(Genesis, 1, 4)

1. Introduction

The electromagnetic wave (light) exists only in motion. Nevertheless, people have known how to stop light for a long time without contradicting this basic fact. Here is a description of the formation of a *standing* light wave — one of the versions of *localizing* the fragment of electromagnetic field in space: "If you put a lamp by night between two smooth mirrors, separated by a distance of one cubit, you will see in each countless light reflections, each smaller than the next, and so to infinity, as if an infinite number of mirrors is inside each mirror". This description is attributed to Leonardo da Vinci and is contained in the notebook eventually named "The Atlantic Codex".

Incidentally, the epigraph points to a more ancient mention of the same problem. One of the principal actions of the first day of creation was the *spatial localization of the electromagnetic field*: the primordial light filling the world with an infinite number of unbounded waves was singled *out* from space deprived of field.

In spite of so venerable age, the question of features of the electromagnetic field localized in space in the form of the standing wave is still relevant today. An attentive observation of such *stopped light*, i.e. a field containing standing components, reveals features of photons that are customarily attributed to particles with nonzero inertial and gravitational rest mass. These features appear in a set of thought experiments, the results of which necessarily imply, for physically realizable fields, the concept of nonzero proper photon mass. Moreover, it proves impossible to suggest some experimental way to distinguish the quantity defined in this way from the standard concept of rest mass of usual massive bodies.

Another result is the formulation of the dependence of electromagnetic resonance properties in ring structures on the topology of space filled with a standing wave field, with a topology uniquely determining the resonance belonging to one of two classes, boson- or fermion-like.

An experiment — even a thought one — does not allow the use of an ideal unbounded plane wave: experimentally realizable waves are always spatially bounded. Correspond-

L A Rivlin Moscow State Institute of Radio Engineering, Electronics, and Automation (Technical University), prosp. Vernadskogo 78, 117454 Moscow, Russia
Tel./Fax (7-095) 434 93 17
E-mail: rla@superlum.msk.ru

Received 19 March 1996, revised 28 November 1996
Uspekhi Fizicheskikh Nauk **167** (3) 309 – 322 (1997)
Translated by K A Postnov; edited by A Radzig

ingly, photons represented by such waves are investigated here. This is the fundamental assumption of all further analysis. The field of modes in a hollow metallic waveguide and the photons filling them are considered as the simplest model for spatially bounded fields.

A metallic waveguide, an instrument now customarily used in experimental physics and applied radioengineering, was not at all such a trivial device a hundred years ago. In 1893, even such a penetrating physicist as Heaviside still denied the possibility of light propagation through tubes [1]. But only four years later Rayleigh published a study entitled "On the Passage of Electric Waves through Tubes or the Vibrations of Dielectric Cylinders" [2]. So we can celebrate the centenary of coming-to-be the notion of a waveguide as a real physical object.

In fact, despite its simplicity, this physical image of a two-dimensional potential well of infinite depth for photons turns out to be heuristically very productive. Since, as mentioned above, it serves only as a model, in what follows the term 'waveguide' should be thought of as such a potential well, putting aside the question as to how realizable are the corresponding ideally-metallic boundary conditions over different frequency ranges. In addition, to avoid confusion, it is worthwhile emphasizing that the model does not include the concept of a dielectric waveguide, i.e. of a bounded-depth potential well (this confusion may result from radically different phenomena in waveguides of both types sometimes being denoted by the same terms [3]).

A nontriviality of the ideal metallic waveguide model manifests itself in the well-known analogy between the Hamiltonian function for a free massive particle and the dispersion equation for the wave mode. Another instructive analogy appears when comparing the aforementioned properties of real bosons and fermions with those of electromagnetic resonances in ring waveguides of different topologies.

Whenever apparently chance analogies and coincidences occur in physical phenomena, it is worth thinking about what could be behind them. It is sufficient to recall what resulted, for example, from thinking about the 'noncausal' coincidence of the inertial and gravitational masses of a body. Nature with all its diversity is organized sufficiently economically for genuinely random coincidences capable of damaging our conviction of its deep uniformity to arise.

Today, it may be reasonably said that "the photon as an elementary particle of the optical field has no reasonable strict definition... Nevertheless it helps ... to qualitatively predict the results of new experimental situations. In general, the avoidance of the axiomatic approach sometimes helps in going forward" [4]. Putting aside the discussion of the physical meaning of the photon concept, it seems necessary to use this approach in this study.

2. The analogy between a photon in a waveguide and a free particle with nonzero mass

The analogy between the relativistic expression for the Hamiltonian function of a free particle

$$\mathcal{E}^2 = (mc^2)^2 + (cp)^2 \tag{2.1}$$

and the dispersion equation of a wave in a hollow metallic waveguide

$$\omega^2 = \omega_{nm}^2 + (c\zeta)^2 \tag{2.2}$$

has been noted by different authors for a long time. The following pairs of quantities are to be compared: the particle total energy $\mathcal{E}$ and the wave frequency ω, the particle proper mass m and the critical (cutoff) frequency ω_{nm} of the waveguide mode with integer indices n and m, particle momentum p and wave propagation constant ζ (c is the speed of light). The association becomes especially visualized if equation (2.2) is multiplied by the square of the Plank constant $\hbar$.

Referring in passing to this analogy, de Broglie notes: "All occurs as if the photon has its own mass determined by the form of the waveguide and the eigenvalue considered ...; in a given waveguide, a photon may possess a whole series of proper masses" [5]. Then, unfortunately, he continues: "Let us put aside these considerations bringing us away from the subject ..." Feynman, making the same comparison, restricts himself to remarking: "It's interesting, isn't it?" [6]. This analogy is already mentioned in textbooks, where one sometimes notes (see, for example, Ref. [7]) its informal character and gives a real physical meaning to the concept of mass of a photon in a waveguide.

Is it possible, however, to find really substantial differences in the behaviour of the quantity

$$M = \frac{\hbar\omega_{nm}}{c^2}\,, \tag{2.3}$$

which is proportional to the critical frequency of the waveguide mode ω_{nm}, and the particle proper mass $m \neq 0$ in the traditional sense of this concept? The answer to this question is contained in the following set of thought experiments, where the role of the quantity M (2.3) is analyzed in different physical situations. Not to anticipate it, it is useful to note that in the light of definition (2.3) the critical wavelength of a waveguide mode

$$\lambda_{nm} = \frac{2\pi c}{\omega_{nm}} = \frac{2\pi\hbar}{Mc}$$

takes the simple meaning of the Compton wavelength of a photon in a waveguide [8].

The standard theory of electromagnetic wave propagation through a waveguide is based essentially on the kinematic approach. To analyze the concept of the mass M of a photon in a waveguide in definition (2.3), one needs to solve dynamical problems in the thought experiments suggested, including dynamics in a gravitational field. But firstly one should verify that the physical reality in the form of the equivalent energy accumulated in some visual process and capable of being transformed into other types of energy corresponds to the quantity M (2.3).

3. Mode field evolution during a change of the waveguide cross-section [9]

As is well known [5], the existence of a finite critical mode with frequency $\omega_{nm} > 0$ is a basic feature distinguishing the propagation of a wave with $\omega > \omega_{nm}$ through a waveguide from the wave picture in free space. To find the essence of these differences, one needs to trace the evolution of the waveguide mode field as the waveguide cross-section changes.

From the first moment a wave with frequency $\omega_0 > \omega_{nm0}$ propagates through the waveguide with a critical frequency ω_{nm0}. If the cross-section changes slowly enough so that the field structure of the observable mode is conserved, the principal field behaviour is determined by the invariance of

the propagation constant $\zeta = \text{const}$ or, equivalently, by the phase interval invariance. The latter implies that the phase difference of the wave of a given mode over a fixed finite segment of a regular infinite-length waveguide remains unchanged with any variation of its cross-section geometry. Thus according to (2.2) it follows that

$$\omega^2 - \omega_{nm}^2 = \omega_0^2 - \omega_{nm0}^2 = \text{const}\,. \tag{3.1}$$

A special case of this relationship is an initial waveguide of infinite cross-section with infinitely remote walls, i.e. a *free space* with $\omega_{nm0} = 0$:

$$\omega^2 = \omega_0^2 + \omega_{nm}^2\,. \tag{3.2}$$

This situation can be considered as the starting point for the process of constructing a waveguide of finite cross-section by bringing closer its initially infinitely remote walls. Then the initial field with frequency ω_0 generates a wave inside the waveguide with frequency ω according to (3.2).

Special consideration should be given to the case when the initial field frequency is vanishingly small ($\omega_0 \to 0$), i.e. the initial field is purely *static*, and (3.2) by $\zeta = 0$ goes over into the equality

$$\omega = \omega_{nm}\,. \tag{3.3}$$

In this limiting case the initial *static* field in free space plays the role of a field generating a *wave* field with the critical frequency in the waveguide. If one considers free space as the limiting case of an infinite-cross-section waveguide, then the initial frequency $\omega_0 = 0$ naturally takes the meaning of the critical value, as this is the static field in free space that possesses the main signature of the critical frequency field: it is unable to propagate.

It stands to reason, in the opposite course of events from a critical waveguide of finite cross-section to the free space the degradation of the wave field into a purely static state is clearly seen.

Further, in addition to the phase interval invariance, one should admit that the photon occupation number in mode is constant ($\nu = \text{const}$) with a slowly changing waveguide cross-section, which seems to be quite justified in the absence of dissipation and nonlinearities. At the same time, the field energy $W' = \hbar\omega(\nu + 1/2)$ of the observed mode does not remain unchanged due to the ω-frequency variations (3.1). The 1/2 term reflects the contribution of the zero-point vacuum fluctuations. The total field energy in the waveguide with transverse structure and critical frequency peculiar to the given mode, equals

$$W = \hbar\omega(\nu + 1)\,, \tag{3.4}$$

where $1 = 1/2 + 1/2$ corresponds to the total energy of the zero-point fluctuations, which are unavoidably present in both modes of the same name with alternative propagation directions, even when the second possesses no real photons.

Therefore the waveguide evolution is accompanied not only by the wave frequency change according to equation (3.1), but also by the field energy W change in complete agreement with the adiabatic invariant of the oscillator energy-frequency ratio:

$$\frac{W}{\omega} = \frac{W_0}{\omega_0} = \hbar(\nu + 1)\,, \tag{3.5}$$

where W_0 is the starting energy at $\omega = \omega_0$.

Obviously, the field energy W increase is due to the work expended on moving the waveguide walls. In the present context, neither the character of external forces causing the motion, nor the nature of the energy source are significant; essential is the fact of energy storage in the waveguide mode.

In the special case of waveguide formation from free space with $\omega_{nm0} = 0$ containing the initial static field with $\omega_0 = 0$, a wave field of critical frequency emerges, whose energy of the quantum, according to (3.5), is

$$\hbar\omega_{nm} = \frac{W}{\nu + 1} \equiv w\,, \tag{3.6}$$

where w is the energy required to fill the mode by a single critical-frequency photon.

4. The mechanism of field evolution in a waveguide with moving walls [9]

The essence of field evolution in a waveguide with time-variated cross-section lies in the Doppler effect and relativistic change of the reflection angle θ by a moving mirror. These phenomena most clearly are revealed by a model of a simplest planar waveguide composed of two parallel metallic planes separated by a vacuum gap of width a and possessing a critical frequency $\omega_{n0} = \pi c n/a$ for the modes of two polarizations TE_{n0} ($n = 1, 2, \ldots, m = 0$) and TM_{n0} ($n = 0, 1, 2, \ldots, m = 0$). The field in such a waveguide is represented by the superposition of two partial plane waves incident on the walls at the angle θ, with $\cos\theta = \omega_{n0}/\omega$.

An accurate account of the frequency changes mentioned above

$$\omega' = \omega(1 + 2\beta\cos\theta + \beta^2)(1 - \beta^2)^{-1} \tag{4.1}$$

and of the angle

$$\cos\theta' = \cos\theta\left(1 + \frac{2\beta}{\cos\theta} + \beta^2\right)(1 + 2\beta\cos\theta + \beta^2)^{-1} \tag{4.2}$$

at each reflection of the partial wave from the wall slowly moving in a direction normal to the waveguide plane with velocity $c\beta \ll c$, leads to the following relationships:

$$a_0 \tan\theta = a \tan\theta_0\,, \qquad \omega \sin\theta = \omega_0 \sin\theta_0\,. \tag{4.3}$$

Zero indices denote the initial values.

The exclusion of the angle θ from system (4.3) brings us back to the invariant (3.1). This means that the phase interval invariance is provided essentially by relativistic mechanisms — the Doppler effect and the variation of the angle of reflection from a moving mirror, despite $\beta \ll 1$.

The increase of the field energy W is due to the work done against the radiation pressure on the moving walls:

$$\frac{\mathrm{d}W}{\mathrm{d}a} = -\frac{W}{a}\cos^2\theta = -\frac{W/a}{1 + (a/a_0)^2(\omega_0^2/\omega_{n00}^2 - 1)}\,. \tag{4.4}$$

The integration of (4.4) accounting for (4.3) leads to the adiabatic invariant (3.5). This solution is obtained directly from dynamical considerations of a waveguide with moving walls, not only without admitting the mode photon occupation number to be constant, but also without using the light quanta concept in general, i.e. entirely within the classical framework.

The same result relating to the origin of the field energy accumulated in the mode is also obtained for the general case of a waveguide with arbitrary cross-section changing in time [10]. The radiation pressure on the walls of the generalized cylinder is determined by the corresponding component of the Maxwellian tension tensor, and the cross-section deformation is produced while its contour shape remains homothetic. This leads to the equation

$$\frac{\mathrm{d}W}{\mathrm{d}\omega_{nm}^2} = \frac{W/2}{\omega_{nm}^2 - \omega_{nm0}^2 - \omega_0^2}\,, \tag{4.5}$$

which is equivalent to (4.4), and after its integration — to the adiabatic invariant (3.5). This general conclusion is clearly supported by detailed calculations for waveguides with circular and rectangular cross-sections [10].

5. Mass of photon in a waveguide — an equivalent to the energy of the critical frequency quantum. The seeding role of zero-point vacuum fluctuations [9]

Thus, the quantity w (3.6) coinciding with the energy $\hbar\omega_{nm}$ of quantum of the critical frequency, is equal to the energy accumulated in the waveguide due to the work done against the radiation pressure force of a single photon in a given mode during the formation of its wave field from the static field of the initial free space. No real photons are required to be present in this process of forming the critical frequency ω_{nm} field: even if $\nu = 0$, the energy of the ubiquitous zero-point vacuum fluctuations $\hbar\omega_0 = \hbar\omega_0/2 + \hbar\omega_0/2$, which possess a vanishingly small frequency $\omega_0 \to 0$, i.e. the static field fluctuations, is sufficient for seeding. Such a creation 'from nothing' of a critical frequency quantum reduces to the compression ('raking up') of the initial static fluctuation field from all unbounded space into the waveguide, increasing its frequency from zero to $\omega_{nm} > 0$.

The main feature of this compression process is the accumulation of energy in a selected waveguide mode. The presence of other internal or external forces and fields in addition to the mode field radiation pressure does not relate to this main result. The action of these fields can simply be included into the external force pushing the waveguide walls during the compression.

Turning to the seeding role of zero-point fluctuations during the storage of energy w (3.6) and noting their obvious analogy with the Casimir effect [11], one can easily discover substantial differences. The Casimir force is known [11, 12] to be due to the integral action on the metallic boundary on the part of fluctuation field of *all* modes in the *entire* unbounded frequency range. The resulting force has a different magnitude and even sign depending on the geometry of the metallic boundaries. For example, two plane-parallel metallic walls (as in a planar waveguide) are attracted to each other.

In contrast to this, to calculate the energy w (3.6), only that component of force which is due to vacuum fluctuations in a *single* selected waveguide mode is required and this component is directed from within to its moving walls.

Incidentally, in the geometry where the Casimir force is directed to the approaching walls, it can even be considered as one of the (or perhaps, the only) components of the external force pushing the walls. Then the energy w (3.6) accumulated in a selected mode comes from the redistribution of the energy of the whole unbounded frequency spectrum of the zero-point vacuum fluctuations and modes.

In the light of these results the quantity M (2.3) acquires a deep physical meaning: it is the mass equivalent to the energy w accumulated in a mode during the spatial localization of the field of zero-point vacuum fluctuations and is connected with the scale of the localization region. The equivalence of the quantity M to the real energy, which can be got back by the reverse motion of the waveguide walls, gives it a clear signature of mass.

A separate remark should be made about the mode TM_{00} with $n = 0$ of a planar waveguide; this mode possesses a zero critical frequency at any width a. The reason for this is a simple fact that serves to confirm the above concepts: the wave polarization of this mode is such that the electric vector is always perpendicular to the wall plane, and hence, the radiation pressure on it is lacking. As a result, both the energy w and the critical frequency vanish.

The conclusions made for a waveguide are also valid for a cavity as a limiting case of a waveguide: the photon mass in the resonator mode is simply proportional to the resonance frequency.

6. Longitudinal and transverse photon masses [3]

With the definition of the mass M in the form (2.3), the energy of photon in a waveguide can be obtained from the dispersion equation (2.2):

$$\hbar\omega = Mc^2\left[1 - \left(\frac{u}{c}\right)^2\right]^{-1/2}, \tag{6.1}$$

which is expressed through the standard formula for the energy transport velocity

$$u = c\left[1 - \left(\frac{\omega_{nm}}{\omega}\right)^2\right]^{1/2}. \tag{6.2}$$

The ratio of the latter to the speed of light (u/c) clearly has in (6.1) the same meaning as the usual relativistic ratio β for nonzero mass particles.

The constant of propagation ζ along the waveguide axis is also expressed through the transport velocity u (6.2) and mass M (2.3):

$$\hbar\zeta = Mu\left[1 - \left(\frac{u}{c}\right)^2\right]^{-1/2}. \tag{6.3}$$

Equations (6.1) and (6.3) are trivially coincident with the standard relativistic expressions for the energy and momentum of a massive particle.

Also coincident is the longitudinal mass of photon

$$M_{\parallel} = M\left[1 - \left(\frac{u}{c}\right)^2\right]^{-3/2}, \tag{6.4}$$

which is the coefficient of the longitudinal acceleration in the derivative $\hbar\zeta$ (6.3) with respect to time t:

$$\hbar\,\frac{\mathrm{d}\zeta}{\mathrm{d}t} = M\left[1 - \left(\frac{u}{c}\right)^2\right]^{-3/2}\frac{\mathrm{d}u}{\mathrm{d}t} \equiv M_{\parallel}\,\frac{\mathrm{d}u}{\mathrm{d}t}\,. \tag{6.5}$$

Apparently, the only force to cause such a longitudinal photon acceleration in a waveguide with constant cross-section is the gravitational force, directed along the waveguide axis (see Section 8 below).

The concept of transverse waveguide photon mass is not so simple. To introduce it, one needs to study accelerated field motion in the direction perpendicular to the waveguide axis. In this context, it has a meaning only if one considers the joint displacement of the field together with the waveguide containing it as a whole. The transverse photon mass is then defined as

$$M_\perp = \frac{\Delta F}{A}\,, \tag{6.6}$$

i.e. this is the ratio between the force difference ΔF relevant to the radiation pressure of a one-photon mode field exerted upon the opposite waveguide walls, and the waveguide acceleration A (neglecting, of course, the mass of the walls). The scheme of this thought experiment is somewhat reminiscent of the well-known consideration of light frequency shift in an accelerated elevator (see, for example, Ref. [13]).

The wave frequency changes with accelerated motion according to law (3.1), where the current critical frequency ω_{nm} is a function of the instant velocity $c\beta$ of the waveguide motion in the transverse direction due to relativistic contraction of the waveguide cross-section. The mechanism for this frequency change is the Doppler effect during reflections from the moving walls. For example, for the planar waveguide model $\omega_{nm} = \omega_{nm0}(1-\beta^2)^{-1/2}$ and the frequency varies as

$$\omega = \omega_0\sqrt{1+\left(\frac{\omega_{nm0}}{\omega_0}\right)^2\frac{\beta^2}{1-\beta^2}}\,. \tag{6.7}$$

To calculate the radiation-pressure force difference ΔF it is convenient to assume that the waveguide is moving with a constant acceleration

$$A = c\,\frac{\mathrm{d}\beta}{\mathrm{d}t}\,,$$

normal to the wall plane, with $\beta = 0$ at $t = 0$. At the next moment $t > 0$, when $\beta = At/c$, the photon with frequency ω belonging to the mode field falls on the wall at the angle θ and transfers to it the momentum

$$p_1 = \frac{2\hbar\omega}{c}\,\frac{\cos\theta \pm \beta}{1-\beta^2}\,. \tag{6.8}$$

The next reflection of the same photon from the opposite wall occurs at $t + \Delta t$, when the velocity increases by $\Delta\beta = A\Delta t/c$, transferring to the waveguide the momentum

$$p_2 = \frac{2\hbar\omega}{c}\left(\cos\theta \pm \beta \mp \Delta\beta\,\frac{1 \pm 2\beta\cos\theta + \beta^2}{1-\beta^2}\right) \times \frac{1}{1-(\beta+\Delta\beta)^2} \tag{6.9}$$

in the direction opposite to p_1. Here all frequencies and angles are given in the laboratory frame of reference fixed at $t = 0$ and $\beta = 0$; the upper signs relate to the case when the first of two sequential reflections occurs from the wall moving toward the incident photon, the lower ones — to the opposite case of the reflection from the co-moving wall.

As a result, in each cycle of two successive reflections the waveguide receives the differential momentum directed toward the acceleration vector:

$$\Delta p = \pm(p_1 - p_2) = \frac{2\hbar\omega}{c}\,\frac{1 \mp (\cos\theta \pm \beta)\Delta\beta/(1-\beta^2)}{1-\beta^2(1+\Delta\beta/\beta)^2}\,\Delta\beta \simeq \frac{2\hbar\omega}{c}\,\Delta\beta\,, \tag{6.10}$$

where the small terms quadratic in β and $\Delta\beta$ depending on the sequence of the reflections are omitted in the second equality. The velocity increase $\Delta\beta$ in a time between the two reflections is

$$\Delta\beta \simeq \frac{aA}{c^2\cos\theta}\,.$$

Hence the momentum given to the waveguide in one reflection is

$$\Delta p_0 = \frac{\Delta p}{2} = \frac{\hbar\omega}{c}\,\frac{aA}{c^2\cos\theta}\,. \tag{6.11}$$

Thus, the external force driving the waveguide, together with the field inside it, into accelerated motion is balanced by the force

$$I\Delta p_0 = \frac{\hbar\omega_{nm}}{c^2}(\nu+1)A\left[1-\left(\frac{u}{c}\right)^2\right]^{-1/2}, \tag{6.12}$$

where $I = (cW/a\hbar\omega)\cos\theta$ is the total flux of photons falling on the waveguide walls with a total energy W of the mode field. The comparison of the force $I\Delta p_0$ (6.12) related to a single photon in the mode, $I\Delta p_0/(\nu+1) = \Delta F$, with the definition (6.6) leads to the transverse photon mass in a waveguide

$$M_\perp = M\left[1-\left(\frac{u}{c}\right)^2\right]^{-1/2}, \tag{6.13}$$

coinciding with the relativistic formula for nonzero mass particles.

Equations (6.4) and (6.13) explain the difference between longitudinal and transverse masses: while both being proportional to the proper photon mass M (2.3), $M_\parallel$ and $M_\perp$ are manifestations of the photon inertia in different types of accelerated waveguide motion. The longitudinal accelerated photon motion constitutes simply wave propagation with changing frequency through the waveguide. The transverse photon motion consists in the acceleration of the wave together with the waveguide confining it, as a whole.

Returning to expression (6.2) for the transport velocity, one can note that in full agreement with the relativistic requirements for massive bodies, it is less than the speed of light: $u < c$, which has a clear interpretation (see, for example, Ref. [7]). The wave propagation through the waveguide consists in multiple reflections from walls of partial waves making the mode field. So the broken trajectory of the energy transfer does not coincide with the waveguide axis and the transport of energy along it slows down. Therefore, the reason for both the appearance of nonzero photon mass and the deceleration of the energy transport velocity is the same: it is impossible to represent the field by a single unbounded plane wave. It seems likely that the velocity intrinsic in energy transport (and hence matter transport) is always constant and equal to c (as well as the speed of light in different inertial systems), and only the motion along nonrectilinear trajectories caused by transverse spatial boundaries yields an

apparent velocity decrease, leading simultaneously to the emergence of nonzero mass.

7. A photon in a centrifuge [14]

Continuing to look for phenomena in which the quantity M (2.3) displays features of mass, it is worth using a simple recipe of Feynman [15]: "A qualitative measure of inertia is mass. It can be measured as follows: one simply ties a body with a cord, then rotates it with a certain velocity and measures the force that is necessary to keep it. The mass of any body can be measured in this way." In other words, one should find the coefficient M^* in the expression for the centripetal force

$$F_c = \frac{M^* u^2}{R}\left[1-\left(\frac{u}{c}\right)^2\right]^{-1/2}, \tag{7.1}$$

acting on the body that rotates with a linear velocity u around a circle of radius R. In the present context this means finding the force of reaction from the walls of a waveguide bent around a circular arc, on the wave propagating inside it, and comparing this force with F_c (7.1).

For simplicity, it is more convenient to consider a waveguide bent along an arc intercepted between the sides of the central angle 2π, i.e. a circular toroidal resonator which has a rectangular cross-section with sides a and b and circular cylindrical walls with radii R and $R_0 = R - a$, inside which the frequency of a wave running around the circle takes discrete admissible values ω_{nml}. The sought reaction force is numerically equal to the difference of the total modulus of the radiation pressure forces acting upon the external and internal cylindrical walls: $\Delta F = F(R) - F(R_0)$. The radial vector component of the radiation pressure on the cylindrical surface is equal to that of the Maxwellian tension tensor with opposite sign, while the time-average difference between the modulus of the total radiation pressure force is proportional to the field mode energy W:

$$\Delta F = \frac{W}{R}\left(\frac{c\varkappa}{\omega}\right)^2 \mathcal{F}, \tag{7.2}$$

where for modes of two polarizations

$$\mathcal{F}_{\mathrm{TM}} = \frac{R}{R_0} - \left(\frac{R}{R_0}-1\right)\left\{1-\left[\frac{J_m(\varkappa R)}{J_m(\varkappa R_0)}\right]^2\right\}^{-1},$$

$$\mathcal{F}_{\mathrm{TE}} = \frac{R}{R_0} - \left(\frac{R}{R_0}-1\right)\left\{1-\left[\frac{J'_m(\varkappa R)}{J'_m(\varkappa R_0)}\right]^2 \frac{1-(m/\varkappa R_0)^2}{1-(m/\varkappa R)^2}\right\}^{-1}, \tag{7.3}$$

and the eigenvalues $\varkappa$ are determined by the roots of the equations

$$N_m(\varkappa R)J_m(\varkappa R_0) = J_m(\varkappa R)N_m(\varkappa R_0) \quad (\text{for TM}),$$

$$N'_m(\varkappa R)J'_m(\varkappa R_0) = J'_m(\varkappa R)N'_m(\varkappa R_0) \quad (\text{for TE}) \tag{7.4}$$

and J_m, N_m are integer m-order Bessel and Neumann functions, J'_m, N'_m their derivatives with respect to the arguments. The admissible frequencies of waves of both polarizations follow from the relationship

$$\omega_{nml}^2 = (c\varkappa_{nm})^2 + \left(\frac{\pi c l}{b}\right)^2 = \omega_{n0l}^2 + c^2(\varkappa_{nm}^2 - \varkappa_{n0}^2), \tag{7.5}$$

where the indices n, m, l are integers, and ω_{n0l} makes sense of the critical frequency of a wave running around the circle, since at $m = 0$ its propagation in the azimuthal direction stops.

To compare ΔF (7.2) with the centripetal force F_c (7.1), one should take into account that in Feynman's mechanical model described by (7.1) all the mass M^* is concentrated around a circle of radius R, whereas the field energy in the toroidal waveguide is distributed over a ring of thickness a [just this fact is accounted for by the factor $\mathcal{F}$ (7.3)]. Therefore, the comparison proposed is valid only for such modes of the torus in which the field closely adjoins the outer cylinder of radius R. This requirement is fulfilled by modes with multiple field periodicity around the circle ($m \gg 1$) and a small number of field nodes along the radius ($n \sim 1$), when all the field components rapidly decrease with reducing radius, and without appreciably affecting the result, one can accept $R_0 \ll R$ and for the first roots of (7.4) we obtain approximately $\varkappa R \simeq m$ [16].

Then the difference of the moduli of the total radiation pressure forces (7.2), which is proportional to the square of the transport velocity $u = c^2 m/\omega R$, is

$$\Delta F = \frac{W}{R}\left(\frac{cm}{\omega R}\right)^2 = \frac{W}{R}\left(\frac{u}{c}\right)^2, \tag{7.6}$$

and the wave frequency, according to (7.5), is

$$\omega = \omega_{n0l}\left[1-\left(\frac{c}{\omega}\right)^2(\varkappa_{1m}^2 - \varkappa_{10}^2)\right]^{-1/2} \simeq \omega_{n0l}\left[1-\left(\frac{u}{c}\right)^2\right]^{-1/2}. \tag{7.7}$$

Finally, taking into account (3.4), we arrive at

$$\Delta F = \frac{\hbar\omega_{n0l}}{c^2}\,\frac{u^2(v+1)}{R}\left[1-\left(\frac{u}{c}\right)^2\right]^{-1/2}, \tag{7.8}$$

where ω_{n0l} is the critical frequency of the toroidal waveguide, and the comparison $\Delta F = F_c$ leads to

$$M^* = \frac{\hbar\omega_{n0l}}{c^2}(v+1) = M(v+1). \tag{7.9}$$

Thus, the experimental determination of the inertial mass of a body by Feynman's method suggests that the quantity M (2.3) is the mass of photon in a waveguide mode.

8. A heavy photon in a waveguide. Gravitational resonator [8, 17]

The answer to the question as to whether the quantity M (2.3) manifests features of both inertial and gravitational mass, is given by studying wave propagation through a waveguide fixed vertically in a gravitational field.

By definition (2.3), the photon mass is proportional to the critical frequency, which in turn depends upon the linear size a of the cross-section:

$$M = \frac{\hbar\omega_{nm}}{c^2} = \frac{\hbar\varkappa}{c} = \frac{\hbar\mu_\varkappa}{ca}, \tag{8.1}$$

where the root of the characteristic equation $\mu_\varkappa$ is a mathematical invariant.

The cross-section a, as any linear dimensions of the body, depends on the gravitational potential Φ_G as $a = a_\infty(1+\Phi)$, where the modulus of the normalized gravitational potential

$|\Phi| \equiv |\Phi_G|/c^2 \ll 1$, and the sign ∞ here and below marks the values at infinity from the gravitational field sources, where the potential $\Phi = 0$, so $\Phi \leqslant 0$. The speed of light also changes: $c = c_\infty(1 + 2\Phi)$ [13].

Accounting for these dependences, expression (8.1) yields the waveguide mass of photon in a gravitational field

$$M = \frac{\hbar\mu_\varkappa}{c_\infty a_\infty}\frac{1}{(1+\Phi)(1+2\Phi)} = M_\infty(1 - 3\Phi)\,, \tag{8.2}$$

where $M_\infty = \hbar\mu_\varkappa/c_\infty a_\infty$. This dependence, obtained from definition (2.3) as a result of the gravitational field effect on the factors constituting M, coincides with the relativistic mass change of a usual heavy body in a gravitational field [13], also evident from the dimensional analysis.

In the gravitational field the propagation constant ζ also changes due to both the critical frequency $\omega_{nm} = \omega_{nm\infty}(1 + \Phi)$ and the wave frequency $\omega = \omega_\infty(1 - \Phi)$ dependences on the gravitational potential Φ [13]:

$$\zeta^2 = \left(\frac{\omega}{c}\right)^2 - \left(\frac{\omega_{nm}}{c}\right)^2 \simeq \zeta_\infty^2\left\{1 - 2\Phi\left[1 + 2\left(1 - \frac{\omega_{nm\infty}^2}{\omega_\infty^2}\right)^{-1}\right]\right\}. \tag{8.3}$$

Here the second term can be not small even for $|\Phi| \ll 1$, if $\omega_\infty \to \omega_{nm\infty}$, i.e. near the critical conditions, however ζ does not vanish because $|\zeta_\infty| > 0$ at $\Phi < 0$.

If one rejects the limitation $|\zeta_\infty| > 0$ and traces the propagation constant ζ changing along the waveguide from the point with potential Φ to the point with potential $\Phi + \Delta\Phi$, then

$$\zeta^2 = \zeta_0^2\left[1 - 2\Delta\Phi\left(1 + \frac{2}{1 - \omega_{nm}^2/\omega^2}\right)\right], \tag{8.4}$$

where ζ_0, ω_{nm} and ω are the values at the point with potential Φ. The propagation constant decreases as the wave travels towards the point with higher gravitational potential ($\Delta\Phi > 0$).

Most interesting is the case close to the critical conditions, when $1 - (\omega_{nm}/\omega)^2 \ll 1$ and the initial value ζ_0 is small (in terms of the mechanical analogy: when the kinetic energy of the body is small in comparison to the proper potential energy, i.e. within the nonrelativistic approximation). Then, despite the condition $|\Delta\Phi| \ll 1$, for

$$\Delta\Phi \simeq \frac{1}{4}\left[1 - \left(\frac{\omega_{nm}}{\omega}\right)^2\right] \tag{8.5}$$

a situation emerges which reproduces, in a certain sense, a tabletop laboratory 'black hole': the r.h.s. of (8.4) vanishes, $\zeta = 0$, and the wave propagation upwards (i.e. against the gravity force) through the waveguide stops. Then, together with the sign of the square root $\pm\sqrt{\zeta^2}$, the direction of propagation reverses. Such a gravitational reflection of a wave in a homogeneous regular waveguide is similar to the pause at the uppermost point of a trajectory and subsequent fall of a heavy body that has been thrown upward and has exhausted its initial momentum.

Condition (8.5) defines the gravitational height of the waveguide, i.e. the vertical distance H to the wave's turning point. In particular, if $\Delta\Phi_G = c^2\Delta\Phi = gz$, where g is the acceleration of free fall and z the vertical coordinate, then the waveguide gravitational height is

$$H = \frac{c^2}{4g}\left[1 - \left(\frac{\omega_{nm}}{\omega}\right)^2\right] = \frac{u^2}{4g}\,, \tag{8.6}$$

where values in the r.h.s. are taken at $z = 0$.

Thus, the cross-section of the waveguide with vertical coordinate $z = H$ serves as an opaque gravitational mirror for the wave. [It should be noted that an attempt to observe the gravitational mirror effect under terrestrial gravity would require very stable experimental conditions (frequency, geometry, etc., $\sim 10^{-15}$) and ideal conductivity (superconductivity) of the metallic waveguide walls.]

The gravitational height H (8.6) obtained is two times smaller than the height of elevation of a heavy body with initial velocity u in a gravitational field with constant acceleration of free fall g, calculated in the framework of Newtonian mechanics. The reason for this discrepancy is the same as when comparing the results of the calculation of gravitational light trajectory bending in general relativistic theory, confirmed experimentally, with those obtained by Newtonian mechanics using the equivalence principle (see, for example, Ref. [13]). The gravitational reflection of a wave in a vertical waveguide is in essence one version of gravitational light bending. This is seen, for example, in a simple model of a planar vertical waveguide (Fig. 1). Using the well-known technique of scanning a light beam, reflected many times from flat mirrors, one can imagine the zig-zag broken line with constant step that the light beam follows in the waveguide at $g = 0$ as a line inclined by the angle θ to the normal to the waveguide walls. A gravitational field with $g > 0$ makes this initially straight line bend as shown in Fig. 1 by the thick line, with the maximum corresponding to the gravitational height H (8.6). This curve represents a smoothed scanning of the light beam path in the waveguide shown by the thick broken line with decreasing step, which illustrates the phenomenon of gravitational wave stopping and reflection. In the same figure, the dashed curve shows the scanning of beam deflected with half the angle variation rate, calculated in the Newtonian approximation. The maximum of this curve, which is essentially a parabolic nonrelativistic trajectory of a heavy body, equals $2H$.

From the same figure one can note, incidentally, the distributed character of the gravitational mirror: the wave vector of a partial wave gradually reverses over all the waveguide height. Therefore, perhaps, it is more relevant to speak not about the reflection, but about the refractive

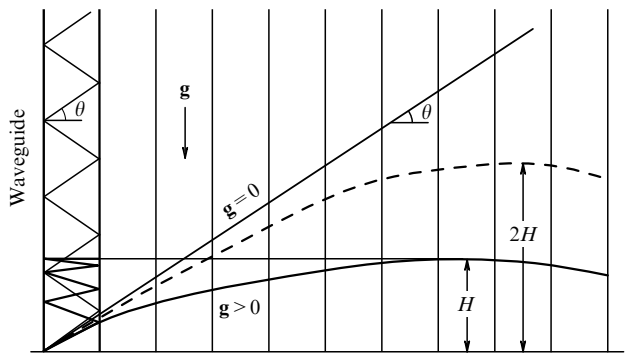

Figure 1.

inversion of a wave in a waveguide exposed to a gravitational field.

If there is a mirror — a gravitational one in the present case, then it is easy to construct a resonator in the form of a waveguide with vertical dimension exceeding the gravitational height H, the bottom end closed by a common metallic mirror. The resonance condition

$$\int_0^H \zeta(z)\,\mathrm{d}z = \pi N, \qquad N = 1, 2, \ldots$$

in the particular case (8.6), when $\zeta = \zeta_0(1 - z/H)^{1/2}$, yields the resonance frequency measured at the bottom mirror:

$$\omega_N = \sqrt{\omega_{nm}^2 + \left(\frac{3\pi c N}{2H}\right)^2} = \omega_{nm}\sqrt{1 + \left(\frac{6\pi g N}{c\omega_{nm}}\right)^{2/3}}\,. \quad (8.7)$$

9. Weighing the photon [17]

To weigh a photon directly, one should put a resonator with one photon in the mode on a balance platform and determine the difference of forces of radiation pressure on the lower and upper reflecting surfaces, which will yield the photon weight. This procedure is especially visual in a gravitational resonator, in which the radiation pressure on the lower and upper surfaces is transmitted to different bodies: the pressure on the lower metallic surface is transferred to the balance platform, and the upper gravitational mirror transmits the reaction force, bypassing the balance, directly to the body of high mass (source of gravity), for example, the Earth. Thus, the radiation pressure on the metallic bottom of a resonator, not being balanced by the opposite force, is defined as the photon weight in the gravitational resonator.

At each reflection from the bottom, the single photon of the mode transfers a momentum to the balance platform

$$\Delta p = \frac{2\hbar\omega_N}{c}\sqrt{1 - \left(\frac{\omega_{nm}}{\omega_N}\right)^2}\,,$$

and the whole period of the photon circulation in the waveguide resonator (up and down) is equal to twice the time of the electromagnetic energy transfer to a height H with the velocity u:

$$\Delta t_H = 2\int_0^H \frac{\omega}{c^2\zeta}\,\mathrm{d}z\,,$$

where the integrand is given by (8.4). For the case $\Delta\Phi_{\mathrm{G}} = gz$ this yields

$$\Delta t_H = \frac{2}{u}\int_0^H \left(1 - \frac{z}{H}\right)^{-1/2}\mathrm{d}z = \frac{4H}{u} = \frac{u}{g}\,.$$

Hence the photon weight as the Δp to Δt_H ratio equals

$$F_g = \frac{\Delta p}{\Delta t_H} = 2\,\frac{\hbar\omega_N}{c^2}\,g = 2Mg\,, \quad (9.1)$$

where $M = \hbar\omega_N/c^2$ is the mass (2.3) of a photon with frequency ω_N (8.7). It should be noted that the photon weight is two times higher than the standard classical value, and the period of circulation Δt_H is two times smaller than that for a heavy ball that reflects elastically from an immobile base. The reason for this is the relativistic character of the photon motion.

Putting aside the photon for a while, how can one weigh a heavy particle of mass m moving with a relativistic velocity $\mathbf{u}$, not lying on the balance platform? Let the particle trajectory in the gravitational field be a continuous sequence of arc-like segments based on the horizontal balance platform, and the reflections at the points of fall be absolutely elastic. Then the particle weight can be determined as the ratio of the twofold normal component of the particle momentum $2p_z$ near the platform surface to the time interval Δt between two successive reflections, i.e. $F_g = 2p_z/\Delta t$. The equation of motion of the particle in a field with gravitational potential Φ_{G} determining the particle trajectory is (see, for example, Ref. [13])

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = m\left\{-\left[1 + \left(\frac{u}{c}\right)^2\right]\nabla\Phi_{\mathrm{G}} + \frac{\mathbf{u}}{c}\left(\frac{\mathbf{u}}{c}\cdot\nabla\Phi_{\mathrm{G}}\right)\right\}. \quad (9.2)$$

If $\nabla\Phi_{\mathrm{G}} = \mathbf{g}$, then in Cartesian coordinate system this equation splits into two ones:

$$\frac{\mathrm{d}p_z}{\mathrm{d}t} = -mg\left[1 + \left(\frac{u_x}{c}\right)^2\right],$$

$$\frac{\mathrm{d}p_x}{\mathrm{d}t} = mg\,\frac{u_x u_z}{c^2}\,, \quad (9.3)$$

where indices x and z denote the corresponding horizontal and vertical vector components.

Two limiting cases are possible. For a strictly vertical motion $u_x = 0$ and the particle weight $F_g = mg$ coincides with the Newtonian definition. For a very slanting, nearly horizontal trajectory, one can assume $u_z \ll c$, $u_x \simeq u$ and then

$$F_g = mg\left[1 + \left(\frac{u}{c}\right)^2\right] \leqslant 2mg\,.$$

The general case of a moderately slanting trajectory corresponds to a weight value between the two limiting ones:

$$mg \leqslant F_g \leqslant mg\left[1 + \left(\frac{u}{c}\right)^2\right] \leqslant 2mg\,. \quad (9.4)$$

In a similar way, the photon weight in a vertical waveguide, depending on how close the critical regime is, lies within the range from Mg at $\omega \gg \omega_{nm}$ to $2Mg$ at $\omega \to \omega_{nm}$. The last case corresponds to the phenomenon of gravitational light deflection in general relativistic theory.

10. The proper photon mass and the momentum defect [3, 8]

The above-considered thought experiments with an electromagnetic wave in a waveguide lead to the conclusion that de Broglie's analogy has not at all a formal but a clear physical meaning. The energy of a quantum of the critical frequency, which corresponds in this analogy to the proper photon mass in the waveguide, is real. It is accumulated in the mode by field compression during the evolution from an unbounded plane wave of free space to the transverse-localized mode field of a finite cross-section waveguide. This evolution is accompanied by doing the work against the radiation pressure forces of the

zero-point vacuum fluctuations, originally of zero frequency. The proper photon mass in the waveguide, defined as the equivalent of the energy mentioned above, serves in different situations precisely as the inertial and gravitational mass in the standard meaning of these concepts. Therefore, the quantity M (2.3) must necessarily be identified with the nonzero proper photon mass in a waveguide, since no thought experiment has so far been suggested to disprove this statement.

Going back to the basic relativistic relationship (2.1), it is necessary to stress once again in the spirit of a very instructive analysis of the mass concept [18, 19] that the proper photon mass is, according to (2.1), identically zero only in so far as the photon momentum modulus $|\mathbf{p}| = \hbar\omega/c$ is equal to the energy $\hbar\omega$ with a factor of c^{-1}. However, the equivalence of these quantities corresponds only to a photon represented by an ideal, infinite and unique plane wave, which cannot be physically realized.

In fields with complex structures, which differ from the ideal image above, separate components of the vector quantity $\mathbf{p}$ can be mutually balanced to yield zero in the sum, so that $|\mathbf{p}| < \hbar\omega/c$.

Such a mutual balancing occurs in a waveguide mode comprising ν photons with total energy $\nu\hbar\omega$. The partial waves with the transverse momentum components

$$p_\perp = \pm\frac{\hbar\omega_{nm}}{c}\,\frac{\nu}{2}$$

form a standing wave in the direction normal to the waveguide axis with total momentum

$$\frac{\nu}{2}\left[\frac{\hbar\omega_{nm}}{c} + \left(-\frac{\hbar\omega_{nm}}{c}\right)\right] = 0\,.$$

Comparing this to the case of an unbounded plane wave, in the energy-momentum balance (2.1) they produce a *momentum defect*

$$\nu\delta p = \frac{\nu}{2}\left[\frac{\hbar\omega_{nm}}{c} - \left(-\frac{\hbar\omega_{nm}}{c}\right)\right] = \nu\,\frac{\hbar\omega_{nm}}{c}\,.$$

As a result, the photon mass in a waveguide with $\nu = 1$ equals the momentum defect divided by the speed of light:

$$M = \frac{\delta p}{c}\,, \tag{10.1}$$

and the electromagnetic field in the waveguide consists of a propagating travelling wave and an oscillator *resting* in space — a standing wave.

Clearly, the standing field components are inherent not only in the waveguide or resonator modes stemming from metallic boundary conditions. They are considered as examples only because of their simple boundary structure. One can give many examples of standing-wave components in fields, which are bounded in space only partially and not reduced to the ideal image of a unique plane wave and, hence, they possess the momentum defect and nonzero proper mass [8]. Among these examples are: a plane wave reflected obliquely from a flat mirror; the diffraction field behind a screen hole; a single-mode laser beam, etc.

Although, of course, any standing wave can be decomposed into the superposition of two oppositely travelling waves, it is, in essence, not a wave in a strict sense, which assumes a process accompanied by energy transportation, because the space occupied by the standing wave is split into a sequence of isolated domains with no energy exchange between them. An image of a standing wave is a series of oscillators resting in space.

Each domain of a standing wave contains an *immobile* energetic fragment — the *rest energy*. From here one makes one step to obtain its equivalent — the nonzero proper mass generated by such a stopped energy.

Thus, the necessary and interconnected attributes of nonzero photon proper mass M (2.3) are: the spatial localization of the wave field excluding the possibility of its representation by a single unbounded plane wave; the presence of the standing field component ('stopped' light) and the momentum defect; the decrease of the energy transport velocity to a sublight level and the appearance of dispersion.

All mentioned above is, of course, one of the manifestations of the coordinate-momentum uncertainty relation as the real reason for the appearance of nonzero proper mass. In the light of the concept considered, the energy of discrete bound states, which appear during the spatial localization of matter, is due to the work done in raking up (compressing) it from the initial unbounded space [3, 8, 9, 20].

The vital role in this process is played not so much by the spatial restriction as by the standing wave emergence accompanying it: the energy of the lowest state in a potential well of width a is equal to the energy of the nth state in a well of width na although the degree of localization in the first case is n times higher than in the second one, but the standing-wave domain size is the same in both cases. In other words, the energy of the bound states is determined not by the size of the localization region, but the scale of the standing-wave domain.

It is worth noting that the same considerations about the physical nonreality of the ideal image of a plane wave lead to the inversion of the continuum of canonical field variables into a denumerable set of radiation oscillators (the free space modes) [21].

11. From the wave equation to the Schrödinger equation [3]

Thus, the nonzero photon proper mass emerges as a result of the wave's spatial bounding, which leads to the appearance of a standing field component. One can expect that the transformation of the massless wave equation should simultaneously evolve into a Schrödinger-type equation containing mass explicitly without introducing it as *a priori* definition.

To follow this conversion, the wave dispersion equation in a waveguide (2.2) should be written, in view of the nonrelativistic character of the Schrödinger equation, in the approximation of proximity to the critical regime:

$$\omega \approx \omega_{nm} + \frac{(c\zeta)^2}{2\omega_{nm}}\,, \tag{11.1}$$

which corresponds in the mechanical analogy to the kinetic energy being small with respect to the proper energy. Simultaneously, the application of differential operators to a monochromatic wave running along the z-axis allows one to write them down in the form $\partial/\partial t = \mathrm{i}\omega$ and $\partial^2/\partial z^2 = -\zeta^2$.

Therefore, the dispersion equation (11.1) can be interpreted as operator equations

$$\frac{\partial^2}{\partial z^2} + 2\,\frac{\omega_{nm}}{c^2}\,(\omega - \omega_{nm}) = 0\,, \tag{11.2}$$

$$\frac{1}{\mathrm{i}}\,\frac{\partial}{\partial t} = -\frac{c^2}{2\omega_{nm}}\,\frac{\partial^2}{\partial z^2} + \omega_{nm}\,. \tag{11.3}$$

Applying these equations to any of the six field components χ_i accounting for $\hbar\omega = \mathcal{E}$ and $\omega_{nm} = Mc^2/\hbar$, we arrive at

$$\frac{\partial^2 \chi_i}{\partial z^2} + 2\,\frac{M}{\hbar^2}\,(\mathcal{E} - \hbar\omega_{nm})\chi_i = 0\,, \tag{11.4}$$

$$\frac{\hbar}{\mathrm{i}}\,\frac{\partial \chi_i}{\partial t} = -\frac{\hbar^2}{2M}\,\frac{\partial^2 \chi_i}{\partial z^2} + \hbar\omega_{nm}\chi_i\,. \tag{11.5}$$

As stated above, the energy $\hbar\omega_{nm}$ is basically the potential energy U accumulated in the waveguide, which can be returned by its reverse evolution to a free space. Apart from this potential energy of field compression in the waveguide, one can find only one possible additional wave potential energy, namely, the gravitational potential energy. As shown above, the energy of quantum of the critical frequency in the gravitational field with potential difference $\Delta\Phi_{\mathrm{G}}$ is

$$\hbar\omega_{nm} = \mathrm{const} + \hbar\omega_{nm0}\,\frac{\Delta\Phi_{\mathrm{G}}}{c^2}\,,$$

i.e., assuming the normalization const $= 0$, the quantity $\hbar\omega_{nm}$ in (11.4) and (11.5) can be interpreted as the potential energy $\hbar\omega_{nm} = U$. Then these equations take the familiar form

$$\frac{\partial^2 \chi_i}{\partial z^2} + 2\,\frac{M}{\hbar^2}\,(\mathcal{E} - U)\chi_i = 0\,, \tag{11.6}$$

$$\frac{\hbar}{\mathrm{i}}\,\frac{\partial \chi_i}{\partial t} = -\frac{\hbar^2}{2M}\,\frac{\partial \chi_i}{\partial z^2} + U\chi_i\,. \tag{11.7}$$

The absence of the transverse Laplacian in these equations coinciding with the one-dimensional Schrödinger equation, is not surprising: the integration over transverse coordinates under the waveguide boundary conditions has already been done, and the photon proper mass (2.3) simply results from this integration.

Despite the discovered genetic connection of the Schrödinger equation to the classical wave equation, their solutions are traditionally thought to have qualitatively different meanings. The difference is that the first equation has a complex-function solution, for example, like $\chi_i \sim \exp\bigl[\mathrm{i}(\omega t - \zeta z)\bigr]$, whereas in the solution of the second equation only purely real $\mathrm{Re}\,\chi_i$ or imaginary $\mathrm{Im}\,\chi_i$ parts are considered as having physical meaning.

This difference disappears and the complex solutions of the classical wave equations become meaningful when considering the unique entity of the electromagnetic field represented by interconnected vectors $\mathbf{E}$ and $\mathbf{H}$ and the electromagnetic field tensor construction containing both real and imaginary components. This suggests the introduction of a complex radiation field vector

$$\mathbf{R} = \frac{\varepsilon_0\varepsilon}{2}\,\mathbf{E} + \mathrm{i}\,\frac{\mu_0\mu}{2}\,\mathbf{H}\,, \tag{11.8}$$

where the electric component is described by the real and the magnetic by the imaginary parts of the six-dimensional vector (ε and μ are the relative dielectric permittivity and magnetic permeability, respectively). The normalization in (11.8) is chosen so that the field energy density is

$$w_0 = \mathbf{R}\cdot\mathbf{R}^*\,, \tag{11.9}$$

i.e. equal to the scalar product of the vector $\mathbf{R}$ by its complex conjugate $\mathbf{R}^*$, and the Poynting vector is equal to their vector product

$$\mathbf{S} = \frac{\mathrm{i}c}{\sqrt{\varepsilon\mu}}\,\mathbf{R}\times\mathbf{R}^*\,. \tag{11.10}$$

It is worth noting that in expression (11.9) for the field energy density, vectors $\mathbf{R}$ and $\mathbf{R}^*$ play the same role as functions ψ and ψ^* in the quantum-mechanical definition of the probability density. An instructive illustration of the similarity between behaviour of the photon with nonzero proper mass and a massive quantum-mechanical particle is the detailed analogy between electromagnetic wave reflection from a boundary separating two dielectrics in a waveguide and particle reflection from a potential barrier. This analogy explains the sense of periodic and aperiodic solutions emerged [3].

12. A waveguide with energy dissipation

The threshold character of the critical phenomena in a waveguide expressed by the dependence of the propagation constant ζ on the frequency ω (2.2) is discovered only in a model with infinite-conductivity walls σ. In a real waveguide with $\sigma^{-1} > 0$ and, correspondingly, with wave energy dissipation, the lower σ, the more the dependence $\zeta(\omega)$ loses its threshold shape and the critical frequency concept loses its strict determinacy.

Thus, the concept of proper mass M (2.3) acquires a precise meaning together with ω_{nm} only asymptotically at $\sigma \to \infty$, essentially as a result of ignoring the dissipation processes. Neglecting the dissipative processes by no means affects the reliability of the waveguide electrodynamics conclusions: even Newton's first law is the result of neglecting friction during the motion of material objects.

The dissipative processes pose the question about the infinite electromagnetic length definition for a geometrically limited waveguide [22]. To formulate such a definition without invoking the tautology 'infinite — having infinite length', one needs to rely upon the basic functional waveguide property — the ability of electromagnetic energy transfer: a waveguide of unknown length is assumed to be infinite by definition if there is no response to a finite probing input signal. More precisely, if the reflected signal expected turns out to be indistinguishable against the noisy background, which is present at the input in the absence of the probing signal, too.

The thermodynamically equilibrium noise radiation of a waveguide differs from black body radiation by having a lower frequency limit. The criterion of the response being indistinguishable determines the geometrical extension of the waveguide defined as thermodynamically infinite in the form of the following length:

$$z_T = \frac{2}{AG}\,\ln\left\{N_0\left[\frac{2\pi u}{c}\,\mathcal{F}_{\mathrm{BE}}(1 + \mathcal{F}_{\mathrm{BE}})\right]^{-1/2}\right\}, \tag{12.1}$$

where $A(\omega, T)$ is the absorption ability of the metallic walls used in Kirchhoff's law, N_0 is the photon number in the probing signal, $\mathcal{F}_{\mathrm{BE}}(\omega, T)$ is the Bose–Einstein distribution for the absolute temperature T, and G is a geometrical factor depending upon the wave polarization [22]. The concept of an infinite geometrically limited waveguide does not lose its meaning by the thermodynamic fluctuation level's tending to zero: for very low temperatures and/or very high frequencies, quantum fluctuations prevail.

The fluctuation-noise essence is apparently characteristic for the notion of infinity in more general sense, and sometimes extinguishes its content accessible to testing in thought experiments. What looks more infinite than the 'foolish infinity' of positive integers? One can try to define it using some thought experiment. Let positive integers be realized as records in a sequence of cells. Then one determines experimentally the probability of discovering in an arbitrary cell a number exceeding any given one, i.e. arbitrarily large. If this probability is not zero, then the series of numbers under study should be considered as infinite, i.e. containing arbitrarily large numbers or consisting of an arbitrarily large number of cells enumerated with these numbers. However, for a number-limited series with cell records affected by noise, there is also a nonzero probability of discovering by this procedure of arbitrary sampling any number, including an arbitrarily large one. If one finds a thermodynamically infinite length of such a series using some procedure similar to that for the waveguide above, then obviously, the weaker the signal corresponding to the record and the higher the background noise level, the smaller the number of physical terms that this length contains. Similar considerations are valid for other objects and phenomena pretending to be infinite [22].

13. Boson- and fermion-like resonances in waveguide ring-structures [23]

Properties of wave fields spatially bounded in transverse coordinates are radically dependent upon the topology of the surrounding space. The waveguide ring-structures clearly illustrate this.

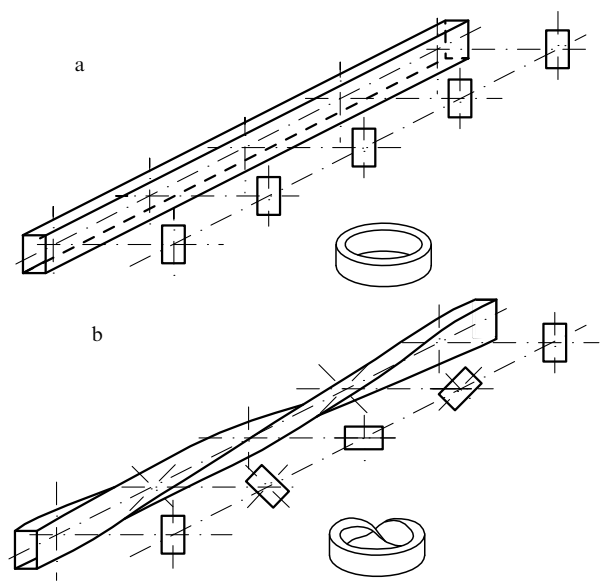

Figure 2.

A waveguide looped in a ring such that its input and output cross-sections coincide, forms a resonator (the ring of this type is shown in Fig. 2a together with its linear evolvent and cross-sections at different points along its axis). It is assumed that the ring is formed smoothly enough for the general field structure to be infinitesimally disturbed. The resonance condition here is in equating the optical ring length to an integral number of wavelengths:

$$\zeta_{nm} L = 2\pi N, \qquad N = 1, 2, \ldots \tag{13.1}$$

(L is the geometrical ring-perimeter length).

It turns out, however, that this resonance condition is valid only for a space with trivial topology and the integer-wavelength rule is only one possibility. To transform the waveguide ring topology, it is sufficient, before joining its initial and end cross-sections, to twist the waveguide along its axis by the angle $\Delta\theta = \pm\pi$ (see Fig. 2b in which the ring evolvent and its cross-section sequence are also presented). This operation is fully equivalent to the formation of a unilateral nonoriented surface with zero Euler characteristics like a Möbius band. The consequences of the twisting procedure for the resonance conditions are as follows.

If the sum of indices $n + m$ characterizing the transverse field structure of the initial waveguide is an even number, then all the field components in the joining point in no way differ from those for the case $\Delta\theta = 0$, and, therefore, the resonance condition (13.1) is not broken. In contrast, if the index sum $n + m$ is an odd number, then twisting by an angle $\Delta\theta = \pm\pi$ (on retention of the transverse structure of the fields) results in all the components at the output waveguide cross-section changing their signs, i.e. in their being multiplied by $-1 = \exp(\pm i\pi)$. In other words, by twisting the waveguide through the angle $\Delta\theta = \pm\pi$, an additional phase shift $\pm\pi$ appears, which should be summed with the travelling wave phase factor. This additional shift corresponds essentially to the Berry geometrical phase (see, for example, Ref. [24]). Hence a new resonance condition emerges in such a Möbius ring waveguide, amounting to the ring's optical length being equal to a half-integer number of wavelengths:

$$\zeta_{nm} L = 2\pi\left(N \pm \frac{1}{2}\right), \qquad N = 1, 2, \ldots \tag{13.2}$$

The dependence of this result upon the evenness or oddness of the transverse index sum $n + m$ gives evidence for field mode symmetry in the transverse cross-section affecting it [23].

The twisting operation can be made according to both the right- and left-screw rule ($\Delta\theta = \pm\pi$) resulting in the Möbius ring waveguide as a whole acquiring a positive or negative polarity shown in Fig. 3 by the vector $\boldsymbol{\mu}$ with absolute value

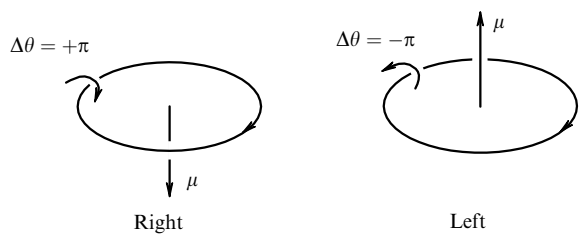

Figure 3.

$|\boldsymbol{\mu}| = N \pm 1/2$. At the same time, the ring-structures with $\Delta\theta = 0$ have no signatures of polarity.

As a result, the set of resonances in the waveguide ring-structures fall into two *topologically* different classes: class $\mathscr{B}$ with integer-type resonances (13.1), and class $\mathscr{F}$ with half-integer resonances (13.2).

Topological differences between the waveguide rings of classes $\mathscr{B}$ and $\mathscr{F}$ also become apparent when one tries to 'glue' them along the waveguide generatrices with the aim of constructing more complicated structures with a unified field composed without demolition of the identical individual mode fields. It is assumed that if by this procedure one removes the metallic walls along the 'glueing' surface, the spatial structure of the fields remains unchanged, and the result of their sewing together can be considered as some new supermode, common for all the individual rings being 'glued', with the resonance frequency coinciding with that of the individual mode.

This general statement becomes more clear if one traces the 'glueing' procedure described above for the example of a rectangular cross-section ring of both classes. The rings of class $\mathscr{B}$ allow the construction of a superstructure with p-fold width of the walls from an unlimited number p of identical individual rings. This is illustrated in Fig. 4a for the example of the mode TE_{11} in a waveguide of rectangular cross-section with sides a and b. The superstructure resulting from the 'glueing' (its linear evolvent is shown) also has a rectangular cross-section with sides a and pb, and the sewing together of individual fields, after the intermediate walls have been removed, forms a supermode TE_{1p} (electric force lines are shown by the arrows) with the resonance frequency coinciding with the initial frequency of the individual mode.

In the alternative case of rings of class $\mathscr{F}$, it is possible to 'glue' only one pair of rings with the opposite polarities along the waveguide generatrices and with the fields being sewed together into a single supermode. This is seen in the example of individual modes TE_{12} in Fig. 4b representing a linear evolvent of the 'glued pair' and its cross-sections at different points along the ring, where the supermode can be denoted by TE_{22}, $TE_{12} + TE_{21}$, TE_{41}, $TE_{21} + TE_{12}$, TE_{22} with the same resonance frequency value.

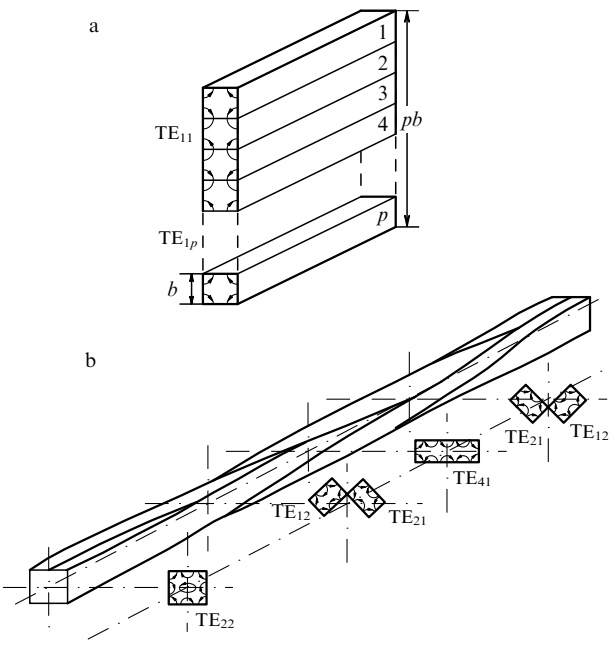

Figure 4.

The resonances considered can be called *boson-like* ($\mathscr{B}$) or *fermion-like* ($\mathscr{F}$) according to the collection of alternative features of the resonating structures of classes $\mathscr{B}$ and $\mathscr{F}$ such as:

(a) the longitudinal index being integer or half-integer;

(b) the transverse field structure being even or odd;

(c) a fixed polarity being absent or present;

(d) the ability of the field to coalesce ('glue') with the formation of a single supermode from an unlimited number of individual rings or from only a pair of rings with opposite polarity.

In addition to the resonances $\mathscr{B}$ and $\mathscr{F}$, there exist also two classes of hybrid resonances, which cannot be identified as yet. One of them ($\mathscr{H}_B$) conserves the boson-like features at $\theta = 0$ despite the oddness of the index sum $n + m$. In another class ($\mathscr{H}_F$) integer resonances emerge for an even sum $n + m$, despite the fact that $\theta = \pm\pi/2$.

It is worth emphasizing the role of the transverse field bounding which, of course, is always met in the waveguide, in constructing the ring-structures, and, in particular, in observing the Berry geometrical phase motion in them: ring formation is only possible if there is no superposition of the opposite-side ring fields. This condition corresponds to the topological requirement for the surface, from which the Möbius band is formed, to have an edge.

Not identifying, of course, the waveguide model considered with some elementary particle classes, it is worth, in conclusion, citing the following words of Dirac [25]. Recalling how "the idea of existence ... of bosons and fermions had firmly been established", he says: "It turned out that all known particles belong to one of these two types. I don't know whether some obvious reason for this exists, but this is the case". In the waveguide ring-structures, differences in field topology and symmetry which, as is well known (see, for example, Refs [12, 26]), are used to construct the theory of elementary particles, are such a visual reason for this.

14. Conclusions [27]

The following facts result from the series of thought experiments suggested:

(1) The nonzero photon proper mass M (2.3) in a waveguide is equivalent to the energy of the critical frequency quantum, accumulated in the mode during the work done against the pressure force of the zero-point vacuum fluctuations with zero initial frequency in the course of evolution from an unbounded free space to a waveguide with finite cross-section. It is this simple electromagnetic energy (a Cheshire-cat smile) that serves as the original source of mass considered as a measurable physical reality.

(2) In different problems of dynamics with some types of accelerated motion, the quantity M (2.3) is equivalent to the inertial mass in the corresponding standard relativistic equations.

(3) In a gravitational field, a photon in a waveguide behaves as a heavy particle with gravitational mass M (2.3): having been directed upwards along a vertical waveguide, it falls back after achieving a limiting height; it can be weighed using a balance, etc.

(4) The adequacy of the accepted photon mass concept is illustrated by the evolution of the classical mass-free wave

equation into a one-dimensional Schrödinger-type equation containing the mass explicitly, without its *a priori* introduction.

(5) The appearance of the nonzero photon mass is due to a standing field component ('stopped' or 'resting' light), generating the momentum defect in the Hamiltonian function, which determines the term corresponding to mass in this function.

(6) We have failed to suggest some experimental means of distinguishing the quantity M (2.3) from the standard mass concept. Therefore, the attempt to give meaning to de Broglie's analogy (Section 2) goes so far away, that it makes us think of the quantity M as a physical reality corresponding to the traditional mass concept.

(7) Depending on the topology and symmetry of the field in the waveguide ring-structures, two classes of resonances emerge, which have boson- and fermion-like features.

Do all these results need to be tested under laboratory conditions? The methodological essence of the thought experiment involving well-known and irrefutable laws (of course, under their correct use) predetermines the negative answer. The approach suggested by the present analysis pretends rather to a heuristical significance.

Finally, it is difficult to resist the temptation of issuing a provoking hypothesis [3, 9, 20]: the proper masses of any particles (not only the photon) are equivalent to the energy of the material they are made of when compressed from unbounded space into a channel with waveguide properties. This natural channel (a two-dimensional potential well) is generated by the balance between the internal pressure of matter and the counter acting proper gravitational field. The eigenvalues of the corresponding self-consistent problem yield a discrete mass spectrum. This concept differs from a spherical model of similar nature (see, for example, Refs [12, 28, 29]) by having a two-dimensional potential well construction with the continuum along the third coordinate, which apparently corresponds to a greater extent to features of a free particle.

However, Einstein's sceptical statement: "All these fifty years of hard considerations have not brought me closer to the answer of what light quanta are. Of course, today everybody thinks he knows the answer, but he deceives himself" prevents us from falling into excessive optimism.

The author sincerely acknowledges B M Bolotovskiĭ for his constant attention.

References

1. Heaviside O *Electromagnetic Theory* (New York, London, 1893)
2. Rayleigh J W *Phil. Mag.* **43** 125 (1897)
3. Rivlin L A *Kvantovaya Elektron.* (Moscow) **20** 919 (1993) [*Quantum Electron.* **23** 799 (1993)]
4. Klyshko D N *Usp. Fiz. Nauk* **164** 1187 (1994) [*Phys. Usp.* **37** 1097 (1994)]
5. de Broglie L *Problems de Propagations Guidees des Ondes Electromagnetiques* (Paris, 1941) [Translated into Russian (Moscow: GIIL, 1948)]
6. Feynman R, Leighton R, Sands M *The Feynman Lectures on Physics* Vol. 2 (London: Addison-Wesley, 1964) [Translated into Russian Vol. 6 (Moscow: Mir, 1966) p. 230]
7. Meshkov I N, Chirikov B V *Élektromagnitnoe Pole* (Electromagnetic Field) (Novosibirsk: Nauka, 1987)
8. Rivlin L A *Kvantovaya Elektron.* (Moscow) **19** 830 (1992) [*Quantum Electron.* **22** 771 (1992)]
9. Rivlin L A *Usp. Fiz. Nauk* **161** (3) 143 (1991) [*Sov. Phys. Usp.* **34** 259 (1991)]
10. Rivlin L A *Kvantovaya Elektron.* (Moscow) **22** 292 (1995) [*Quantum Electron.* **25** 275 (1995)]
11. Casimir H B G *Proc. Kon. Nederl. Akad. Wet.* **51** 793 (1948)
12. Mostepanenko V M, Trunov N N *Éffekt Kasimira i Ego Prilozheniya* (Casimir Effect and Its Applications) (Moscow: Energoatomizdat, 1990)
13. Bouler M *Gravitation and Relativity* (Oxford: Oxford University Press, 1976) [Translated into Russian (Moscow: Mir, 1979)]
14. Rivlin L A *Kvantovaya Elektron.* (Moscow) **22** 289 (1995) [*Quantum Electron.* **25** 272 (1995)]
15. Feynman R, Leighton R, Sands M *The Feynman Lectures on Physics* Vol. 1 (London: Addison-Wesley, 1964) [Translated into Russian (Moscow: Mir, 1967) p. 156]
16. Jahnke P R E, Emde F *Tablitsy Funktsiĭ* (Table of Functions) (Moscow – Leningrad: GITTL, 1948)
17. Rivlin L A *Kvantovaya Elektron.* (Moscow) **22** 625 (1995) [*Quantum Electron.* **25** 600 (1995)]
18. Okun' L B *Usp. Fiz. Nauk* **158** 511 (1989) [*Sov. Phys. Usp.* **32** 629 (1989)]
19. Okun L B *Physics Today* **42** 31 (1989)
20. Rivlin L A *Kvantovaya Elektron.* (Moscow) **6** 1087 (1979) [*Sov. J. Quantum Electron.* **9** 1320 (1979)]
21. Rivlin L A *Kvantovaya Elektron.* (Moscow) **16** 2159 (1989) [*Sov. J. Quantum Electron.* **19** 1345 (1989)]
22. Rivlin L A *Kvantovaya Elektron.* (Moscow) **18** 526 (1991) [*Sov. J. Quantum Electron.* **21** 479 (1991)]
23. Rivlin L A *Opt. Acoust. Rev.* **1** 125 (1990 – 1991)
24. Klyshko D N *Usp. Fiz. Nauk* **163** (11) 1 (1993) [*Phys. Usp.* **36** 1005 (1993)]
25. Dirac P A M *Vospominaniya o Neobychaĭnoĭ Épokhe* (Memoirs of the Exceptional Age) (Moscow: Nauka, 1990)
26. Sakharov A D, in *Problemy Teoreticheskoĭ Fiziki. Pamyati I E Tamma* (Problems in Theoretical Physics. In Memory of I E Tamm) (Moscow: Nauka, 1972) p. 242
27. Rivlin L A *Proc. SPIE* **2799** 182 (1996)
28. Markov M A, Preprint No. P-0208 (Inst. Nucl. Res. Acad. Sci. USSR, 1981)
29. Bykov V P *J. Sov. Laser Res.* **12** 38 (1991)