Diffraction of short pulses with boundary diffraction wave theory

Z. L. Horváth and Zs. Bor
Department of Optics and Quantum Electronics, University of Szeged, P.O. Box 406, H-6701 Szeged, Hungary
(Received 11 May 2000; published 11 January 2001)

The diffraction of short pulses is studied on the basis of the Miyamoto-Wolf theory of the boundary diffraction wave, which is a mathematical formulation of Young's idea about the nature of diffraction. It is pointed out that the diffracted field is given by the superposition of the *boundary wave pulse* (formed by interference of the elementary boundary diffraction waves) and the *geometric (direct) pulse* (governed by the laws of geometrical optics). The case of a circular aperture is treated in details. The diffracted field on the optical axis is calculated analytically (without any approximation) for an arbitrary temporal pulse shape. Because of the short pulse duration and the path difference the geometric and the boundary wave pulses appear separately, i.e., the boundary waves are manifested in themselves in the illuminated region (in the sense of geometrical optics). The properties of the boundary wave pulse is discussed. Its radial intensity distribution can be approximated by the Bessel function of zero order if the observation points are in the illuminated region and far from the plane of the aperture and close to the optical axis. Although the boundary wave pulse propagates on the optical axis at a speed exceeding c, it does not contradict the theory of relativity.

DOI: 10.1103/PhysRevE.63.026601 PACS number(s): 42.25.Fx, 42.25.Gy, 42.25.Bs

I. INTRODUCTION

Thomas Young was the first who made an attempt to explain the phenomena of diffraction on the basis of the wave theory [1–5]. He assumed that the diffraction pattern arises from the interference of the incident wave propagating in accordance with the laws of geometrical optics and the boundary diffraction wave originated from the edge of the diffracting body. Because of its qualitative formulation and the success of Fresnel's theory Young's idea had been forgotten for a long time. Young's views was independently rediscovered and formulated in a quantitative manner by Maggi and Rubinowicz [2–5]. The theory of the boundary diffraction wave was improved by Miyamoto and Wolf [3–6]. For more historical details of the topic see the works of Rubinowicz [2,3] and Wolf [4,5].

The temporal and spatial behavior of a focused short pulse was studied in Refs. [7–9]. Perhaps the most unexpected result of that investigation is that a spike appears on the optical axis in front of the horseshoe-shaped pulse front (see Fig. 2 in Ref. [7]). If the illumination of the lens is not homogeneous but spatially Gaussian with a negligible truncation of the input beam the forerunner pulse disappears [8] (see Fig. 2 in Ref. [8]). If the truncation of the incoming spatially Gaussian beam is not negligible the spike appears on the optical axis again with the amplitude being proportional to the amplitude of the incoming field on the lens aperture [10,11]. From this fact and the other properties of the spike (position, speed, and radial intensity distribution) we concluded that the intensity spike (pulse) is caused by the superposition of the boundary waves generated by the lens aperture [7–11]. In order to emphasize the origin of the pulse on the optical axis it was called *boundary wave pulse*. The other pulse front was named *main pulse*.

However, these are only indirect marks for the origin of the boundary wave pulse. We have owed for its direct derivation up to now. The aim of this paper is to give a direct explanation of the formation of the boundary wave pulse. For the sake of simplicity all the aberrations will be ignored, that is, we will assume that the pulse front which fills the aperture is perfectly spherical or in special case plane.

II. BOUNDARY WAVE PULSE

Consider an (arbitrary) spherical or plane scalar wave diffracted at an aperture in an opaque plane screen S. We take the Cartesian reference system with the origin O in the aperture (Fig. 1) and with the axis z being perpendicular to the plane of the aperture pointing into the half space into which the light propagates.

If the incident wave is monochromatic with angular frequency ω the diffracted wave at a point $P=(x,y,z)$ in the region $z>0$ can be expressed in the form [6]

$$U(P,\omega)=U_G(P,\omega)+U_B(P,\omega), \tag{1}$$

where $U_G(P,\omega)$ represents a wave propagating in accordance with the laws of geometrical optics and $U_B(P,\omega)$ describes a disturbance emerging from the points of the edge of

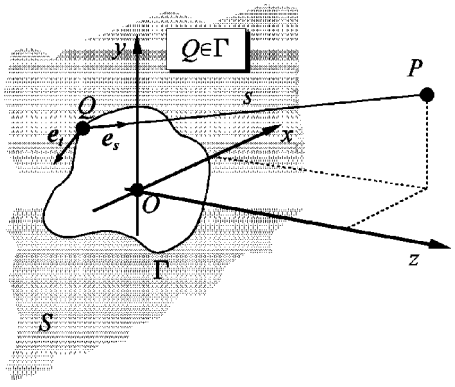

FIG. 1. Notations relating to the calculations.

the aperture. Within the accuracy of the Kirchhoff diffraction theory the $U_B(P,\omega)$ boundary diffraction wave is given by [6]

$$U_B(P,\omega)=\oint_\Gamma U_i(Q,\omega)\frac{e^{-iks}}{4\pi s}\frac{(\vec{e}_s\times\vec{p})\vec{e}_t}{1+\vec{e}_s\vec{p}}d\Gamma, \qquad (2)$$

where Γ denotes the boundary of the aperture and Q is a typical point of Γ, U_i is the incident field, s is the distance QP, $k=\omega/c$ is the wave number (c is the speed of light), $\vec{e}_s$ is the unit vector pointing from Q to P, $\vec{p}$ is a unit normal vector of the phase front of the incident wave at point Q pointing in the direction of the propagation and $\vec{e}_t$ is the unit tangent vector of Γ at point Q (Fig. 1). As it is known [6], in case of monochromatic spherical or plane waves Eq. (1) with Eq. (2) gives the exact solution of the Kirchhoff diffraction integral and it can be regarded as a good approximation for other types of monochromatic fields provided that the incident field can be treated by geometrical optics.

A nonmonochromatic field can be represented as a composition of monochromatic waves:

$$u(P,t)=\mathcal{F}^{-1}\{U(P,\omega)\}, \qquad (3)$$

where $U(P,\omega)=\mathcal{F}\{u(P,t)\}$ and the symbols $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the Fourier transformation and its inverse, respectively. Decomposing the incident field by the Fourier transform into its monochromatic components and using Eq. (1) for each spectral component one can get a similar expression for the diffracted field:

$$u(P,t)=u_G(P,t)+u_B(P,t), \qquad (4)$$

where

$$u_G(P,t)=\mathcal{F}^{-1}\{U_G(P,\omega)\}, \qquad (5a)$$

$$u_B(P,t)=\mathcal{F}^{-1}\{U_B(P,\omega)\}, \qquad (5b)$$

and $U_B(P,\omega)$ is defined by Eq. (2). Equation (4), likewise Eq. (1), is valid for spherical or plane waves and can be regarded as a good approximation in case of other types of incident fields if the input field can be described by geometrical optics. We will show that under the circumstances treated in Refs. [7–9], that is, in case of a circular aperture and convergent spherical pulse, $u_G(P,t)$ and $u_B(P,t)$ can be associated with the main and boundary wave pulses (mentioned in Sec. I), respectively. Therefore, we will refer to the field described by Eq. (5b) [with Eq. (2)] as the boundary wave pulse.

The special case of a circular aperture, with incident fields having axial symmetry around the optical axis (symmetry axis of the circular aperture), could be of special importance. In this case the amplitude of U_B and U_G could be comparable along the optical axis [12]. In the following we will study such cases. The three cases (plane, convergent, and divergent spherical waves) in which Eq. (4) gives the exact solution will be treated in detail. The radius of the aperture will be denoted by a and the origin of the reference frame will be taken in the center of the aperture.

A. Plane wave

Let us assume that the incident wave is a plane wave with normal incidence, that is

$$u_i(P,t)=u_0h(t-z/c), \qquad (6)$$

where u_0 is a constant and $h(t)$ describes the arbitrary time evolution of the pulse. It is evident that the wave propagating in accordance with the laws of geometrical optics is given by

$$u_G(P,t)=\begin{cases}u_i(P,t), & \text{if } P \text{ is in the direct beam } (r<a),\\ 0, & \text{if } P \text{ is in the shadow } (a<r).\end{cases} \qquad (7)$$

By calculating the Fourier transform of $u_i(P,t)$ one can obtain that

$$U_i(P,\omega)=\mathcal{F}\{u_i(P,t)\}=u_0H(\omega)e^{-ikz}, \qquad (8)$$

where $H(\omega)=\mathcal{F}\{h(t)\}$. At the plane of the aperture $z=0$ so $U_i(Q,\omega)=u_0H(\omega)$. After a straightforward calculation the monochromatic boundary diffraction wave at a point P in the region $z>0$ is given by

$$U_B(r,z,\omega)=\frac{u_0H(\omega)}{2\pi}\int_0^\pi e^{-iks(\psi)}\left(1+\frac{z}{s(\psi)}\right)g(K,\psi)d\psi, \qquad (9)$$

where $r=\sqrt{x^2+y^2}$ is the distance of point P from the optical axis,

$$s(\psi)=\sqrt{z^2+a^2+r^2-2ar\cos\psi}, \qquad (10)$$

$K=r/a$ is a dimensionless variable and

$$g(K,\psi)=\frac{K\cos\psi-1}{1+K^2-2K\cos\psi}. \qquad (11)$$

From the definition of K one can see that $0\leqslant K<1$ if P is in the direct beam, and $K>1$ when P is in the geometrical shadow. The field of the boundary wave pulse is the inverse Fourier transform of Eq. (9):

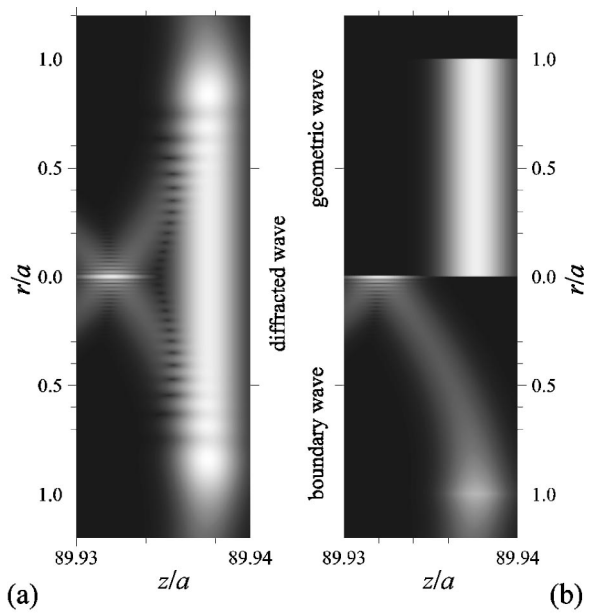

FIG. 2. (a) Diffraction pattern of a short pulse having plane pulse front incident normally on a circular aperture (a denotes the radius of the aperture). (b) The intensity distribution of the geometric (upper half) and boundary wave (lower half) pulses. The diffracted field is given by the sum of the fields of the geometric and the boundary wave pulses. The intensity is measured in arbitrary units [a.u.].

$$u_B(r,z,t)=\frac{u_0}{2\pi}\int_0^{\pi}h(t-s(\psi)/c)\left(1+\frac{z}{s(\psi)}\right)g(K,\psi)d\psi \tag{12a}$$

$$=\frac{u_0e^{i\omega_0t}}{2\pi}\int_0^{\pi}v(t-s(\psi)/c)\times e^{-ik_0s(\psi)}\left(1+\frac{z}{s(\psi)}\right)g(K,\psi)d\psi, \tag{12b}$$

where in the last step the usual expression of

$$h(t)=v(t)e^{i\omega_0t} \tag{13}$$

was used. Here $v(t)$ and ω_0 denote the temporal envelope and the central angular frequency of the input pulse, respectively, and $k_0=\omega_0/c$ is the wave number at ω_0. It is easy to see that on the optical axis the integrand in Eq. (12) does not depend on ψ so for $r=0$ the boundary wave pulse is given by

$$u_B(z,t)=-A_B(z)u_0h(t-s_0/c), \tag{14}$$

where

$$s_0=\sqrt{a^2+z^2}, \tag{15}$$

and $A_B(z)$ is the relative amplitude of the boundary wave pulse (compared to the incident pulse):

$$A_B(z)=\frac{1+z/s_0}{2}=\frac{1+z/\sqrt{a^2+z^2}}{2}. \tag{16}$$

It is worth introducing a new variable defined by

$$T(z)=\frac{s_0-z}{c}=\frac{\sqrt{a^2+z^2}-z}{c}. \tag{17}$$

Then one can write Eq. (14) in the form of

$$u_B(z,t)=-A_B(z)u_0h[t-z/c-T(z)] \tag{18a}$$

$$=-A_B(z)u_i[z,t-T(z)]. \tag{18b}$$

As mentioned before, the diffracted field in the half space $z>0$ can be calculated by $u(P,t)=u_G(P,t)+u_B(P,t)$ [Eq. (4)] where u_G and u_B are given by Eqs. (7) and (12), respectively. It follows from Eq. (18) that for the points P at the optical axis we have a simple expression:

$$u(z,t)=u_i(z,t)-A_B(z)u_i[z,t-T(z)], \tag{19}$$

which means that as a result of the diffraction two pulses propagate on the optical axis with the same temporal shape. The time difference between the two pulses at an axial point P is given by $T(z)$. One can see from Eq. (17) that $0<T(z)<a/c$ and $T(z)$ is a monotonously decreasing function of z, i.e., larger values of z yield smaller temporal separation. This means that the boundary wave pulse always arrives later than the geometric pulse, but it is catching up as time evolves. The minus sign shows that a phase shift of π occurs between the geometric (direct beam) and the boundary wave pulses. It follows from Eq. (16) that the amplitude $A_B(z)$ varies between 1/2 and 1 and asymptotically $A_B(z)\approx 1$ if $a^2\ll z^2$.

Figure 2(a) shows the diffraction pattern of a $\tau=10$ fs pulse with a Gaussian temporal shape and $\lambda_0=2\pi/k_0=800$ nm central wavelength at time $t=600$ ps calculated from Eqs. (4), (7) and (12). The incoming pulse reaches the plane of the diffracting aperture at $t=0$. The radius of the aperture was assumed $a=2$ mm. In Fig. 2(b) the intensity distribution of the geometric [calculated from Eq. (7)] and the boundary diffraction waves [calculated from Eq. (12)] was plotted in the upper and the lower halves of the figure, respectively. The correspondence between the intensity and the shading is indicated by the gray scale next to the linear intensity axis in Fig. 4(a). The diffraction pattern contains two types of interference stripes. One of them is caused by the interference of the elementary boundary diffraction wavelets. These stripes can be seen both in Figs. 2(a) and (b). The other interference stripes are caused by the interference between the geometric and the boundary diffraction waves. Obviously this type of stripes occurs only in Fig. 2(a).

B. Divergent spherical wave

Consider a divergent spherical wave generated by a point source F located on the optical axis, that is, $F=(0,0,-d)$,

where $d>0$ is the distance of the source from the plane of the aperture. Then the incident wave is given by

$$u_i(P,t)=u_0\frac{h(t-R/c)}{R}, \tag{20}$$

where u_0 is a constant, $h(t)$ is an arbitrary function describing the temporal shape of the pulse, and $R=\sqrt{x^2+y^2+(d+z)^2}=\sqrt{r^2+(d+z)^2}$ is the distance FP. It is clear that the wave propagating according to the law of geometrical optics is described by

$$u_G(P,t)=\begin{cases} u_i(P,t), & \text{if } P \text{ is in the direct beam } [r<ad/(d+z)], \\ 0, & \text{if } P \text{ is in the shadow } [ad/(d+z)<r]. \end{cases} \tag{21}$$

The monochromatic components of $u_i(P,t)$ is given by

$$U_i(P,\omega)=\mathcal{F}\{u_i(P,t)\}=u_0H(\omega)\frac{e^{-ikR}}{R}, \tag{22}$$

where $H(\omega)=\mathcal{F}\{h(t)\}$. After a long, but straightforward calculation one can obtain

$$U_B(r,z,\omega)=\frac{u_0H(\omega)e^{-ikf}}{2\pi(d+z)}\int_0^{\pi}e^{-iks(\psi)}\times\left(1+\frac{dz-a^2+ar\cos\psi}{s(\psi)f}\right)g(K,L,\psi)d\psi \tag{23}$$

for the monochromatic boundary diffraction wave at a point P in the region $z>0$, where $f=\sqrt{a^2+d^2}$ is the radius of the wave front that fills the aperture, $s=s(\psi)$ is the same as in Eq. (10), K a dimensionless variable given by

$$K=\frac{d}{d+z}\frac{r}{a}, \tag{24}$$

$L=a/d$ is a dimensionless parameter and

$$g(K,L,\psi)=\frac{K\cos\psi-1}{1+K^2-2K\cos\psi+K^2L^2\sin^2\psi}. \tag{25}$$

From Eq. (24) it follows that $0\leqslant K<1$ if P is in the direct beam, and $K>1$ when P is in the geometrical shadow, likewise in the case of plane waves. By calculating the inverse Fourier transform of Eq. (23) one can obtain the field of the boundary wave pulse:

$$u_B(r,z,t)=\frac{u_0}{2\pi(d+z)}\int_0^{\pi}h\left(t-\frac{f+s(\psi)}{c}\right)\times\left(1+\frac{dz-a^2+ar\cos\psi}{s(\psi)f}\right)g(K,L,\psi)d\psi \tag{26a}$$

$$=\frac{u_0e^{i(\omega_0t-k_0f)}}{2\pi(d+z)}\int_0^{\pi}v\left(t-\frac{f+s(\psi)}{c}\right)e^{-ik_0s(\psi)}\times\left(1+\frac{dz-a^2+ar\cos\psi}{s(\psi)f}\right)g(K,L,\psi)d\psi, \tag{26b}$$

where in the last step $h(t)=v(t)e^{i\omega_0t}$ is used again. On the optical axis ($r=0$) the integrand in Eq. (26) is a constant so the integration results in a multiplication by π:

$$u_B(z,t)=-A_B(z)u_0\frac{h[t-(f+s_0)/c]}{d+z}, \tag{27}$$

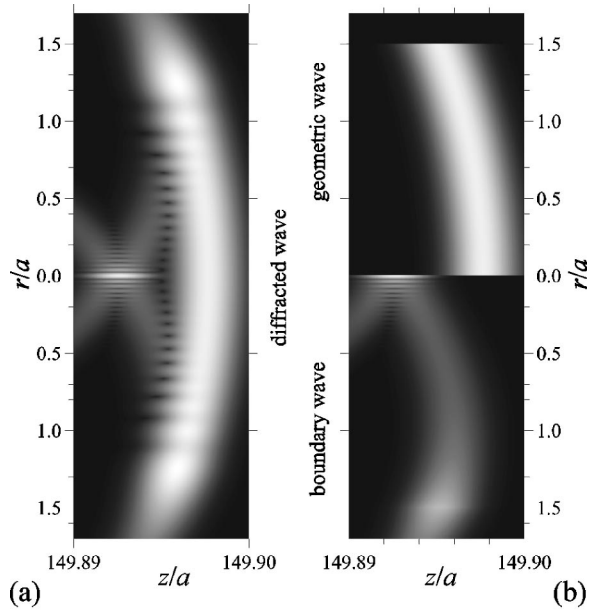

FIG. 3. (a) Diffraction pattern of a short divergent spherical pulse incident on a circular aperture. The center of the incident pulse front is located on the optical axis. (b) The intensity distribution of the geometric (upper half) and boundary wave (lower half) pulses. The interference of the geometric and the boundary wave pulse yields the diffraction pattern plotted on the left half (a). The intensity is measured in arbitrary units [a.u.].

where $s_0=\sqrt{a^2+z^2}$ [Eq. (15)] and the relative amplitude of the boundary wave pulse is given by

$$A_B(z)=\frac{1}{2}\left(1+\frac{dz-a^2}{f\sqrt{a^2+z^2}}\right). \tag{28}$$

If we introduce a variable with a dimension of time defined by

$$T(z)=\frac{f-d}{c}+\frac{s_0-z}{c} \tag{29a}$$

$$=\frac{f-d}{c}+\frac{\sqrt{a^2+z^2}-z}{c}, \tag{29b}$$

Eq. (27) can be written in the form of

$$u_B(z,t)=-A_B(z)u_i[z,t-T(z)]. \tag{30}$$

The diffracted wave behind the screen is the sum of the fields of the geometric (direct) beam and the boundary wave pulse [Eq. (4)]. Using Eqs. (21) and (30) on the optical axis it can be written in a form of

$$u(z,t)=u_i(z,t)-A_B(z)u_i[z,t-T(z)]. \tag{31}$$

One can conclude from Eq. (31) that, as a result of the diffraction, two pulses with the same temporal shape propagate along the optical axis and the temporal shape of the pulses is identical with the input pulse shape. It follows from Eq. (29) that $0<(f-d)/c<T(z)<(f-d)/c+a/c$ and $T(z)$ decreases monotonically. From this we draw a conclusion similar to the plane wave case, that is at a point given by $z>0$ the boundary wave pulse always arrives later than the geometric pulse. The temporal separation of the two pulses is always larger than $(f-d)/c$, and it decreases with increasing z.

Figure 3(a) shows the diffraction pattern of a $\tau=10$ fs pulse with a Gaussian temporal shape and $\lambda_0=800$ nm central wavelength at the moment 1000 ps after the incoming pulse reached the edge of the aperture (i.e., $t=f/c+1000$ ps $\approx$3000 ps) calculated from Eqs. (4), (21), and (26). The calculation was done by assuming that $a=2$ mm and $f=600$ mm. The intensity distribution of the geometric [Eq. (21)] and the boundary diffraction waves [Eq. (26)] were depicted in the upper and the lower halves of Fig. 3(b). The intensity distributions were plotted with the same gray scale used in the previous case. Again we can observe the two types of interference stripes. The ones, occurring both in Figs. 3(a) and (b), correspond to the interference of the elementary boundary diffraction wavelets, whereas the stripes observable only in Fig. 3(a) correspond to the interference between the geometric and boundary waves.

C. Convergent spherical wave

Consider now a spherical wave converging towards an axial focal point F in the region $z>0$. Then $F=(0,0,d)$, where $d>0$ is the distance of the focus from the plane of the aperture. In front of the focal plane ($z<d$) the incident wave is represented by a converging spherical wave. When the input wave passes through the focus it becomes a divergent spherical wave and a phase change of π occurs [4,12]. This behavior (known as phase anomaly or Gouy shift) is a geometrical optical effect that occurs along the rays passing through either of the two principal centers[1] of the incident wave front [4,6]. Hence the incident wave can be written in a form of

$$u_i(P,t)=\begin{cases} u_0h(t+R/c)/R, & \text{if } z<d,\\ -u_0h(t-R/c)/R, & \text{if } d<z, \end{cases} \tag{32}$$

where u_0 is a constant, $h(t)$ is an arbitrary function and $R=\sqrt{x^2+y^2+(z-d)^2}=\sqrt{r^2+(z-d)^2}$ is the distance FP and the sign change represents the phase shift of π mentioned above. It is easy to show that the wave propagating according to the law of geometrical optics is given by

$$u_G(P,t)=\begin{cases} u_i(P,t), & \text{if } P \text{ is in the direct beam } [r<ad/|z-d|],\\ 0, & \text{if } P \text{ is in the shadow } [ad/|z-d|<r]. \end{cases} \tag{33}$$

On the optical axis Eq. (33) yields

$$u_G(z,t)=u_i(z,t)=-u_0\frac{h(t-(z-d)/c)}{z-d}. \tag{34}$$

In the same way as treated in the two previous cases one can calculate the boundary wave pulse. It is given by

$$u_B(r,z,t)=\frac{u_0}{2\pi(d-z)}\int_0^{\pi}h\left(t-\frac{s(\psi)-f}{c}\right)\times\left(1+\frac{dz+a^2-ar\cos\psi}{s(\psi)f}\right)g(K,L,\psi)d\psi \tag{35a}$$

$$=\frac{u_0e^{i(\omega_0t+k_0f)}}{2\pi(d-z)}\int_0^{\pi}v\left(t-\frac{s(\psi)-f}{c}\right)e^{-ik_0s(\psi)}\times\left(1+\frac{dz+a^2-ar\cos\psi}{s(\psi)f}\right)g(K,L,\psi)d\psi, \tag{35b}$$

where $f=\sqrt{a^2+d^2}$ is the focal length (the radius of the pulse

[1] In our special case the two principal centers coincide.

front filling the aperture), the dimensionless variable K is now defined by

$$K=\frac{d}{d-z}\frac{r}{a}, \tag{36}$$

and in the last step $h(t)=v(t)e^{i\omega_0 t}$ is used again. The definitions of $H(\omega)$, $s=s(\psi)$, L and $g(K,L,\psi)$ are the same as the ones for the divergent spherical wave, given by $H(\omega)=\mathcal{F}\{h(t)\}$, Eq.(10), $L=a/d$ and Eq. (25), respectively.

It follows from Eq. (36) that $0\leqslant|K|<1$ if P is in the direct beam, $|K|>1$ when P is in the geometrical shadow,[2] and $0\leqslant K$ if P is between the plane of the aperture and the focal plane $(0<z<d)$ or $K\leqslant 0$ for points behind the focal plane $(d<z)$.

On the optical axis $(r=0)$ Eq. (35) yields

$$u_B(z,t)=A_B(z)u_0\frac{h[t-(s_0-f)/c]}{z-d}, \tag{37}$$

where $s_0=\sqrt{a^2+z^2}$, and the relative amplitude of the boundary wave pulse is given by

$$A_B(z)=\frac{1}{2}\left(1+\frac{dz+a^2}{f\sqrt{a^2+z^2}}\right). \tag{38}$$

Again, we introduce a variable with a dimension of time defined by

$$T(z)=\frac{d-f}{c}+\frac{s_0-z}{c} \tag{39a}$$

$$=\frac{d-f}{c}+\frac{\sqrt{a^2+z^2}-z}{c}, \tag{39b}$$

then Eq. (37) can be written in the form of

$$u_B(z,t)=-A_B(z)u_i(z,t-T(z)). \tag{40}$$

As it was mentioned before, the diffracted wave behind the screen is the sum of the fields of the geometric (direct) beam and boundary wave pulse [Eq. (4)]. It follows from Eqs. (33) and (40) that the field on the optical axis can be written in a form of

$$u(z,t)=u_i(z,t)-A_B(z)u_i(z,t-T(z)). \tag{41}$$

From Eq. (41) we see that two pulses with the same temporal shape propagate on the optical axis and the temporal shape of the pulses is identical with the input pulse shape as was the case before. It follows from Eq. (39) that the temporal separation of the two pulses $(T(z))$ decreases monotonously between $a/c-(f-d)/c>0$ and $-(f-d)/c<0$ and $T(z)=0$ at the focal point $(z=d)$. That is, $T(z)$ is positive in front of the focal point $(0<z<d)$ and negative after the focal point $(d<z)$, which means that at an axial point behind the aperture the boundary wave pulse arrives later than the geometric pulse in front of the focal point, and it precedes the geometric pulse after the focus (see Figs. 4 and 5). The boundary wave pulse overtakes the geometric pulse at the focal point. At the focal point both $u_G(z,t)$ and $u_B(z,t)$ has singularity [see Eqs. (34) and Eq. (37)], but the sum of the two fields is finite. Calculating the limit of Eq. (41) one can obtain the field at the focus $(z=d)$ by

$$u(F,t)=u_0\frac{f-d}{cf}h'(t). \tag{42}$$

The statements stated above are in full agreement with our previous results published in Refs. [7,9,11]. Using the approximations applied there,[3] Eqs. (41) and (42) turn into the equation for the focused field published previously [Eq. (18) in Ref. [9] and Eq. (18b) in Ref. [11]]. In this treatment no approximation was used. Equation (41) gives the exact solution of the diffraction problem within the validity of Kirchhoff's diffraction theory.

Figures 4(a) and 5(a) show the diffraction pattern of a $\tau=10$ fs long pulse with a Gaussian temporal shape and $\lambda_0=800$ nm central wavelength at the moments $t=-30$ ps and $t=30$ ps, respectively [calculated from Eqs. (4), (33), and (35)]. The gray scale next to the linear intensity axis shows the relation between the intensity and the shading. The shaded image on the right was depicted with proportional scaling of the spatial coordinates $(z-d)$ and r. The calculation was done by assuming that $a=2$ mm and $f=50$ mm. The incoming pulse reaches the edge of the aperture at $t=-f/c$ and passes through the focus at $t=0$. In Figs. 4(b) and 5(b) the intensity distribution of the geometric [Eq. (33)] and the boundary diffraction wave [Eq. (35)] was depicted in the upper and the lower halves of the figure. Comparing Fig. 5 with Fig. 4 one can see that at time $t=30$ ps the boundary wave pulse is a little bit closer to the geometric pulse than at time $t=-30$ ps. This is a manifestation of the asymmetry of the focused field which becomes more and more considerable for smaller Fresnel numbers [13]. In our case the Fresnel number is $N=a^2/(\lambda_0 f)=100$.

D. Properties of the boundary wave pulse

From Figs. 2–5 one can conclude that the shape of the boundary wave pulse is similar to a letter X. This X shaped profile is not the only similarity between the boundary diffraction wave pulse and the so-called X wave [14–16]. We will later see that the boundary wave pulse propagates at superluminal velocity likewise an X wave [15,16]. Besides the similarities there are differences between the two waves. The most important one may be that in contrast to an X wave, which is a nondiffracting beam (i.e., it preserves its

[2]When the observation point is in the geometrical shadow, $K\rightarrow\pm\infty$ if $z\rightarrow d$ and r is fixed. Then, because of large values of K, Eq. (35) is not convenient for the computation of the boundary diffraction wave. A more Appropriate equivalent formula can be found in the Appendix [Eq. (A6)].

[3]In Refs. [7,9] paraxial and Debye approximations, in Ref. [11] paraxial approximation was used.

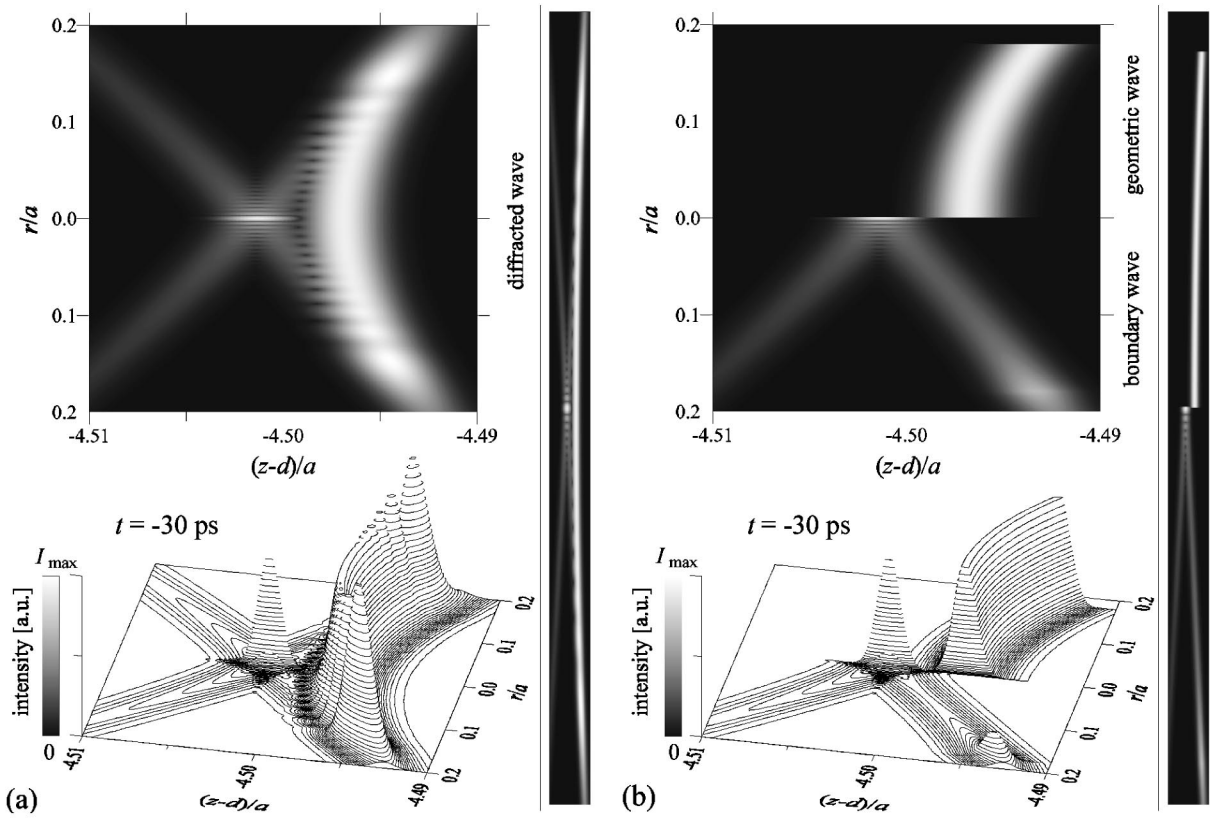

FIG. 4. (a) Diffraction pattern of a short convergent spherical pulse passing through a circular aperture before the pulse reaches the focus. The inset on the right shows the intensity distribution with proportional scaling of the spatial coordinates. (b) The intensity distribution of the geometric (upper half) and boundary wave (lower half) pulses. On the optical axis in front of the focus ($z<d$) the boundary wave pulse is closer to the aperture than the geometric pulse.

radial profile), the boundary wave pulse spreads during the propagation [see Eq. (53), Figs. 4 and 5].

The boundary wave pulse has a significant intensity on the optical axis and the boundary of the geometrical shadow. This behavior is expected. Because of the cylindrical symmetry the interference of the elementary boundary diffraction waves is constructive along the optical axis. Since the diffracted wave is a continuous function of the position and the geometrical wave is discontinuous across the edge of the geometrical shadow, the boundary pulse should also be discontinuous in order to compensate for the discontinuity.

It has been shown that for all the three previous cases the boundary wave pulse on the optical axis is given by

$$u_B(z,t) = -A_B(z)u_i[z,t-T(z)], \tag{43}$$

where $A_B(z)$ is the relative amplitude and $T(z)$ the time delay compared to the direct (geometric) wave [see Eqs. (18), (30), and (40)]. Both of these quantities have a simple geometrical meaning. Using the notations of Fig. 6, it is obvious that $\cos\alpha = z/s_0$, $\sin\alpha = a/s_0$, $\cos\beta = d/f$ and $\sin\beta = a/f$. Hence, Eqs. (16), (28), and (38) can be summarized in a uniform expression:

$$A_B(z) = \frac{1+\cos\chi}{2} = \cos^2\frac{\chi}{2}, \tag{44}$$

where χ is the angle which the scattered ray starting from a typical point Q of the aperture and passing through the axial point given by z makes with the direction of the propagation of the incoming wave (see Fig. 6). If t_Q and t_O denote the moments when the incoming pulse reaches the edge of the aperture and the origin of the reference frame (Fig. 6) then the time delay can also be expressed in a uniform expression of

$$T(z) = \Delta t + \frac{s_0 - z}{c}, \tag{45}$$

where $\Delta t = t_Q - t_O$ is a constant (depending on geometry). The position of the boundary wave pulse on the optical axis at the moment $t \geq t_Q + a/c$ can be calculated from Eqs. (43) and (45). It is given by

$$z_B(t) = \sqrt{[c(t-t_Q)]^2 - a^2}. \tag{46}$$

Hence the peak of the boundary wave pulse moves along the optical axis at a velocity of

$$v_B = \dot{z}_B = \frac{cc(t-t_Q)}{\sqrt{[c(t-t_Q)]^2 - a^2}} = c\frac{s_0(z_B)}{z_B} = \frac{c}{\cos\alpha}, \tag{47}$$

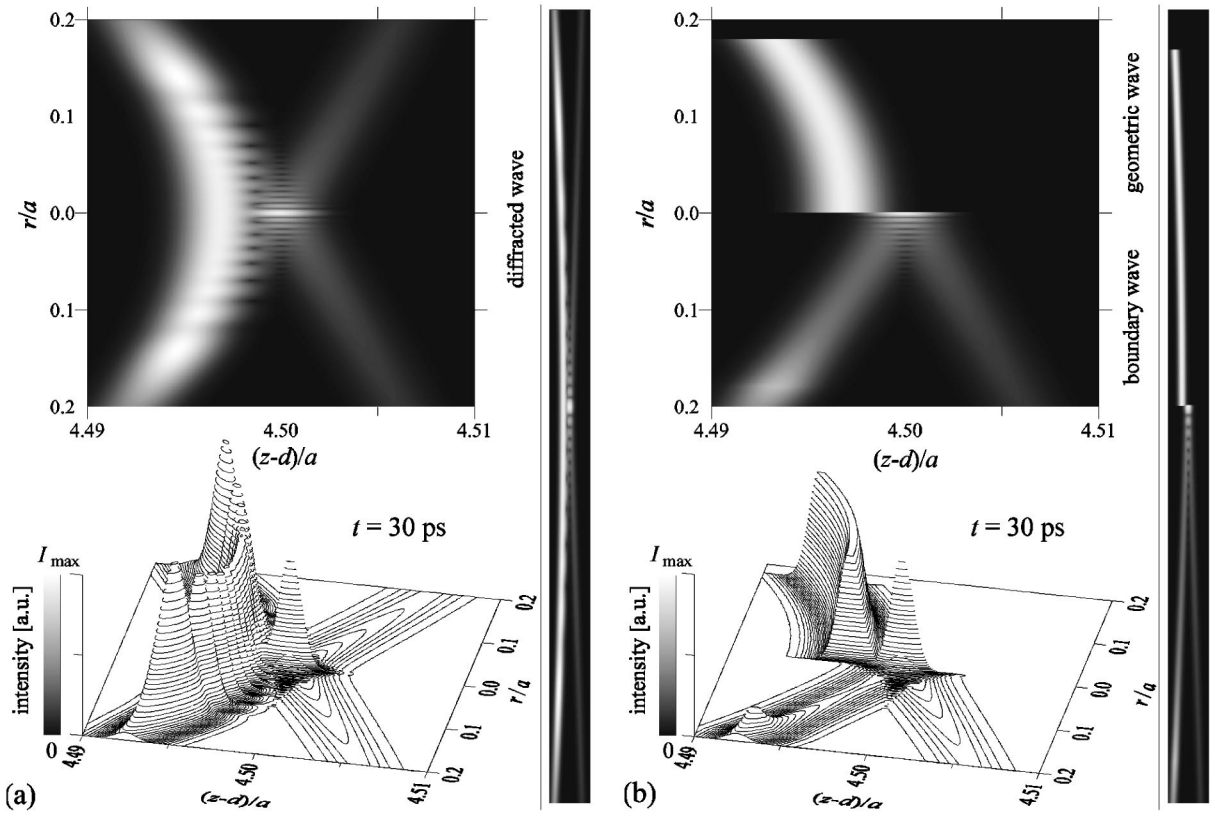

FIG. 5. (a) Diffraction pattern of a short convergent spherical pulse passing through a circular aperture after the pulse has passed through the focus. The inset on the right shows the intensity distribution with proportional scaling of the spatial coordinates. (b) The intensity distribution of the geometric (upper half) and boundary wave (lower half) pulses. On the optical axis behind the focus ($d<z$) the boundary wave pulse is farther from the aperture than the geometric pulse.

where α is marked in Fig. 6. Although this velocity *exceeds c*, it does not violate the relativistic causality principle, since the elementary boundary wavelets originated from the edge of the aperture propagate *exactly* at a speed of c along the scattered rays. Because of this an interference phenomenon on the optical axis caused by boundary diffraction waves (and so the boundary wave pulse) moves at a speed of $c/\cos\alpha$, where α is the angle which the scattered ray makes with the optical axis at an axial point (see Fig. 6).

As mentioned in the previous section the interference among the elementary boundary waves produces an interference pattern consisting of concentric rings in a plane being perpendicular to the optical axis. The structure of that pattern can hardly be seen in Figs. 2–5. In order to expose the details, a portion of Fig. 4(a) (neighboring the optical axis) is depicted again in Fig. 7(a) with a proportional scaling of the spatial coordinates. The position of the boundary wave pulse on the z axis at $t=-30$ ps is denoted by z_0 (i.e., $z_0=z_B(t)$). The radial distribution of the boundary wave pulse resembles the one of a Bessel beam [17] of zero order, as its radial intensity distribution is given by the square of the Bessel function of zero order (J_0^2). The radial intensity distribution of the boundary wave pulse can be explained as follows: The expression of the boundary wave pulse for all the three cases can be written in a form of

$$u_B(r,z,t)=\frac{u_0 e^{i\omega_0(t-t_Q)}}{\pi}\int_0^{\pi} v(t-t_Q-s(\psi)/c)e^{-ik_0 s(\psi)} \times A(r,z,\psi)g(K,L,\psi)d\psi, \tag{48}$$

where $A(r,z,\psi)$ is a given function (and $L=0$ for plane waves). If the observation point is in the direct beam and it is close to the optical axis, $|K|\ll 1$ and hence $g(K,L,\psi)\approx -1$. Furthermore, if the observation point is far from the plane of the aperture, by expanding the square root in Eq. (10), s can be approximated by

$$s=\sqrt{s_0^2+r^2-2ar\cos\psi}\approx s_0+\frac{r^2-2ar\cos\psi}{2s_0}. \tag{49}$$

The last approximation is valid if $|r^2-2ar\cos\psi|\ll s_0^2$, where $s_0=\sqrt{a^2+z^2}$. Under these conditions $A(r,z,\psi)$ and $g(K,L,\psi)$ are slowly varying functions of ψ. We will approximate them in the integrand by their values at the optical axis (that is $A_0(z)=A(0,z,\psi)$ and -1, respectively). The variation of the integrand caused by envelope $v(t)$ is also negligible if $ra/s_0=r\sin\alpha\ll c\tau$, where τ denotes the temporal duration of the pulse. That is, if the observation point is in the direct beam, and it is far from the plane of the aperture and close to the optical axis and $r\sin\alpha\ll c\tau$, the field of the boundary wave pulse can be approximated by

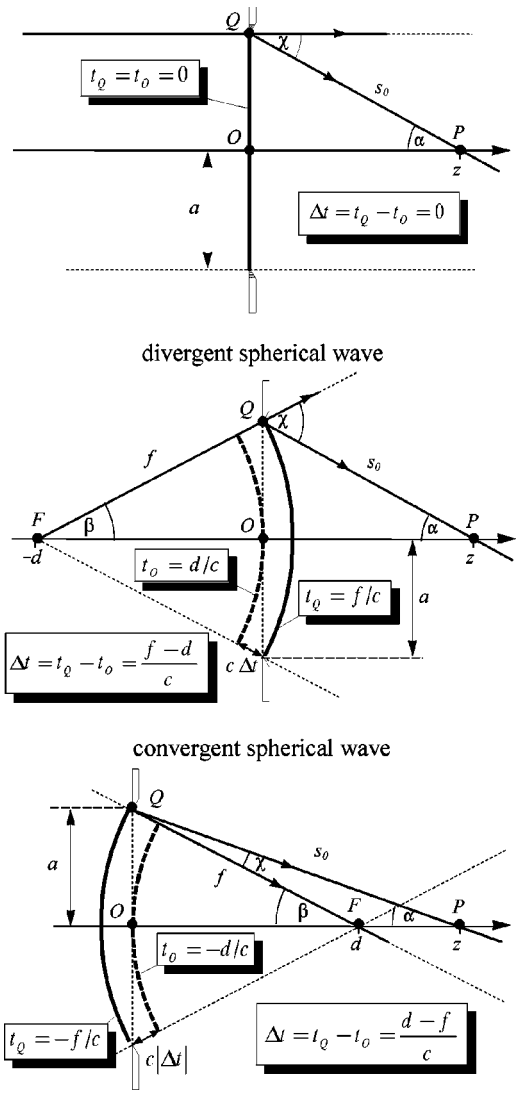

FIG. 6. Notations relating to the calculations for the axial behavior of the boundary wave pulse. t_Q and t_O denote the time instants when the pulse front reaches the edge of the aperture and the origin of the reference frame, respectively. The pulse fronts belonging to these instants are plotted with solid and dashed line.

$$u_B(r,z,t) \approx -A_0(z)u_0 h\left(t - t_Q - \frac{s_0}{c} - \frac{r^2/(2s_0)}{c}\right) \times \left[\frac{1}{\pi}\int_0^{\pi} e^{ik_0 r \sin\alpha \cos\psi} d\psi\right], \quad (50)$$

where the expression of $h(t)=v(t)e^{i\omega_0 t}$ is used in the last step. Since, the integral in the square bracket equals to $\pi J_0(k_0 \sin\alpha r)$ and $A_0(z)u_0 h(t-t_Q-s_0/c)=A_B(z)u_i(z,t-T(z))$, we have

$$u_B(r,z,t) \approx -A_B(z)u_i\left(z, t - T(z) - \frac{r^2}{2c\sqrt{a^2+z^2}}\right) \times J_0(k_0 \sin\alpha r), \quad (51)$$

where $\sin\alpha = a/\sqrt{a^2+z^2}$. Owing to our assumptions ($r \sin\alpha \ll c\tau$ and $r \ll a$) the radial variation caused by the factor $r^2/(2c\sqrt{a^2+z^2}) = (r/a)r\sin\alpha/(2c)$ is negligible with respect to J_0, that is the field of the boundary pulse can be approximated by

$$u_B(r,z,t) \approx -A_B(z)u_i(z,t-T(z))J_0(k_0 \sin\alpha r). \quad (52)$$

Figure 7(b) shows the comparison of the radial intensity distribution at $t=-30$ ps in plane $z_0 = z_B(t)$ calculated from the exact [Eq. (35)] and the approximate [Eq. (52)] formulas for different values of the incoming pulse duration τ. The rest of the parameters of the calculation were the same as the ones used for Fig. 7(a). The results of the exact expression were displayed by dotted ($\tau = 10$ fs) and solid ($\tau = 30$ fs) lines. The intensity belonging to the approximate formula was indicated with hollow circles. The approximation is much better for larger pulse duration. For smaller pulse duration the intensity decreases faster than the approximate intensity, but the approximation gives the radii of the dark rings accurately. The radius of the mth dark ring is given by

$$r_m = \frac{x_m}{2\pi}\sqrt{1+\left(\frac{z}{a}\right)^2}\lambda_0 \approx \frac{x_m}{2\pi}\frac{z}{a}\lambda_0, \quad (53)$$

where x_m is the mth positive root of the equation $J_0(x)=0$.[4]

III. CONCLUSIONS

The diffraction of femtosecond pulses at a circular aperture has been studied on the basis of the Miyamoto-Wolf theory of the boundary diffraction wave which is a mathematical formulation of Young's idea about the nature of diffraction. It has been pointed out that the diffracted pulse can be represented as a sum of the boundary wave pulse (generated by the edge of the aperture) and the ''geometric'' pulse (represented by geometrical optics). Because of the pulsed illumination and the path difference between the geometric and boundary waves the geometric and boundary wave pulses appear separately. This is the main difference between the monochromatic and pulsed illumination. The monochromatic waves are infinite waves hence the boundary wave and the geometric wave appear simultaneously. There is no way to separate the two waves. The boundary wave can directly be observed only in the geometric shadow where it equals to the diffracted wave. In case of pulsed illumination, because of the finite and short pulse duration, the boundary waves are manifested not only in the geometrical shadow but in the illuminated region, too.

The properties of the boundary wave pulse have been studied. The diffracted field on the optical axis has been calculated analytically for arbitrary temporal pulse shapes. It has been pointed out that the boundary wave pulse propagates on the optical axis at a speed larger than c. An approximate expression has been derived for the intensity of the boundary wave pulse in a plane perpendicular to the optical

[4] $x_1 = 2.405$, $x_2 = 5.5201$, $x_3 = 8.6537$, $x_4 = 11.7915$.

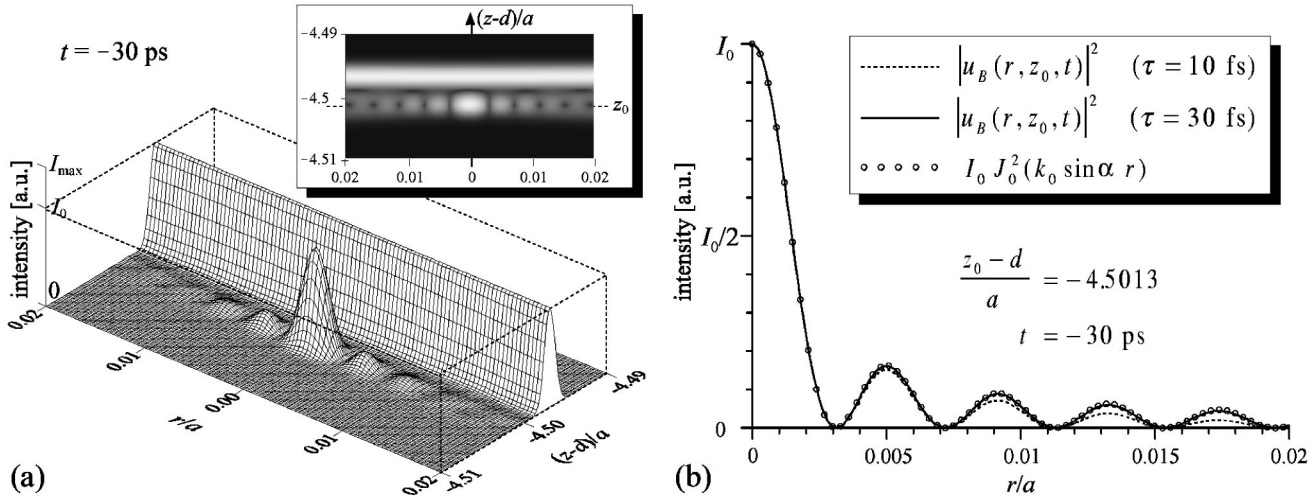

FIG. 7. (a) In order to expose the details of the radial distribution of the intensity the axial region of Fig. 4(a) is depicted again with proportional scaling of the coordinates. The radial intensity distribution of the boundary wave pulse resembles the one of a zero order Bessel beam. (b) Radial intensity distribution of the boundary wave pulse for different values of the pulse duration τ. Close to the axis the intensity can be approximated by the Bessel function of zero order (J_0). The larger the pulse duration the better the approximation.

axis. The radial intensity distribution can be approximated by the Bessel function of zero order if the observation points are in the illuminated region (in the sense of geometrical optics) far from the plane of the aperture and close to the optical axis.

Although the Miyamoto-Wolf theory of the boundary diffraction wave gives the exact result of the Kirchhoff diffraction integral only for spherical and plane waves, it can be regarded as a good approximation for other types of waves. Therefore the boundary wave pulse can appear for other types of incoming waves (or in presence of aberrations). As it was mentioned in the introduction, the boundary wave pulse appears when a short pulse is focused by a lens and the incoming beam is truncated by the lens aperture for both homogeneous or spatially Gaussian illumination and in absence of aberrations or in presence of chromatic aberration [7,9–11].

ACKNOWLEDGMENTS

We thank Katalin Varjú for her valuable remarks and suggestions. This work has been supported by FKFP Grant No. 208/97.

APPENDIX

Remarks on the numerical calculations

As it was mentioned before the boundary wave is discontinuous across the edge of the geometrical shadow. This discontinuity arises from the factor in the integrand given by

$$g(K,L,\psi)=\frac{K\cos\psi-1}{1+K^2-2K\cos\psi+K^2L^2\sin^2\psi}, \qquad \text{(A1)}$$

where the constant L is positive for a spherical wave and zero for a plane wave. At the edge of the geometrical shadow $|K|=1$. If $|K|\neq 1$ but it is close to unity then $g(K,L,\psi)$ has a sharp spike at $\psi=0$ when $K>0$ or at $\psi=\pi$ when $K<0$ with the peak values of $g(K,L,0)=1/(K-1)$ and $g(K,L,\pi)=-1/(K+1)$, respectively. But, despite of such a diverging behavior, the integral of $g(K,L,\psi)$ is finite and it is given by

$$\int_0^{\pi} g(K,L,\psi)\,d\psi=\begin{cases}-\pi/\sqrt{1+K^2L^2}, & \text{if } |K|<1,\\ 0, & \text{if } |K|>1,\\ -(\pi/2)/\sqrt{1+K^2L^2}, & \text{if } |K|=1.\end{cases} \qquad \text{(A2)}$$

Moreover the integral of $g(K,L,\psi)$ can be written in the form of

$$\int_0^{x} g(K,L,\psi)\,d\psi=\sum_{i=1}^{2} C_i \arctan\left(q_i\tan\frac{x}{2}\right), \qquad \text{(A3)}$$

where $0\leqslant x\leqslant\pi$ and the parameters C_1, C_2, q_1, and q_2 depend on K and L. The fact that the former integral can be expressed in a closed form could be useful for the numerical evaluation of the field of the boundary diffraction wave. If the observation point is in the vicinity of the edge of the geometrical shadow ($|K|$ is close to unity) it is worth evaluating the integral in Eqs. (12), (26), and (35) by using the mean value theorem of integral calculus.

For convergent spherical wave in the region of the geometrical shadow close to the focal plane $|K|\gg 1$ and $1/|d-z|\gg 1$, so Eq. (35), although valid at all points, is not convenient for computations. It is worth writing Eq. (35) into a new form. If we introduce a new function $\tilde{g}(K,L,\psi)$ instead of $g(K,L,\psi)$ given by

$$\tilde{g}(K,L,\psi)=\frac{\cos\psi-K}{1+K^2-2K\cos\psi+L^2\sin^2\psi}, \tag{A4}$$

then

$$\frac{g(K,L,\psi)}{a(d-z)}=\frac{\tilde{g}(1/K,L,\psi)}{rd}, \tag{A5}$$

and so from Eq. (35) the field of the boundary wave pulse is given by

$$u_B(r,z,t)=\frac{u_0 a}{2\pi rd}\int_0^{\pi} h\left(t-\frac{s(\psi)-f}{c}\right)\times\left(1+\frac{dz+a^2-ar\cos\psi}{s(\psi)f}\right)\tilde{g}(1/K,L,\psi)\,d\psi \tag{A6a}$$

$$=\frac{u_0 a e^{i(\omega_0 t+k_0 f)}}{2\pi rd}\int_0^{\pi} v\left(t-\frac{s(\psi)-f}{c}\right)\times e^{-ik_0 s(\psi)}\left(1+\frac{dz+a^2-ar\cos\psi}{s(\psi)f}\right)\tilde{g}(1/K,L,\psi)\,d\psi. \tag{A6b}$$

[1] T. Young, Philos. Trans. R. Soc. London **92**, 26 (1802).
[2] A. Rubinowicz, Nature (London) **180**, 160 (1957).
[3] A. Rubinowicz, Prog. Opt. **4**, 201 (1965).
[4] M. Born and E. Wolf, *Principles of Optics*, 6th (corrected) ed. (Pergamon Press, Oxford, 1987), Chaps. 8.8 and 8.9.
[5] K. Miyamoto and E. Wolf, J. Opt. Soc. Am. **52**, 615 (1962).
[6] K. Miyamoto and E. Wolf, J. Opt. Soc. Am. **52**, 626 (1962).
[7] Zs. Bor and Z. L. Horváth, Opt. Commun. **94**, 249 (1992).
[8] Z. L. Horváth and Zs. Bor, Opt. Commun. **100**, 6 (1993).
[9] Z. L. Horváth and Zs. Bor, Opt. Commun. **108**, 333 (1994).
[10] Z. L. Horváth, Ph.D. thesis, JATE University, 1997 (unpublished).
[11] Z. L. Horváth and Zs. Bor, Phys. Rev. E **60**, 2337 (1999).
[12] A. Rubinowicz, Phys. Rev. **54**, 931 (1938).
[13] Y. Li and E. Wolf, J. Opt. Soc. Am. A **1**, 801 (1984).
[14] J. Lu and J. F. Greenleaf, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 19 (1992).
[15] J. Fagerholm, A. T. Friberg, J. Huttunen, D. P. Morgan, and M. M. Salomaa, Phys. Rev. E **54**, 4347 (1996).
[16] P. Saari and K. Reivelt, Phys. Rev. Lett. **79**, 4135 (1996).
[17] J. Durnin, J. Opt. Soc. Am. A **4**, 651 (1986).